13

Language Diversity as a Resource for Understanding Cultural Evolution

Nicholas Evans

Abstract

The study of linguistic diversity, and the factors driving change between language states, in different sociocultural contexts, arguably provides the best arena of human culture for the application of evolutionary approaches, as Darwin realized. After a long period in which this potential has been neglected, the scene is now set for a new reconnection of evolutionary approaches to the astonishingly diverse range of languages around the world, many on the verge of extinction without trace.

This chapter outlines the various ways coevolutionary models can be applied to language change, and surveys the many ways diversity manifests itself both in language structure and in the organization of diversity beyond the language unit. Problems of establishing comparability and characterizing the full dimensions of the design space are discussed, including the distribution of characters across it, the correlations between them, and the challenge of establishing diachronic typologies (i.e., establishing the likelihood of different types of transition, including the insights that could be reached through properly focused studies of micro-variation). It concludes by surveying the main types of selection that mold the emergence of linguistic diversity—psychological/ physiological, system/semiotic, and genetic/ epidemiological—and spells out seven major challenges that confront further studies of linguistic diversity within an evolutionary framework.

Introduction

Languages occupy a central role in studies of human cultural diversity, whether viewed through the prism of social, cultural, historical, or psychological variability. With something between 6,000 and 7,000 distinct languages spoken today, they offer a kaleidoscope of largely independent natural experiments in evolving complex cultural systems without formal planning. This has led to

widely diverse outcomes both in terms of linguistic structures themselves and their articulation with culture and society.

The complexity of linguistic systems, their amenability to rigorous modeling, the assessability of these models against external data,[1] and increasing sophistication and commitment to gathering matched data on all of the world's languages has allowed linguists to accumulate a vast body of information on patterned diversity. With current moves to recast linguistics away from universalizing accounts that relegate diversity to a bit part, and toward coevolutionary models which assign it the central role, the time is now ripe to reconnect linguistics with evolutionary theory in a range of ways.

Linguistics, as the discipline most centrally occupied with language, has had an ambivalent relationship with evolutionary biology, and now with studies of cultural evolution. On one hand, at various moments in its history, linguistics has been particularly interfertile with evolutionary biology (going back to Darwin) and the development of rigorous modeling of evolutionary processes (Atkinson and Gray 2005). The historical linguist Brugman's 1884 distinction between shared retentions and shared innovations in language, for example, anticipated Willi Hennig's comparable distinction between symplesiomorphies and synapomorphies in systematic biology by nearly 70 years (Hennig 1950). On the other hand, for half a century, throughout both the structuralist era of Saussure, Sapir, and Bloomfield and the Chomskyan generativist era which followed, these classical connections of linguistics to biology were allowed to languish. Gains in the elaboration of techniques for "describing each language on its own terms" (structuralists) and for modeling the open-ended nature of linguistic systems in mathematically tractable terms (generativists) were achieved at the expense of developing procedures for comparing large matched data sets that respect languages' individuality while still allowing meaningful comparability.

It is only with the growing sophistication of linguistic typology, now amassing large bodies of cross-linguistically comparable material, that comprehensive ontologies of linguistic design choices, and statistically well-grounded testing of the relations between them, and to nonlinguistic factors such as group size or population genetics, have become possible. Since Labov's pioneering work on linguistic variation (Labov 1972, 1994, 2001, 2010) and the development of techniques by functionalist linguists like Bybee (2007) for studying the impact of use on form, the field has begun to have a better understanding of what promotes micro-variation and how social factors impact upon it.

---

[1]   Different types of linguists will interpret this as testing the characterizations in a reference grammar (written in disciplined normal language rather than a formalism) or a dictionary against a corpus, or of testing formal syntactic models against speaker intuitions. Though lively debate rages around these differences, a (rare) feature that unites most linguists is the belief that, in principle, the highly complex models we build do permit objective validation of some sort; linguistics has not gone down the postmodern road of seeing these models as simply subjective and untestable.

Coevolutionary Models of Language Change

Before plunging into an examination of language diversity, it is worth sketching a basic coevolutionary model of the language change which engenders it (cf. Evans 2003b; Evans and Levinson 2009a, b; Levinson and Evans 2010). First, languages are what Rudi Keller (1994, 1998) calls "objects of the third kind." They are neither straightforward results of biological evolution, like eyes or kangaroo hops, nor the intentional results of conscious human planning, like suspension bridges or constitutions. Rather, like equal-length supermarket queues or paths worn straight to the front door, they are the "unintended outcomes of intentional behavior," where the intention might be to get through the supermarket checkout or to the front door as quickly as possible, or (in the case of language) to persuade, deceive, sound like or unlike a particular person, or avoid homophony with an obscene sound-alike word. Second, the notion of coevolution applies here at a number of levels, of which I mention the three most important here.

Biology–Culture Coevolution

The biology–culture interaction is familiar in studies of human evolution (Boyd and Richerson 1985; Durham 1991). At the most gross level, this refers to the spiraling interactant effects between the hardware of the human brain plus its physical input-output system (e.g., larynx, tongue, ear) and the cultural software of language. Physiological changes created affordances which supported more complex language, while the cultural evolution of linguistic systems in its turn placed increasing demands for complex cognitive skills and a streamlined input-output system. Changes to the vocal tract that allow more consonant distinctions to be articulated permit the evolution of more complex consonantal phonologies, which favor individuals whose vocal tracts and perceptual acuity can best produce and discriminate the sophisticated new phonologies. In a more interesting guise, to which the eld is just returning after many years of dogma-driven neglect, the possibility is opening up that different genes favor different linguistic structures (e.g., pitch perception and tone, Dediu and Ladd 2007) and that some of the distribution in language structures over the design space can be sourced to genetic differences in the populations that have shaped them through history. I will discuss this more later (see section on Mechanisms of Selection), and it is also addressed by Levinson and Dediu (this volume).

Social–Psychological Coevolution

Languages lead a double life, as social institutions and individual representations/dispositions. This gives a double set of mechanisms by which change can occur:

1. At the individual level: imperfect learning or generalization of new forms, streamlining of pronunciation leading to phonological reduction, modi cation of accent accompanying social reaf liation.
2. At the societal level: norm resettings weed out or repurpose some variants through such social mechanisms as stigmatization, revaluation (e.g., the recentering of crucial vowel norms on the mercantile class in Shakespeare's time), or identi cation of variants with a particular social group.

Normally, individuals construct their own linguistic representations from a complex society of people around them. To be sure, some caregivers have disproportionate input in early childhood, but as life proceeds, the set of sources expands—try the thought experiment of going through your own vocabulary, including variants of pronunciation or meaning, and thinking the source from which you learned each item. Societal norms about language, to a large extent, are "out there" in some generalized sense, but there are important subsets of society (e.g., parents, other kindred, neighborhood, clan, professional networks) which may be disproportionately important.

During the structuralist-generativist period, dominant theories focused on just one or the other of these two loci. For Saussure, the most tractable object of study was *langue*, a set of social conventions which he ctively conceived of as a set of identical dictionaries deposited in the brains of all speakers. For Chomsky, the true object of linguistic study was (*individual*) *competence*, the knowledge of an idealized speaker-hearer in a homogeneous speech environment. Each was an abstracting move, with Saussure sidelining *parole* and Chomsky sidelining *performance*, both deemed too chaotic to be analytically tractable.

By focusing on just one linguistic locus, both approaches disfavored the adoption of a coevolutionary approach. More recent work, however, is leading to a view that synthesizes both of these loci, using the sorts of variationist methods developed by Labov to nd systematicity in social variation as well as functionalist methods to nd systematicity in individual variation in production. The interaction of evolution at these two sites can then amplify small selection biases through evolutionary funneling effects. Christiansen and Chater (2008:507) argue that "language has evolved to be learnable": this works through "C-induction" (Chater and Christiansen 2010) by which "cultural transmission delivers the restricted search space needed to enable language learning, not by constraining the form language takes on an innate basis, but by ensuring that the form in which language is presented to the learner is learnable" (Merker 2009:461).

Culture–Language–Cognition Coevolution

A third bidirectional conception of coevolution is that (a) cultural design choices are an important selector for language (typically over a scale of centuries or

millennia), and that (b) linguistic design choices are an important selector for cognition (over the scale of the lifetime or shorter). The latter phenomenon, whereby language is argued to have some inﬂuence on thought, is typically labeled neo-Whorﬁanism (or the Sapir–Whorf hypothesis) after two of its major mid-C20 proponents. It is useful to distinguish the former as Vico–Herder effects, after two early romantic proponents of the view that languages express unique aspects of the history and worldview of their cultures (or peoples, or nations). Though the two effects are frequently conﬂated loosely under the "Sapir–Whorf" rubric, they operate on very different timescales and require very different evaluation methods, which makes it handy to have different labels for the two types of effect.[2]

It has taken decades to develop suitably operationalized domains of language and experimental paradigms to evaluate the Sapir–Whorf hypothesis, and it remains a topic of ﬁerce debate. Proponents of its "weak form" (i.e., choices in language structure have some impact on habitual thought) have pointed to effects of grammar or lexical structure on such issues as color perception, preferred mode of spatial orientation (absolute vs. egocentric/body-centered), or categorization and representation of information about motion events. This is not simply "thinking for speaking" (Slobin 2003), but more generally "scanning and coding for speaking," given that particular grammars may require their speakers to report on particular aspects of reality (say, absolute compass orientation) at any point in the future, so that attention for future use becomes paramount, and this in turn inﬂuences how memories are coded. Methodologically, the focus in neo-Whorﬁan research has been on the effects of language on individual cognition (see the collection of studies in Gentner and Goldwin-Meadow 2003), whether in adult processing or child development. There have also been interesting studies on the ways these are summed across the society of language speakers to produce statistical effects of a language variable (say, grammatical gender) on a social variable (e.g., female workforce participation, Mavisakalyan 2011).

For Vico–Herder effects, two main variants can be distinguished:

1. "Ethnosyntactic" phenomena, where some "cultural preoccupation" ends up shaping the grammar in some way (for examples, see section on Social and Cultural Selection). Appealing case-study type examples abound, but it is fair to say that the problems of evaluating causal effects in this domain have not been overcome; since they are correlation based across large numbers of cultures, they rest on solving coding problems (see section on The Problem of Comparability) in both language and culture.

---

2   Indeed, Whorf explicitly denied any causal link of culture on language: "The idea of 'correlation' between language and culture, in the generally accepted sense of correlation, is certainly a mistaken one" (Whorf 1956:139).

2. Social-scale phenomena, where such issues as size of speech community and number of speakers who have acquired the language late in life are argued to impact on factors like the degree of grammatical elaboration and complexity. Once again, coevolutionary accounts of this type are a new game, and the eld has yet to develop rigorous methods for testing these interesting claims (see discussion in section on Social and Cultural Selection).

These three nested levels of coevolutionary modeling in linguistics give some indication of the wide range of factors which may give rise to the diversity found across the world's languages. We now turn to the languages themselves, with an eye to indicating the many dimensions on which language diversity can be identi ed.

Diversity in Language Structure

All languages are made up of a large set of signs: conventionalized associations of form, meaning, and combinatorics. Signs are usefully partitioned into lexical signs (basically words that would go into a dictionary) and grammatical signs (like af xes, or syntactic facts like word order), as well as less straightforward sign types like intonation and other forms of prosody. The fact that English builds noun phrases in the order demonstrative—adjective—noun ($this_1$ big_2 $book_3$) whereas Indonesian does the reverse ($buku_3$ $besar_2$ ini_1) provides one example of a grammatical sign (signaling how units within a noun phrase, NP, will be assembled and have functions like "modi er" or "determiner" assigned).

For all signs, each of their three dimensions is conventional and largely arbitrary. Consider the English word "know." As Saussure emphasized (using other examples), in a tradition going back to the Greeks, it is largely arbitrary what form is employed to designate concepts, as can be illustrated by nding the (rough) equivalents in other languages:[3]

- Russian знать *znat'* (cognate with English *know* but highly divergent in form)
- Japanese 知る *shiru*
- Tamil தெரி• *teri*
- Kayardild *mungurru*
- Dalabon *bengkan*

---

3   Premodern linguistics was obsessed with written language to the exclusion of the spoken; much modern linguistics has erred on the other extreme, treating writing systems simply as transcription methods. A more useful approach (and interesting from the point of view of incorporating writing systems as cultural innovations potentially impacting on other aspects of language) is to treat sound and written forms as partially independent but partially linked signifying systems, something I emphasize here by giving the Russian, Japanese, and Tamil words in their respective orthographies as well as in a Romanized transcriptional system.

Despite this traditional view of the "arbitrariness of the sign," evidence re-claiming significant levels of non-arbitrariness has been accumulating from a variety of quarters. Onomatopoeic words (e.g., bird names resembling bird calls) are one type of case. A second (see Haiman 1980, 1983) is "grammatical iconicity": if there are two types of possession, e.g., "inalienable" like "my eye" vs. "alienable" like "my yam," the semantically "closer" inalienable type will use a more "direct" syntactic construction, as in the Paamese contrast between *mete-ku* [eye-my] "my eye" vs. *auh aa-k* [yam possessed.edible-my] "my yam" (Crowley 1982). A third is "diagrammatical iconicity" by which the ordering of affixes normally reflects their logical scope.[4] A fourth is the fact that in many languages phonological cues within a word provide probabilistic cues to word-class, helping the child learn the combinatorics of new words on the basis of their forms (Monaghan et al. 2007, 2011). In Japanese, for example, all verbs end in *-u* (and a high proportion in *-ru*) whereas for nouns this proportion is much smaller; thus the fact that the Japanese verb for "(come to) know" is *shiru* rather than *shika*, is not arbitrary. A more balanced view, therefore, is to see linguistic forms as exhibiting a mixture of arbitrariness and nonarbitrariness in their forms across the whole vocabulary and grammar.

The meaning of signs is likewise highly variable; we are not just dealing with pinning different labels on the same conceptual object. To translate English *know* into Russian we need to distinguish (roughly) factual knowledge (*znat'*), procedural knowledge (*umet'*), and acquaintanceship (*byt' znakom s...*). Japanese *shiru* is better translated as "come to know," with "know" a consequence of certain aspectual operators (completed transition) on the basic verb meaning. In Dalabon *bengkan*, even though it is the closest to an equivalent, is better translated as "have in mind on a long-term basis" (Evans 2007); it can include long-term remembering, belief, and lengthy contemplation, and unlike "know" only implicates rather than indefeasibly entails the truth of its cognitive content. Thus, unlike English *know*, you can follow *bengkan* with a word meaning something like "believedly" to indicate skepticism with regard to the accuracy of the mental state.

Finally, the combinatorics can be rather different. Though the corresponding words in some of the above languages (Russian, Japanese, Dalabon) are like English in being transitive verbs, other languages do things differently. The Tamil equivalent *teri* is also a verb, but assigns quite a different case pattern, with the knower in the dative case (literally "to Kumar this place knows" for "Kumar knows this place"). Dalabon *bengkan* is a transitive verb, and its

---

4   As an example, consider Kayardild *karndi-wala-nurru* [wife-many-having] "polygynous" vs. *karndi-nurru-wala* [wife-having-many] "many married (men)." The semantic scope of *wala* and *nurru* seems "compositional" to English speakers even though the ordering is alien, as is the fact that these categories are expressed by suffixes rather than free words. However, diagrammatic scope is not found in all cases. For a discussion of one such case and the unusual evolutionary pathways which gave rise to it, see Evans (1995b).

syntax is not grossly different from English, but its morphology differs radically. Like other transitive verbs in Dalabon, it takes prefixes indexing both subject and object, so "(s)he knows him/her" is the single word *bûkahbengkan*, where *bûkah*- indicates "(s)he acts upon him/her." The verb can also accrue all sorts of modifying prefixes (e.g., *bûkahkakkûbengkan*, "(s)he really knows him"). Passing to languages which do not use verbs to encode this concept, in Kayardild the corresponding word, *mungurru*, is a predicative adjective (much like the English word "'knowledgeable"), though one still capable of taking an object, so that *ngada mungurru ngumbanji* "I know you" is more like "I [am] knowledgeable of you."

Combinatorics is fundamental to describing grammars of languages because the rules grammarians write to build words from their parts, or phrases and clauses from words, refer to classes of entity (like noun, verb, adjective, determiner, root, verb phrase, and so on). The first four of these (noun, verb, adjective, determiner) exemplify "parts of speech" or "word classes." These are sets of words united by common combinatorics, and these sets are indefinitely large, in the case of the "major word classes." It is their common combinatorics which allow us to formulate syntactic rules. To build an English noun phrase up as determiner + adjective + noun rather than the opposite order as in Indonesian, we need to know which lexical items go in which slots: "the" or "this" can fill the determiner slot, "big" or "red" the adjective slot, and "man'" or "building" the noun slot.

An important dimension of variation across languages lies in the patterning of word classes across languages. It is well established that there is considerable variation here. English, for example, lacks developed word classes that correspond to such classes as expressive or mimetics in languages like Semelai or Japanese, or co-verbs in Jaminjung, whereas there are many languages that do without prepositions (e.g., Kayardild) or adjectives (e.g., Lao). Whether this variability extends to the most fundamental distinction of all—that between nouns and verbs—is a topic of continuing debate. For a language like Straits Salish, Jelinek (1995) has argued that there is just one open class of predicates (as in predicate calculus), with meanings like "eat," "be a man," "be Eloise." On this analysis, in a language like Straits Salish "the men ate the fish'" would be rendered as "the (ones) who are.men ate the (ones) which are.fish."

This does not exhaust the ways signs can vary. Consider polysemy, where what looks like a single sign (same form, same combinatorics) has more than one meaning. Languages differ again here, in terms of which meanings they co-link to the same form. The polysemic ranges stretch from near universal at one end (e.g., association of "high'" with happy/active rather than sad/inactive) to highly idiosyncratic at the other. Consider the extension of "see'" to "understand," as in "I see," which is sufficiently widespread to have led Sweetser (1990) to argue this was based on universal metaphors of visual perception as cognition. Yet this particular polysemy is absent in Australian Aboriginal

languages, where it is "hear" rather than "'see" which is the modality under-lying extension into metaphors of understanding (Evans and Wilkins 2000). In many cases, gurative language draws heavily on rather speci c cultural scenarios or presumptions. The celebrated extension from " y" to "steal" in medieval French, retained in these two meanings of *voler*, was mediated by the speci c practice of falconry (*le faucon vole le perdrix*, "the falcon ' ies' the partridge") (Benveniste 1966).

Returning to the form of signs, and now focusing on the syntagmatic[5] dimension (i.e., the dimension of how things combine in sequence), we see once again that what appears super cially to be the same form can result from quite different types of structure. Consider the situation where three languages, A, B and C, all have a sound sequence *kákì*, where ´ represents a high pitch and ` a low pitch (thus a pitch contour like [¯ _], and where this contrasts with a word *kàkí* [_ ¯].[6]

We could class all three languages as tone languages, on the grounds that all use pitch to discriminate meaning. However, lurking under this apparent similarity we nd that pitch is organized in a very different way. If we expand the number of syllables we are looking at, we nd that for A, the number of different pitch combinations is 2^S, where S is the number of syllables; this is what we would expect if we could make an independent two-way choice on each syllable. This is more or less the situation that is found in "classical" tone languages like Mandarin or Vietnamese, except that these have more tone contrasts. Such languages are sometimes called "syllable tone" languages. A second possibility would be that, however many syllables there are in the word, there are just two "melodies"—a rising one and a falling one—which can get squeezed onto just one syllable or stretched out over a long word. On a single syllable, the contrast would now show up as rising versus falling tone as the melody gets compressed. In this case, tonal phonologists generally talk of "word tone" languages.

A third possibility is that the number of contrasts is linearly related to the number of syllables, as s or $s + 1$; the latter is more or less the case for Tokyo

---

5   This is usually contrasted with the "paradigmatic" dimension, or the dimension of opposition as opposed to the "dimension of combination" which is the syntagmatic dimension. We can illustrate the difference with the *p* in the word *sprite*. It is a syntagmatic fact that *p* can be preceded only by the sound *s* (i.e., *hprite*, *mprite*, etc. are impossible in English), and a paradigmatic fact that it could (in hypothetical but perfectly pronounceable words) be substituted by *k* or *t* (*skrite*, *strite*) but not by *m* or *n* (*smrite*, *snrite*). We can transfer these concepts directly to many other informational systems, such as DNA; for example, the fact that any one site is part of a four-way opposition between bases is paradigmatic, while the groupings of bases into longer sequences of various sizes is syntagmatic.

6   In Japanese the rst (meaning "oyster") would be written 蛎, whereas the second could mean either "fence" or "persimmon," and would be written 垣 and 柿, respectively: the character-based writing system distinguishes them totally. It is also worth pointing out that the latter two can, with more phonological nesse, be distinguished, according to whether the high tone is continued onto following postpositions like the topic marker.

Japanese, except that words more complex than the ones shown here would demonstrate that instead of syllables, *S* should be counting slightly smaller units, known as morae. A consequence is that in Tokyo Japanese, just one place has to be marked as the "in ection point," which is the point at which pitch falls.[7]

These possibilities are summarized in Table 13.1. We could think of them as three sets of rules negotiated between a composer and a librettist in terms of how they line up their contributions as they collaborate on an opera. In the rst, the composer is allowed to put any note to any syllable. In the second, the composer hands a tune to the librettist and instructs her to stretch it out, once per verse, for verses of any length. In the third, the composer writes tone rows (here, of just two tones) but gives the librettist choice about where in the lyrics she may move from one tone to the next.

I have chosen an example from the realm of tone because it sidesteps some of the gridlocked debates that have led to rather unproductive standoffs in the realm of syntax. What it should illustrate very clearly is that diversity in how languages are organized (in this case, in how they harness melody as part of meaning-signaling form and link it up to segmental elements) may not be immediately obvious, and that bene t is drawn from having abstract representational mechanisms able to capture different deeper patterns operating in what may seem to be the same form. Conversely, more abstract representations may show similar patterns lying under different forms, as when a monosyllabic falling tone and a disyllable with a high-low pattern are shown to both be instantiations of a falling melody in a word tone language. In other words, the use of abstract representations is neutral in terms of increasing or decreasing the level of variation we postulate as underlying the system.

The role of abstract representations of one form or another has been a major problematic in debates about linguistic diversity over the last half a century. There has been a general tendency for those from the generativist tradition to use more abstract representations and those from the descriptive and typological traditions to use more concrete ones. Both are clearly necessary to do justice to many linguistic phenomena, but they also introduce a dangerous possibility of glossing over signi cant diversity by viewing the "real" phenomena as invariant at some underlying level. Even where they are justi ed (far from a simple point to determine), their use can still cause two major problems for comparative work: (a) since abstract representations often depend on more sophisticated analysis of the data, there will be fewer data points in terms of languages for whose structures we can vouch at this abstract level; (b) if the abstract representations exhibit elements that are speci c to particular languages, this also makes comparison harder.

---

[7] The rising pattern is produced because of a rule that lowers the rst syllable just in case the
in ection point does not produce a low tone on the second syllable.

Table 13.1 Types of linear patterning of pitch choices in three types of language. N(S) is a function from the number of syllables N to the number of word-tone contrasts maximally available for an N-syllabled word.

Hypothetical Language	Form	Process	Representation	N(S)	Type Language
A	*kàkí* [_ ¬]	Concatenation of contrasting syllable tones	ka[_] + ki[¬]	Power of syllables: 2^s	"Classical" tone languages (e.g., Mandarin, Vietnamese)
B	*kàkí* [_ ¬]	Association of melody with whole word	kaki + [_ ¬]	Constant: 2	Word tone languages (e.g., Mian)
C	*kàkí* [_ ¬]	Change in pitch level at in ection point	ka′ki	Linear: S(+1)	Pitch accent languages (e.g., Japanese)

Further Dimensions of Diversity: Beyond the Language Unit

In the previous section I focused on structural diversity, assuming that each language is relatively standardized and internally variant, and then compared given structures for each such language in terms of form, meaning, or combinatorics. Three other types of diversity should not be neglected and are thus brie y discussed here.

Sociolinguistic Variability within Speech Communities

This concerns the way variation is organized within the unit rather imprecisely designated as the speech community: a unit which according to the situation may be smaller or larger than the units we conventionally and uncritically designate as languages.

This kind of variation is fundamental. It is vital to establish, empirically, how robust linguistic systems are against communicative degradation when norms are not shared, an area where modeling can provide vital insights (Hruschka et al. 2009; McElreath et al. 2003). Experience with our own languages makes it clear that Chomskyan levels of idealized homogeneity are not necessary to assure (largely) successful communication, but how far can they drop without ceasing to function as an ef cient shared code? Are there different patterns of internal diversity by domain (grammar vs. lexicon vs. phonology) or by medium (speech, writing, sign vs. spoken language)?

In fact we nd that variation is often less a matter of degraded signal than an additional semiotic layer employed for social-signaling purposes, identifying the speaker's regional af liation, class, caste, religion and so on, as well as aspects of the communicative setting. This may, in turn, have adaptive functions in terms of agging membership in communities with the same norms

of cooperation and coordination (Richerson and Boyd 2010). Once linguistic variability is deployed in the same way, it creates a more complex set of shared norms: not just a <form : meaning : combinatorics> triple, but a link from each of these to some social information. This can drive linguistic change in particular ways. For example, historical linguists have recently become aware of processes like "correspondence mimicry," where speakers are aware of proportionalities between languages and use these to refashion foreign adoptions in their own language to make them look more native-like (Alpher and Nash 1999). Likewise, an increasing number of otherwise inexplicable historical changes are cropping up, for which the best explanation is that a highly unusual variant got promoted in some variety precisely to signal that speech community's distinctness from a neighboring one, in the grammatical equivalent of the *you say tomato, I say tomahto* principle (for an Iwaidja example, see Evans 1998). Such changes can only occur against a background of shared knowledge of both languages.

Not all speech communities are equally diverse internally. Some (e.g., precontact Kayardild) were exceedingly homogeneous, whereas others are highly diverse (e.g., the "dialect chains" of Western Desert in Australia or Numic in the Great Basin). It is sometimes proposed—thus far without clear comparative evidence—that such dialect chains are found among hunter-gatherers in desert regions where the vagaries of rainfall and yield constantly drive small bands of desert-dwellers to recoalesce with others living in areas that got a good recent rainfall. Having a shared grammar then makes it easy to learn whatever new code one needs to blend into the environment (cf. Shaul 1986). We can get situations where there is effectively a common grammar that gets localized by differences in vocabulary (e.g., parts of Northern Vanuatu, or the Western Desert Chain mentioned above). Conversely there are situations with such high levels of shared vocabulary that mutual comprehension is assured despite the grammars being organized in signi cantly different ways (this was the situation between Kayardild and Yukulta; Evans 1995a).

Once writing is added as medium, of course, we can have speech communities of great diversity in speech (effectively many different languages) united by a common writing system; this is the case for Chinese. The converse case involves a single language (perhaps spread across dialects) divided into two or more "political languages" with attendant different literatures, through the use of different writing systems, as between Serbian (written in Cyrillic) and Croatian (written in Latin script). The Hindi–Urdu case is similar though complicated by greater divergence in learned vocabulary, with Hindi stocking from Sanskrit and Urdu from Persian.

On top of this, there are situations where more than one language is deployed inside a speech community for designated roles—so-called diglossia. These languages may be related as a more archaic/classical and modern/vernacular version, such as classical Arabic and local vernaculars throughout the Arab world, or may be unrelated, as is the case between English and many

other languages today (e.g., Swahili, Malay) or, traditionally, between Latin and European vernaculars or Classical Chinese and other languages of the Sinosphere (e.g., Korean, Japanese, Vietnamese).

Responding to diglossic situations simply by treating them as two separate languages misses key insights into how language spans its communicative goals as well as how linguistic systems interact. In particular cases of bilingualism, such as the traditional French–Cree bilingualism of Métis trappers in the Canadian prairies, the severing of ties between the mixed speech community and its two reference groups can produce a peculiar outcome: a so-called mixed language like Michif. In the Métis–Michif case, this happened once racial rede nitions cut off people of mixed descent from both groups. Mixed languages intimately combine features of two languages that would normally not be transmitted by borrowing or contact, in what Bakker (1997) has called language intertwining. It can happen that speakers of mixed languages sometimes know neither of the two source languages, i.e., a group of Michif speakers developed who, unlike their forebears, knew neither French nor Cree but preserved the complexities of both languages in a new mixed code.

A key tenet of linguistics since the structural era has been Meillet's dictum that "language is a system where everything hangs together."[8] This formulation suggests that certain hypothetical points of the design space (characterized by combining elements which would not "hang together") are unpopulated because any system combining them would somehow be dysfunctional. Now that we know more about mixed systems, which often combine elements in highly unexpected ways, an alternative explanation suggests itself; namely that co-transmission of traits is what produces any "Meillet effects" that may be out there, and that once particular social situations are present to engender mixed languages, then we may see the co-occurrence of unexpected traits. At a period when there is growing evidence that traits hitherto believed to be tightly coupled can in fact exhibit more independent evolutionary trajectories (Dunn et al. 2011b), the need to examine the impact of given social deployments of linguistic variation takes on a new signi cance.

Also highly variable is the way given speech communities harness variation to signal social signi cance. English speakers are accustomed to variation (phonological, lexical, grammatical) being used to signal regional origins, as with the different pronunciations of *butter* as [bɛtʰə], [bɛdəɹ], [bɛtʰəɹ], [bɛɾə], [bɛʔə], [bʊʔəɹ], and so forth. We also know that speakers slide between more conscious, educated upstyle variants (e.g., *aren't you*) and more casual ones (e.g., *arntcha* or *arntchas*[9]) according to factors like formality. Since Labov (1966), we also know that patterning can emerge across the whole speech community even though no individual has knowledge of the whole

---

[8] Famously formulated by Meillet (1906) as *une langue est un système où tout se tient*.

[9] In those parts of Australia and elsewhere in the English-speaking world that use *youse* (~ [jəz]) as a (typically low-style) plural second person pronoun.

pattern. However, this sort of patterning only scratches the surface of what languages can do with variation. For Tamil (Andronov 1962), for example, there are choices between classical/literary and colloquial, regional colloquial varieties, and caste varieties (e.g., the variety used by Brahmins) as well as a range of situationally based choices for using honori cs depending on the speaker-hearer or speaker-referent relationships. For Javanese (Errington 1988), there are three main registers—*krama*, *madya*, and *ngoko* (in order of descending "re nement")—with complex conventions for who can use what to whom, taking into account the social position of both speaker and hearer. These three registers are so different that a given sentence may differ in every word between them (e.g., Krama *Menapa nandalem mundhut sekul semanten* vs. Ngoko *Apa kowé njupuk sega semono* for "did you take that much rice?").

For the Australian language Bininj Gun-wok (Evans 2003a), there is an everyday or default variety known as *gunwokduninj* "real language" (though itself divided into half a dozen main dialects); a special variety known as *gun-gurrng* "in-law language" used with certain high-respect af nes like one's wife's mother's brother; an obscene joking variety used between potential (but unactualized) af nes; a series of clan lects (with shibboleth terms down to a much ner grain than the main dialects) used to establish clan identity at the beginning of encounters as well as in addressing ancestral clan sites; and many special lexical items found just in poetic or song language. As with Javanese, there may be total lexical differentiation between registers (though grammatical af xes are typically unchanged): "Have you got any meat? No, nothing" would be *Gun-gunj yigarrme? Gayakki* in *gun-wokduninj* but *Gunmulbui yiwalebonghme? Gayagura* in the decorously multisyllabic *gun-gurrng*.

Multivarietal "lect clouds" may exhibit suf cient systematicity to suggest systems of shared norms that transcend what linguists take as the unit of description when writing grammars of a single variety. In Northeastern Arnhem Land, for example (Wilkinson 1991), a substantial part of the variation is factored across two orthogonal dimensions:

1. a geographical one, running broadly west to east and shown by the optional loss of initial ŋa from pronouns in the western area, and
2. a social one based on the assignment of every clan and language variety to one of two patrilineally transmitted moieties.

Languages spoken by clans of the Yirritja moiety have vowel- nal phonologies (like Italian) whereas those of the Dhuwa moiety have dropped most nal vowels giving a more staccato phonology (like Catalan). The intersection of these two factors (see Table 13.2) creates a four-way matrix in which most stretches of speech in a number of languages in this region can be rapidly located in social and geographical space. This attests to semiotic rules shared across a multilingual region that transcends the boundaries of the single languages (like Djapu or Djamparrpuyngu) which form the normal units for linguistic description.

Table 13.2   Geographic and social patterning of the rst-person plural pronoun form in some Yolngu dialects (after Wilkinson 1991:187).

Geographical [initial ŋa-drop]	Social (patrimoiety) [Final vowel drop]			
	Yirritja moiety (Dhuwala varieties) ŋa		*Dhuwa moiety* (Dhuwal varieties)	
Western	Gupapuyngu:	[ŋa]napuru	Djamparrpuyngu:	[ŋa]napur
Eastern	Gumatj:	ŋanapuru	Djapu:	ŋanapur

A still underexamined area, particularly important for theories in which gesture played at least some part in the original evolution of language, is the semiotic partition between spoken and signed or gestural codes. In Dhuwal (Northeastern Arnhem Land), for example, the conventionalized touching of symbolic body parts instead of verbal reference is a regular practice for some respected types of af ne, so that a man might ask "where is (your) wife's mother?" by saying the equivalent of "where is that kind," slapping his knee while saying "that kind" (Heath 1982:255). Alternatively, gesture may be harnessed to clarify thematic roles like instrument. Another Australian language, Iwaidja, is unusual for Australian languages in not having overt marking for instrumental case, but I have recorded Iwaidja speakers saying things like "I went.out bark.torch" while raising their hand in a gesture of holding a torch concurrently with the word for it. In this case, the gesture marks the thematic role in a way that would be carried out by the grammatical device of case-marking in most other Australian languages. There has been so little work on the integration of language and gesture (or spoken and sign language in communities that have both) that we cannot currently compare the functions and degrees of integration that these two modalities have, but it is an important dimension of variation for future study.

Until recently, linguists have kept the study of variation in the various guises outlined above as a distinct domain from the study of structure. Though this decision had temporary heuristic value, bringing together the study of sociolinguistic variation and the study of structural diversity is likely to yield many interesting ndings in the coming decades.

Phylogenetic and Typological Diversity

Zooming out from our dialect maps to the level of language families and subgroups, we encounter another level of diversity. Once again, this has only been conceived as a phenomenon in need of explanation in the last two decades or so, prompted by Johanna Nichols' seminal book *Linguistic Diversity in Space and Time* (Nichols 1992) and, more recently, books like Daniel Nettle's (1999) *Linguistic Diversity* and the large compilation of data assembled in the

WALS[10] survey. Work like this has drawn attention to striking disparities in the worldwide distribution of diversity, be it phylogenetic (i.e., in terms of numbers of families, subfamilies etc.) or typological (i.e., in terms of the amount of gross typological variation).

Some of these can be attributed to colonial expansion erasing ancient diversity. This may be modern, as in the spread of English or Spanish—the latter case eliminated all indigenous languages (e.g., Cuba and Uruguay). However, it may also be ancient, as in the spread of Latin and its descendants with the attendant elimination of Etruscan, Umbrian, Gaulish, and many other languages, or the spread of Austronesian languages through islands in southeast Asia (which eliminated, e.g., all non-Austronesian languages from the Philippines and Western Indonesia). There has been a tendency by some authors (e.g., Renfrew and Bellwood 2002) to see "demic expansion," typically agriculture driven, as the only cause of widespread language families; that is, the higher populations of farmers simply squeezed out the hunter-gatherers demographically, so that the farmers' language ended up displacing that of the aboriginal hunter-gatherers. However, there are many large language families in the world (e.g., the earlier levels of Niger-Congo, Pama-Nyungan in Australia) where we currently lack a plausible explanation in terms of agriculture-driven demic expansion. This is one type of phenomenon, then, where linguistic facts (distribution of language families) abut scenarios of cultural evolution more generally, leading to multidisciplinary accounts of prehistory that range from canonical and convincing (e.g., the Austronesian expansion) to still enigmatic (Pama-Nyungan, Trans-New Guinea).

Even if we strip away the effects of colonization and expansion and its presumed steamrollering of prior phylogenetic diversity, we are left with major distributional puzzles. In terms of its number of maximal clades (i.e., phylogenetic groupings, including isolates not currently relatable to any other), New Guinea and its immediate surrounds comes in with around 35—a number greater than that found for the whole of Eurasia. This cannot simply reﬂect New Guinea's long-standing isolation from the centralizing effects of larger states. Australia, just next door, joined to New Guinea for most of its human history, and with comparable or greater levels of linguistic diversity in terms of number of speakers per language, is covered by what increasingly looks like a single language family, as more and more evidence accumulates (Evans 2003c). The differing levels of diversity in Australia and New Guinea (and comparable asymmetries elsewhere) remain one of the great unanswered puzzles of linguistics. Its solution will most likely need to bring in both (a) proximal examinations of degrees of sociolinguistic dispersal in traditional speech communities of different kinds and (b) distal research on what social conﬁgurations accelerate or retard microdiversiﬁcation within a speech community, and what determines how variants get invested with social meaning to allow

---

[10] WALS = World Atlas of Linguistic Structures, now available online: http://wals.info/

them to survive as transmitted signs rather than being weeded out as mistakes or stigmatized forms.

The worldwide distribution of typological diversity is likewise shaping up as a key area of research, replete with currently puzzling data. To illustrate the distinction between typological and phylogenetic diversity consider the following: both Vanuatu and New Caledonia contain low-level branches of Oceanic within Austronesian, and in fact Vanuatu is somewhat more diverse phylogenetically since it is normally held to contain several low-level branches. Yet New Caledonia appears to be much more diverse typologically: it contains, for example, several tone languages and has much greater diversity in phoneme inventories than those found in Vanuatu. Once we recognize the independence of phylogenetic and typological diversity, two immediate challenges face us: How can we quantify the degree of typological dispersal in an observer-independent way that is comparable across all languages? What types of models can account for why there seems to be much greater typological diversification in some areas than others? Once again, we need to attend both to linguistic microlevels (measuring amounts of typological variation inside speech communities) and to types of social organization or other coevolutionary selectors which may afford or downplay typological diversification.

The Problem of Comparability

> Coding comparative cultural and linguistic data is a task fraught with the difficult reduction decisions familiar to all scientifically minded anthropologists.
> —Fiona Jordan (2013:47)

Finding ways of coding data in cross-linguistically comparable ways is a difficult but essential task if linguistics is to draw the full power from the vast data sets it is beginning to assemble. In the development of the field, both structuralist/descriptive and generative traditions took intellectual positions that hindered cross-linguistic comparison. Somewhat simplistically, we can say that structuralists were guilty of bongobongoism, exaggerating noncomparability, and generativists were guilty of procrusteanism, smoothing away cross-linguistic differences by using the great representational flexibility afforded by deep structure to surface mappings.

For the structuralists, both the Boas-Sapirian program (i.e., each language should be described on its own terms) and the Saussurean dictum (i.e., the significance of any sign was its place in the system) worked to favor descriptions which could be exquisitely finely tuned to linguistic particularity. However, their premises made cross-linguistic comparison impossible.

For the generativists, the quest to look below the surface to a more abstract "underlying structure" had several effects. One was to downplay surface variation as insignificant. The famous Binding Condition A (Chomsky

1981)—anaphors (unlike pronouns) should be "bound in their governing category"—predicts that there should not be languages with a single item *them* whose meaning would range across "disjunct" pronominals like "they saw them" to reflexives like "they saw themselves" and reciprocals like "they saw each other." This is because the first reading would not be "anaphorically bound within the clause," whereas the second and third would be. Put another way, languages should always have distinct anaphors for the bound and nonbound situations. Now when languages turned up which appeared to violate this (Old English, but also many others, such as Mwotlap and Tinrin in the Pacific, or Javanese), the response was either to postulate two distinct entities at some deeper level of analysis that happens to be homophonous on the surface, saving the putative universality of the Binding Condition.[11] Certainly there may be cases where more subtle investigation reveals justification for this, but to assume that such reanalyses of the data can be taken as the default (or to dismiss apparent counterexamples from less-described languages as likely to yield a deep-level distinction on further analysis) had the effect, within the generative program, of ignoring many populated regions of the design space.

This is not to deny the very real difficulties in comparing phenomena across languages. Suppose we are surveying the world's phoneme inventories and are deciding whether particular languages have a /p/ phoneme or not. First we compare English and French: both have a /p/ contrasting with a /b/, so we could decide affirmatively. However, if we are fussier about phonetic detail, we might decide that the English phoneme is really an aspirate /pʰ/ whereas the French is really a /p/. Then we bring in Hindi, which (leaving aside the voiced aspirate or murmured /bʰ/) has a three-way distinction between a /p/ close to the French sound, a /pʰ/ closer to the English one, and a /b/. Then we add Mandarin, which contrasts /pʰ/ and /p/ but lacks any real voiced sound,[12] and Korean, which contrasts an aspirate /pʰ/ with a lax unaspirated sound (revealingly, sometimes transliterated as *p* and sometimes as *b*) which alternates between these two according to word position. Finally we add Kayardild, which has just one bilabial stop phoneme realized variably as [b], [p] and [pʰ]; though the aspirated pronunciation is marginal enough to make this unappealing as

---

[11] An alternative response has been to split the phenomenon into such a large number of factors that it becomes difficult to gather sizeable cross-linguistic data (e.g., Reuland 2008). A third option, however, may be introduced, as Cole et al. (2008) do for Peranakan Javanese; namely, there are words which are unspecified for the anaphoric vs. pronominal contrast and hence compatible with both conditions, thus making it possible for them to conclude that Peranakan Javanese "does not…constitute evidence that there are languages in which the Binding Theory fails to apply. Indeed, Peranakan Javanese provides compelling evidence that the Binding Theory is active in languages containing forms that appear to be exempt from the Binding Theory" (Cole et al. 2008:585).

[12] Pinyin, the system now most widely used for Romanizing Mandarin, writes p for /pʰ/ and b for /p/. However, earlier systems of Romanization, such as the Wade–Giles system, used a different method, writing these sounds respectively as *p'* and *p*.

the candidate for naming the phoneme, the grounds for choosing between the other two are pretty arbitrary. This gives us the situation shown in Table 13.3.[13]

Depending on how we operationalize the question, we could answer that:

- all six languages have a /p/ (since all have phonemes including this in their allophonic range),
- just one language (Hindi) has a /p/ (since only Hindi in our sample contrasts it with both other bilabials), or
- three languages (Hindi, French, Mandarin) have a /p/ (since just these three languages have a /p/ phoneme that does not stray into the phonetic territory of the other phonemes).

Comparable problems recur in the comparison of every part of the linguistic system. In terms of semantics, does a language have a word for "hand" if the same word extends its denotation to the whole arm? Do we treat this as polysemy (i.e., meaning 1: "hand"; meaning 2: "arm") or do we claim it has just made a different cut on reality that bypasses any concept exactly translating "arm." In terms of syntax, does a language (say, Lao) have adjectives if they are merely a sub-sub-subclass of verbs? Or, comparing word orders and deciding on whether a language has SOV (subject–object–verb) word order, how do we deal with, for example, the following factors?

- Different orders, if one or both elements are pronouns rather than nouns.
- Languages which differ in the degree of rigidity of order, so that in some there is strict SOV order whereas in others it is a mere statistical preference.
- Languages which have different orders in main and subordinate clauses.
- Languages where it is not clear that there is a notion of "subject" in the familiar sense, but which organize their syntax in a way that con ates the patient of a transitive with the sole argument of an intransitive. In this case should the patient count as the subject in assessing order?

The sub eld of linguistics known as linguistic typology, which has for its core goals the systematization of cross-linguistic comparison, has tried to nd a

Table 13.3 Overview of comparability problems when determining which languages have a /p/ phoneme.

Phonetic realization	Hindi	Korean	French	Mandarin	English	Kayardild
[pʰ]	/pʰ/	/pʰ/		/pʰ/	/p/ or /pʰ/	/ pʰ~p~b/
[p]	/p/	/p~b/	/p/	/p/		
[b]	/b/		/b/		/b/	

---

[13] In fact, the problem is much deeper than this, since the "phones" used in the exposition here are themselves idealizations that are not exactly equivalent across languages. For an important discussion of this problem, see Ladd (2011).

way out of these impasses by developing frameworks that suggest replicable ways of getting around some of these problems. The key analytic elements in doing this are:

1. The systematic distinction of language-specic terms (where structuralist methods can be employed so as not to camouage linguistic particularity) from "comparative concepts" that serve as anchor points (or *tertium comparationis*) between languages.
2. The use of exible denitions employing either prototypes or *canonical types* to place the anchor points at those points in the design space which make the comparison with actual language material most useful.
3. The progressive factorization of comparative concepts so as to allow for the formulation of a number of dimensions, which can then be assessed independently, such as the breaking down of criteria for "subject" into grammatical elements (again factorizable into such dimensions as position, government of agreement, case choice etc.), discourse elements (e.g., topicality), and semantic elements (e.g., preferential projection of agentive roles).

None of these steps are analytically simple, and there is typically a substantial time lag between primary descriptive materials on many languages and the assembly of data from those into typological surveys, with the result that key denitional facts may be missing because they were not held to be important at the time of description. For this reason, the typological surveys with the highest-level of coding consistency have resulted from "Leipzig School" projects, (e.g., APICS[14] or the Leipzig LoanWord project), where the project design involved a group of typological masterminds calling iteratively on the expertise of a number of language experts who would be trained into the comparison methods. The obvious disadvantages of this approach, however, are the relatively limited number of sample points and the lack of extensibility: when the project ends, there is typically no new data entry. For these reasons, others advocate a much more open approach where any expert is, in principle, free to add their own data to a permanently updatable database. The advantages and disadvantages of this approach are almost the converse of the Leipzig school: greater coverage and extensibility on the one hand, but potentially lower data reliability on the other.

Whatever the problems outlined above, it is clear that the eld of typology is very signicantly extending our knowledge of the world's linguistic diversity as it lumbers through these various difculties. We now have large databases, like WALS or APICS, and scores of books synthesizing worldwide linguistic data to give reasonably workable ontologies for many of the categories linguists wish to compare (e.g., the Cambridge series with titles like Number,

---

[14]   Atlas of Pidgin and Creole Structures; see http://wwwstaff.eva.mpg.de/~taylor/apics/

Case, Aspect, Gender etc., each addressing a particular morphosyntactic category). Although we are still a long way from having a workable ontology of linguistic categories, there is substantial and accelerating mutual feedback between typologists and descriptivists, which is breaking down many of the old inconsistencies in descriptive practices. The biggest challenge that remains is the basic one of getting descriptive data (in the form of the classical "Boasian trinity" of grammar, texts, and dictionary) in the rst place. Depending on where we set the bar for a reasonable description, we still only have coverage of perhaps 10–20% of the world's linguistic diversity. Without extending this, we are a long way short of having the sorts of data sets we need to study cultural evolution seriously.

Dimensions of the Design Space

Toward a Total Ontology of the Design Space

Perhaps the biggest challenge for linguistic typology is to develop a total ontology of the design space (i.e., a clearly de ned ontology for all possible linguistic phenomena). Although this has in fact been a relatively unarticulated goal of typology for several decades, the recent adoption of the term "ontolinguistics" by some linguists (Schalley and Zaefferer 2007) and the so-called GOLD (general ontology for linguistic description) initiative have begun to make the ontological aims of much of typology more explicit. Within phonetics, of course, this goal has been a driving force for a long time—the explicit design goal of the International Phonetic Alphabet is to unambiguously represent and characterize any attested speech sound—and there are explicit procedures for admitting newly discovered sounds to the set.

An initial and deceptively straightforward example of exhaustively characterizing one subset of the design space for the dimension of main clause word order is Greenberg's famous "word-order typology" (Greenberg 1963). Using three elements—subject (S), object (O), and verb (V)—six possible orders are generated: SOV, SVO, VSO, VOS, OSV, OVS.[15] At the time it was initially formulated, some points in the design space were believed to be empty, but subsequent discoveries (e.g., the nding that the Carib language, Hixkaryana, was OVS; Derbyshire 1977) have shown that all combinations in this set are attested, albeit in highly skewed fashion.

---

[15] Among the assumptions needed, for this to be a complete ontology, are: (a) S and O are unproblematically identi able as units in the language, not at all clear in the case of ergative languages like Dyirbal, for example, or in those where "subject" properties are split between both NPs; (b) the language will have an identi able basic word order (if a language has free order of elements, with roles signaled by case, it must then be treated as uncategorizable or one of its orders chosen as basic on grounds such as frequency).

Our grand global ontology can be built on any of the unit types discussed earlier, or any combination thereof. For example, one can survey whether nasal segments are in fact more common (yes) or universal (no) as cross-linguistic expressions of negation. Two very important lines of research in typology focus on (a) constructions used for expressing particular meanings (e.g., reciprocals comparable to English "each other") and (b) semantic dimensions that get grammaticalized and hence form core building blocks in the semantic machinery of languages. A brief look at each of these will show some of the challenges linguists still face in making their ontologies comprehensive.

First consider reciprocal constructions, a central preoccupation both because of the importance of reciprocity in human culture and ethics (cf. Gintis, this volume) as well as the syntactic intricacy of reciprocals in many languages (leading "anaphors" like *each other* to play leading roles in syntactic analyses in the generative tradition). An obvious question[16] is: How many ways can languages do reciprocals? In other words, how many basic types of construction types are there? WALS is no use here; it simply tells us whether languages have "dedicated" reciprocals (like *they see each other*) or coerce the means used for expressing re exives (*ils se voient*, *sie sehen sich*, etc.). However, there are so many ways of implementing dedicated reciprocals grammatically that we need a whole raft of factors to account for the design space of just this one grammatico-semantic category. In an elegant article, König and Kokutani (2006) proposed four types: *quanti cational* (like each other), *pronominal*, *af xal*, and *deverbal*. However, further research (Evans 2008) shows that this typology is still far from comprehensive.

Among other means, languages can use special reciprocal auxiliaries, call upon symmetric simultaneous signs (in Indo-Pakistani sign language) such as two sts going toward each other, or develop strange "clause-and-a-half" constructions which have shed a "contrastive subject" pronoun out of a tit-for-tat following clause, giving something like "she-him gave and.he.in.turn money" for "the two of them gave each other money." What a more comprehensive investigation shows, then, is that many linguistic phenomena have a "long tail distribution": a small number of regular structural solutions account for most languages. However, to arrive at a full account, we need a much larger design space. Although for some purposes it is convenient just to work with the common types, there are other goals, such as getting the constructions which represent the full semantics most explicitly, or accounting for seemingly unmotivated characteristics found in the common constructions,[17] or simply

---

[16] There are also interesting questions about the semantic content of reciprocals: Do we count "the students followed each other onto the stage" or "the cop and the robber chased each other down the road"? These are nontrivial questions but they would take us too far a eld here; for discussion, see Evans et al. (2011).

[17] See Evans (2010) for examples. There it is argued that a rare construction shown in some highlands New Guinea languages actually come closest to representing the full semantics of reciprocal constructions, and that by making the presence of both transitive and intransitive

accounting for where languages seem to have the most diverse range of structural solutions. In the case of reciprocals, this is arguably because of the engineering challenge of mapping several propositions—an action going in each direction, plus the coordination or feedback joining the actions together—into a single clause. For phenomena like this, then, it may be the rare constructions that are actually more informative.

Second, consider the problem of factorizing the semantic dimensions of the design space, which I will exemplify with a relatively "new" category known as the mirative. General linguistic theory, and logic, have long been aware of the categories of tense, aspect, and mood, each deeply intertwined with the systems of verb in ection in most familiar languages. (Many analytic problems remain in each of these, but the general dimensions are relatively familiar.) As of the 1950s, at an accelerating pace, a further dimension of so-called "evidentials" began to be explored, as accounts from languages from the Amazon to the Andes to the Caucasus to the New Guinea Highlands have come to indicate what a vital role the marking of evidence source plays in the core grammatical systems of many languages. This led to a basic framework for describing the semantics of evidential systems (e.g., direct participation, direct perception, visual vs. other, inference, hearsay), the degrees of evidentiary weight assigned to each of these, who is taken to be the evidential source, and so on. Following hard on the evidential wave, another interesting type of verbal category has begun to come to light, initially on the basis of languages like Turkish and Macedonian on the one hand and Lhasa Tibetan on the other; this category presents the degree to which the information is new (mirative) or already cognitively integrated. The discovery of a category like this, however, immediately raises a whole host of questions: Is it just the speaker's cognitive state that is at issue or could it be, for example, the hearer's in a question, or even the intersubjective cognitive state of both? Must the relevant cognitive state be in the here-and-now of current conversation or can it be projected back to an earlier moment of realization (perhaps tensed)? Does the engagement with the new information occur just at the level of the proposition or can it be narrowed down to particular entities that participants are for example, pointing to? In other words what is the scope of the cognitive attitude at issue? All of these are currently the subject of lively exploration, and it is not my goal to explore these thoroughly here. I simply want to show that the turning up of some new semantic dimension in one or two languages can sometimes open up a whole new multidimensional space for ontological exploration.

An important trend in typology has been to map the semantic topology of the design space by successively smaller points. Indeed, we could say that every time a language is found that makes a previously unreported distinction,

predicates explicit in the semantic structure these rare types then motivate the widespread mixing of transitive and intransitive features found in the commoner types of reciprocal construction (Evans et al. 2007).

then the need to distinguish two points is made clear, across which the use of common forms in a given language is then mapped. Consider two (among many) senses that are formally conflated by the English perfect: *The dog has eaten the roast* and *John has eaten fugu fish*. When we look at how these get rendered in other languages, we find that many (e.g., Japanese, many Sinitic languages) have a special "experiential perfect" that would be used for the second but not the first, which would be expressed by other means. The use of a single construction to express both in English is thus a contingent fact; we accommodate this within a more general account of aspectual typology by postulating distinct points in semantic space (resultant state, experiential perfect) which some languages (English) conflate while others (Japanese) do not. The technique of "semantic maps," which overlays the semantic range of particular forms in a number of languages over a language-independent semantic grid, is used to explore which points in this space attest conflation of this sort. Results in a number of areas of grammatical semantics, such as mood/modality and indefinite pronouns, suggest that semantic extensions are highly motivated (by semantic closeness and possession of similar elements). This makes it possible to set up a general topology such that language-specific semantic ranges always span contiguous points (see Figure 13.1).

These sallies into the grand realm of what a total linguistic ontology would look like should give some idea of what a vast set of dimensions and multifactorial subspaces would be needed to characterize the design space completely in a way that every human language could be accurately characterized within it. The typological work that goes into constructing this ontology has a double
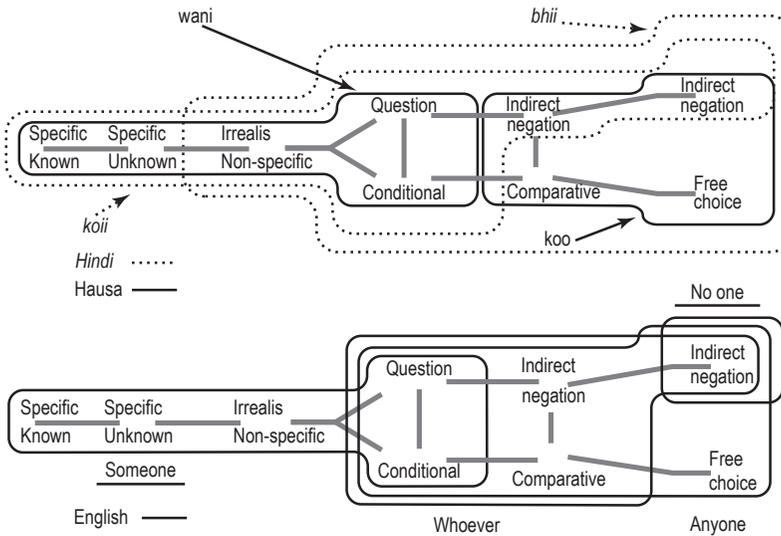

Figure 13.1 Partial semantic map of indefinite pronoun types, comparing English, Hindi, and Hausa (adapted from Haspelmath 1997)

purpose. First, it aims at being able to characterize any linguistic phenomenon, in any language, in a precise way that relates it to what is found in all other languages. Second, it allows any language to be given a comprehensive typological profile, in terms of a clear value (or "character") at each of many hundreds of thousands of dimensions: phonological, combinatoric, semantic, including complex combinations of these such as a given polysemy, or a given construction for expressing a given meaning.[18]

Profiles of this latter type are likely to play an increasingly central role in examining phylogenetic relationships between languages, simply because of the vast amounts of information they include. There have already been a number of interesting attempts to do this (Reesink et al. 2009), and despite some interesting results the field has not yet been willing to put such findings on an equal footing with the gold standard yielded by the comparative method. My personal view is that it is just a matter of time and of experimenting with the right choices of traits before this method really comes into its own. I will close this section by considering three crucial issues which bear on this question.

Distribution of Characters in the Design Space

Most points in the design space have a very skewed distribution of characters. This is the same whether one looks at basic word order (87% of languages are either SOV or SVO), vowel inventories (only about a dozen of the world's vowel inventories have a "vertical" system where only tongue height matters, not frontness or roundedness), the sensory modalities to which "see" and "hear" can extend, and so forth. The great rarity of many theoretically possible characters has given rise to many claims about "universals," but typically what begins as a claimed absolute gap turns into a mere rarity as the sample is extended, another manifestation of the "long tail" phenomenon mentioned above. This is not to say that there may never be gaps which remain absolutely unfilled, but the combination of skewed distributions with discovery lag and sampling issues mean that at any one historical moment in the evolution of typology we will have many "falsely absolute gaps." Instead of stipulating these gaps as cognitively impossible, a more useful universal approach would be to develop general and comprehensive models of selector bias. These models are needed anyway for the rest of the distribution (whether or not absolute gaps are found), and still retain their validity in accounting for rara even where these are "rehabilitated" gaps. As the title of an important article by Dryer (1998) put it: "statistical universals are better than absolute universals" (though personally

---

[18] Here I skirt around the problem of how to convert typical typological characterizations into attribute-value pairs and merely point out that, mathematically, each typological attribute can be defined as having a number of potential values, for which a score can be assigned. Where the choice set is greater than two (as with the Greenbergian word-order characters), it is possible to factorize down the dimensions to yield binary codings (e.g., treating the S|V, the O|V and the S|O orders separately).

I would prefer "statistical trends" to "statistical universals"); cf. Christiansen and Chater's (2008:500) remark that "universals" are more akin to statistical trends tied to patterns of language use.

Many approaches have been taken to explain these asymmetries: generativists in terms of universal grammar stipulating some systems as possible; functionalists and cognitivists appealing to "markedness," "naturalness," frequency, or general cognitive preferences. Regardless of theoretical persuasion, all such skewings ultimately require explanation, with the due caution that some may be due to sampling biases.

Unless one takes the asymmetries themselves to be hardwired, a plausible line of general attack is through selector biases on what can evolve, drawing the selectors from a wide range of types. These range from perceptual or articulatory constraints, to general cognitive constraints, to system constraints favoring combinatoriality and discreteness in the evolution of efficient signaling systems, or conjunctive category characterizations (A and B) over disjunctive ones (A or B).[19] As with other complex systems, such as engineering solutions or complex species adaptations, the number of equivalently adaptive solutions increases with the number of selectors, whose interactions generate large numbers of more-or-less equivalent local optima (in evolutionary biology, cf. Niklas 1994; 2004; for an application of complex adaptive systems perspective to language, see Beckner et al. 2009).[20] Some of these selectors may not be equally distributed across all speech communities (see below), thus opening a chink for genetic, cultural, and sociolinguistic biases to apply.

Coupling of Characters (Implicational Universals)

A rich seam of typological work from the 1960s, originating in the Greenbergian tradition, has explored "implicational universals": statements about the likelihood or possibility of one character in a language, given another (i.e.,

---

[19] For example, many languages have "Eskimo" kin systems like English "father" vs. "uncle" (F ≠ FB = MB) (lineal +1 generation male vs. collateral +1 generation male); many have "Dravidian" kin systems like Kayardild *kanthathu* "F, FB" vs. *kakuju* "MB" (patrilateral +1 generation male vs. nonpatrilateral +1 generation male); many have Hawaiian kin systems which use a single term for all three (F= FB=MB); and many have "Sudanese" systems which distinguish all three (F≠FB≠MB). None, however, are known in which F = MB ≠ FB. Greenberg (1990) attributed this gap to the fact that it would require a disjunctive definition whereas all the others can be characterized conjunctively.

[20] "[E]ngineering theory shows that the number of equally efficient designs for an artifact generally is proportional to both the number and the complexity of the tasks that an artifact must perform" (Niklas 1994:6772), and with respect to simulated biological evolution, "morphological diversification became easier on complex as opposed to simple fitness landscapes. Likewise, it is biologically reasonable to suppose that the morphological diversity manifested by extant species occupying similar or identical habitats vouchsafes that very different phenotypes can have equivalent capacities for growth, survival, or reproductive success" (Niklas 2004:65).

pertaining to the coupling or linkage of traits). Three well-known examples, one each from grammar, semantics, and phonology, are:

1. Word-order universals: linking basic clause order (SVO, SOV, etc.) to the order of adpositions or possessive expressions with respect to the NP they are connected to. For example, Greenberg Universal #4 claimed that SOV order implicates postpositions rather than prepositions.
2. Semantic universals: claims about the structure of the color lexicon to the effect that a language would, for example, only possess a distinct word for orange or purple if it also possesses distinct words for green and blue. The implicational hierarchy for the elaboration of color term terminologies, proposed by Berlin and Kay (1969), generates a large number of speci c implicational universals of this type.
3. Phonological universals: if a language has N distinct oral vowel phonemes, the number of distinct nasal vowel phonemes will not exceed N.

Like unconditional[21] universals, implicational statements generally tend to turn out to be statistical tendencies as research proceeds. Often this takes longer to discover because implicational statements generate a greater number of logical cells that require independent testing—the product of possibilities in each dimension under examination.

As with absolute universals, each statistical association needs explanation. In the case of word-order associations, for example, explanations have been advanced in terms of consistent parsing or processing orders across units of different types; in the case of color terms, explanations appeal to the differential sensitivity of color receptors at different wavelengths.[22] Each of these has its own trajectory of unfolding debate. The point is that implicational universals introduce additional phenomena (i.e., those de ned by trait linkage), which need external explanation above and beyond the selector biases operating on individual traits.

Implicational universals thus relate to the correlated evolution or "coupling" of characters in more general evolutionary terms. We can view correlations as a static result (i.e., languages are more learnable or processable if particular traits cohere) or as the result of correlated evolution (i.e., the evolution of one trait favors or disfavors another). In either case, because trait correlation must result from correlated evolutionary processes, any claims about coupling are particularly susceptible to Galton's problem (i.e., apparent statistical correlations may merely re ect oversampled inheritances from a clade in which they happen to be correlated). In addition, recent work (e.g., Dunn et al. 2011b) has

---

[21] I avoid the term "absolute universal" here, since it incurs an ambiguous double opposition: to statistical (tendency) and to conditional.

[22] See Loreto et al. (2012) for a recent evolutionary multiagent simulation producing the expected emergence of color vocabularies based on the attested uneven distribution of just noticeable-difference across the wavelength.

turned up evidence for what might be called "pseudo-coupling" (i.e., apparent universal correlations which turn out to be lineage speci c, such as a word-order coupling which holds in Austronesian but not in Bantu or vice versa).

Among the interesting reasons for this sort of nding are (a) possible noncomparability of units (e.g., that what gets counted as an adposition is not the same, functionally, from one group to another), (b) possible noncomparability of syntactic environments which act as selectors on units, and (c) the possibility that extrinsic selectors (e.g., discourse styles) are the real drivers of directed evolution, rather than the co-occurring features under examination. The latter is an example of the phenomenon Sapir called "drift," which has rather a different meaning in linguistics than in evolutionary biology, and denotes the mysterious phenomenon of related languages following parallel evolutionary pathways for no evident reason.

A more precise picture of character coupling than we now possess would bring many advantages. In terms of harnessing character pro les to phylogenetic inferencing, the more independent (uncoupled) or weakly coupled characters we can nd, the more likely we are to detect signals of deep relatedness. It can help us evaluate the degree to which linguistic systems really are tightly integrated from a functional point of view (which tends to be the default assumption by linguists), as opposed to just looking like this because hundreds of thousands of traits usually get transmitted together. On the other hand, because our shallow time barrier for demonstrating linguistic relatedness will fail to detect cases where many tips grow out of the same iceberg, there is enormous potential for Galton's problem to go undetected, and for co-inherited traits to appear to result from correlated evolution.

Diachronic Typology

The grand ontology described above conceives of a design space of language states: what is their phoneme inventory, what is the meaning and syntax of "know," what is the basic clause order, and thousands upon thousands of other questions. Equally, we can turn our attention to the transitions between states. For example, how does a language develop tone (tonogenesis), or how does it get from one basic word order to another? Are all logical types of transition equally probable? What preconditions need to be in place for a particular evolutionary step to occur? These are the concerns of the eld of diachronic typology.

A general evolutionary postulate here is that every synchronic phenomenon has a diachronically understandable pathway. To understand tonogenesis, for example, we need to understand how particular features of the consonants impact on the pitch of the vowels they adjoin. To understand word-order change we need to understand how other processes either move around some elements (topicalization or focalization) or reduce them to the point where they are out of the word-order game (e.g., the reduction of unstressed pronouns to af xes).

Conversely, at least some synchronic gaps or rara can be explained by the lack of ready diachronic pathways for evolving them (Evans 1995b) or the need to combine a number of rare features in the springboard state (Harris 2008), rather than any intrinsic lack of processability or learnability of the phenomenon itself.

For only a tiny fraction of the world's languages do we possess actual records of how the language has changed through millennia. Languages like Chinese, Japanese, Greek, Latin, or Sanskrit and their daughters are invaluable because they allow us to see what actually happened down the centuries and what features went together at any one moment. However, they represent such a small sample of the world's languages that for most diachronic changes we need to draw on other types of evidence, typically by comparing closely related languages. This is particularly the case when dealing with rara, where the diachronic evidence may, like the phenomenon itself, be con ned to just one family. Synchronic rara thus create a particular aura of urgency around documenting the closely related languages that can help us understand them.

Changes, like states, tend to display stochastic distributions, and diachronic typological databases aim to characterize frequency as well as possibility, both in terms of topology (between A and B) and direction (from A to B, or B to A?). Important new developments here include the synthetic studies of attested and unattested patterns of sound change and their mechanisms, under the banner of "evolutionary phonology" (Blevins 2004), and growing databases of semantic shifts.

Macro- and Micro-Variation

Another postulate of diachronic typology is that categorical change is typically preceded by variation: within an individual, across a speech community, or both. As Labov has frequently argued (e.g., Labov 1994), this makes microvariation a crucial part of understanding linguistic change, since it gives a way of detecting the seeds of large shifts which may themselves be unobservable because of the lengthy time spans involved.

Two types of methods have, since the 1960s, given us a very detailed picture of a number of changes and how they proceed. Functionalist studies, such as those by Bybee and Scheibman (1999), have carried out very ne-grained studies of how particular tokens are pronounced (e.g., *don't* and its various reductions) with regard to the frequency with which they occur in particular environments (e.g., before *know*, as in *dunno*). These studies tend to focus on the effects of repetition and frequency on form and emergent structure. Another approach has come from the Labovian variationist school, which focuses on the dynamic distribution of variants through speech communities. In general, variationist models of change involve three steps: generation of the change itself (e.g., within an individual, during learning, the streamlining of production

or social positioning), propagation (e.g., through social networks), and valuation (i.e., the valorization or stigmatization of a variant, or its categorization as associated with a particular social group). In terms of the types of linguistic phenomenon covered, there has been a strong bias toward the form end of language, often with detailed studies of micro-variation in pronunciation, with less on grammar and very little at all on semantics.

A key question here is: Do mechanisms of social selection interact with the sorts of structural preferences outlined above? For example, will two variants, A and B, one of which is cross-linguistically highly preferred, be treated equally by processes of social selection, and will these processes apply equally in all types of speech community? I have already suggested above (see section, Diversity in Language Structure) that some processes may depend on high levels of shared bilingualism, such as the promotion of one highly marked gender form to shibboleth status by Iwaidja speakers and its subsequent generalization. More generally, it has been argued that a process of "esoterogeny" (i.e., the elaboration of difficult difference) operates in certain types of speech communities, such as the multilingual and highly metalinguistically aware communities of much of New Guinea (Thurston 1992). In such cases, the "expensive" options may actually be selected, due to interests of signaling group affiliation from childhood. If, further, bilingual-awareness-driven change of this type is particularly characteristic of small speech communities, then we would have a situation where the sociolinguistic characteristics of the speech community impacts upon the type of change and, ultimately, on the type of language found in particular communities. I will return to this particular variant of the society–language structure coevolution hypothesis below (see section on Social and Cultural Selection).

These examples should make clear that studies of micro-variation need to be connected to studies of macro-variation. Unfortunately, micro-variation in small-scale speech communities is one of the most neglected aspects of current linguistic research, and the last two decades of intensive language documentation have by and large failed to treat the organization of micro-variation as one of the dimensions in which we will discover significant cross-linguistic diversity.

Mechanisms of Selection

In the preceding sections I have argued that a central goal of linguistics is to give a full account of how and why language diversity is skewed across the design space—whether of synchronic structures or of attested diachronic pathways. In evolutionary terms, this requires us to identify the various types of selectors involved. When they are postulated, we must ultimately provide experimental or other detailed replicable evidence of how they work, such as studies

showing acoustic similarities or articulatory adjustments as responsible for certain types of sound change. Here are a number of broad types of selectors.

Psychological and Physiological

These are the traditional staple of selection-based accounts. They include cognitive biases in cognition/processing of various types, the acquisition process, and the physiology of perception and production. Data on specic conditions facilitating selection have come from laboratory studies (especially in phonetics), quantitative studies of linguistic corpora (e.g., looking at the frequency with which new vs. established mentions occur in different grammatical roles), and, increasingly, computer simulations. The general logic of these accounts is to attribute some skewing in the cross-linguistic data to the greater ease or frequency with which certain traits co-occur, leading to correlations via Zipan effects (Zipf 1935) or amplied cross-generational learnability. Well-known examples are the complementarity of palatal versus velar articulations before high front versus other vowels; the frequency of "up" polysemies with health, dynamism, activity, and happiness; or the association of ergative/absolutive case patterning with nouns and of nominative/accusative patterning with pronouns (on this last point, see DuBois 1987).

System Selection

This includes properties which favor the overall efciency of languages as intricately interconnected semiotic systems. An interest in system architecture goes back to the grammatical tradition of Paṇini in ancient India which already had a deep concern with informational parsimony and developed the rst data compression algorithms for representing complex and variable linguistic forms with extreme economy. This has resurfaced over and over in the history of linguistics through both structuralist phases (e.g., the importance of structured sets of orthogonal feature oppositions) and generative phases (including the overt reconnection of generative methods with the Paṇinian tradition, such as rules deriving variable realizations of single underlying forms).

In his seminal identication of various "design features" of language, Hockett (1960) pinpointed a number of key properties of languages as systems. These include double articulation (i.e., phonological units are inherently meaningless but then combine into meaningful morphemic units which are the level at which semantic composition begins) and arbitrariness (freeing sign form from referent properties, which is useful, e.g., in coining short names for giant numbers or creatures). Though some linguists in the generative tradition have wanted to attribute these properties to the human mind, in the form of a putative "Universal Grammar," computer simulations (Kirby 2002; Zuidema and De Boer 2009; Reali and Christiansen 2009) have shown that properties

like arbitrariness, discreteness, compositionality, and consistency in branching can emerge through recursive transgenerational selection.

Another type of system selection concerns linked typological traits. Traditions of doing this go back a long way in linguistics; for example, to Wilhelm von Humboldt in the nineteenth century, with his postulations of agglutinating, fusional, incorporating, and isolating types. More recently they have recurred through claims about implicational universals in the typological tradition or as ow-ons from "parameter settings" in the generative tradition. An example of the latter approach was taken by Baker (1996) to link a number of traits of polysynthetic languages (e.g., incorporation, lack of non nite forms). I have already outlined the dif culties in evaluating such claims empirically, because of the large numbers of traits that need to be checked across the sample, but typology is increasingly able to do this. Although many linked-trait hypotheses may be wobbling or tumbling (Dunn et al. 2011b), the precarious complexity of natural languages makes it likely that at least some will survive, at least as statistical correlations. A compelling example concerns the way modality (spoken vs. signed) impacts on a range of language structures, from word classes to semantic structures: Sign language typically includes a type of "classi er" word class not found in spoken language, and many classi ers lack a de ned class comparable to pronouns in spoken language. The structure and semantics of reciprocal constructions may also show signi cant differences from that found in spoken language (Zeshan and Panda 2011).

Social and Cultural Selection

Explanations of this type appeal to properties of the social setting or of the culture in which a language is spoken. Many linguists, such as Perkins (1995), and most recently Bentz and Christiansen (2010), Lupyan and Dale (2010), and Trudgill (2011), have suggested that the most complex and unusual linguistic systems are most likely to evolve in such small communities (see also Bentz and Christiansen 2010). This is partly because speakers in small face-to-face communities can draw on a wider range of mutual knowledge, facilitating the grammaticalization of, for example, detailed kinship information, and partly because widespread multilingualism produces a different semiotic in which sounds, words, and grammatical items are positioned in a complex multilingual space—establishing linguistic identity, in this complex semiotic space, may involve promoting forms which are "marked"—and hence unlikely to arise if psychological and perceptual/articulatory selectors were given a free rein.

The parallels with sexual selection in biology are intriguing here. For certain social-signaling purposes (e.g., signaling lifelong group membership), the most "expensive" structures (i.e., those that are hard to learn or process) may be the most suitable.

Cultural selection is most likely to operate on the semantic dimension of language organization, in the realm sometimes known as "ethnosyntax"

(Enfield 2002). Examples are the impact of particular marriage rules and family structures on kinship semantics, of social categories on honorific behavior, of certain ways of talking on the development of "kintax" in Australian languages (Evans 2003c; Blyth 2013), or of patterns of journey discussion on the emergence of "associated motion" categories (Simpson 2002).

There may be interesting and complex feedbacks between social and other selectors. An interesting example comes from Alipur sign language in southern India (Panda 2013). There, an Urdu-speaking Muslim community transplanted from Andhra Pradesh to a predominantly Hindu area in Karnataka became reproductively isolated for religious and linguistic reasons, and consequently developed high rates of congenital deafness, leading to the evolution of a village-level sign language. Here cultural features (religion) impacted on social unit size, affected some genetic traits in the population, and selected for the development of a particular linguistic modality. Given what we know of other sign languages, it is likely that the evolution of Alipur sign language would be accompanied by the selection of particular linguistic characteristics, but so far we lack detailed descriptions of this sign language.

A further example, also involving sign, comes from Warlpiri, where cultural traditions enforce a speech ban on widows for around a year after their husbands' deaths. During this time, widows dwell in women's camps where the use of sign is widespread owing to the large number of widows. This situation makes it one of the few sign languages to be primarily used by speaking people (though, of course, its existence is also a bonus to deaf Warlpiri) and, most likely as a consequence of this, it displays far greater parallelism to the structures of spoken language than is normal in sign languages (Kendon 1988).

Genetic and Epidemiological Selection

For more than a century, since Boas's forceful denials of any link between language and race, linguists and anthropologists have placed the various mechanisms outlined above beyond the effects of genetic variability. This was based on the "self-evident" fact that children raised in any culture appear to learn that culture's language flawlessly. However, it neglects to consider the possibility that iterated selector effects, minor in any one generation but amplified through coevolutionary bottlenecks over many generations, can produce significant biases over time in the emergence of particular types of system. The genie was let out of this particular bottle by Dediu and Ladd's (2007) finding that the distribution of tonal phoneme systems correlates significantly with genes (*ASPM* and *Microcephalin*) that code pitch discrimination; other vocal tract differences such as lip and palate configurations are now emerging as candidates for genetic biases that may have loaded the dice in selecting for the emergence of particular types of phoneme system (or phonetic realization of particular phonemes). We now need to contemplate, and thoroughly investigate, the possibility that a significant part of the world's linguistic variability

is due to genetic differences in the populations of speakers (for further details, see Levinson and Dediu, this volume).

Biological features of speaker populations are not con ned to genes. An intriguing series of studies by Butcher (2006) and his colleagues (Stoakes et al. 2011) raises the possibility that particular properties of the phonologies of Australian languages may have been shaped by longstanding epidemiological particularities of the Australian indigenous population (in the original, medical sense of "epidemiological"), notably chronic otitis media. Butcher suggests that a number of phonological features may be adaptations to systematic gaps in frequency perception in the high spectral range in much of the population as a result of chronic otitis media. These include "long at" phoneme inventories lacking fricatives and phonotactics that permit large numbers of "heterorganic" nasal + stop clusters; for example, the Kayardild triplet ŋa**nk**i "temple," ŋaɳki "beach (LOC)" and kaŋki "word (LOC)."

In other words, against a background of prevalent chronic otitis media knocking out perception in the high-frequency range, Aboriginal languages could have evolved (or maintained) "long at," otitis-robust phonologies that concentrate the bulk of the perceptual discriminations in the mid-frequency range. Evaluating this hypothesis is a fraught enterprise. To be fully convincing it would need to correlate detailed data on prevalence of otitis media across large numbers of language groups, and deal with the dif culty of determining prevalence in past populations whose health status may have been different.[23] Thus, in principle, epidemiological as well as genetic factors may play a role in weighting selective processes in a particular direction.

Prospects and Challenges

The extraordinary diversity of the world's languages is being reconceptualized from noise to signal. As a result, linguistics is now entering an exciting phase where language diversity and the evolutionary processes which shape it are assuming center stage. A steadily expanding set of descriptions from languages across the globe, coupled with increasingly successful methods of setting up cross-linguistic comparison, is coming to furnish the most detailed, far-reaching, falsi able and interrogable data sets we have for the world's many thousands of culture-de ned groupings. Still, as the eld stands on this threshold, it faces seven great challenges: the rst two empirical in nature, followed by ve theoretical issues:

1. *Basic descriptive coverage*. Despite recent advances, we are a long way from having anything like comprehensive descriptive coverage of

---

[23] Interestingly, the only existing paleopathological study, by Roche (1964), nds extremely high rates of aural exostoses in Aboriginal skulls, attributing much of the effect to prevalent otitis media in earlier populations.

the world's 6,000–7,000 languages. Getting beyond our current levels of around 20% coverage is a giant challenge, particularly at a time when major international programs (DoBeS, HRELP) and departments within institutions (e.g., Max Planck Institutes for psycholinguistics and evolutionary anthropology) are coming to the end of their funding lifetimes.

2. *Sociolinguistic studies of micro-variation in small-scale speech communities*. This represents a huge gap, not just in terms of coverage but also at the level of theorization and investigative tools.[24] But without studies of micro-variation in communities of the type that have shaped most of our human past, many central questions cannot be evaluated rigorously: these centrally include those concerning diversi cation, or the impact of population size on structural options.

3. *A complete characterization of the design space.* This is intimately tied up with developing an ontology for calibrating coding (and for quantifying uncertainty/ambiguity of analysis), and functional accounts of distributions. Typologists have been steadily working away at this, but this remains a work-in-progress, leaving linguists "like chemists without a list of the elements, or physicists with no account of particles" (Corbett 2012).

4. *A comprehensive categorization of transitions* compiled by induction from known changes across the world's languages and supplemented by experimental and computational studies of what leads to change. More extensive knowledge of which changes are common and which are rare will be of great help in weighting the likelihoods of alternative phylogenies. Historical linguists already do this implicitly and intuitively; it is a central part of their unquanti ed art and craft. However, explicit measures need to be developed and tested.

5. *A complete phylogenetic tree of the world's languages.* Linguistics is in the sad state, unthinkable to geneticists, of having hundreds of unconnected families with no methods for joining them up at deep levels. This is analogous to a vast collection of twigs and small branches without any larger tree that joins them together. The ±10 ky time barrier set by the comparative method, just like the time barrier set by radiocarbon dating, can only be overcome through the development of new methods capable of picking up heritable information in more sensitive ways. Here, the most promising techniques are new methods that apply phylogenetic algorithms to huge numbers of traits. As we get cross-linguistic data on an ever-growing number of characters, this will become increasingly informative.

---

[24] An honorable exception is the series of studies by James Stanford (2008a, b) on how Sui speakers in Southern China organize tonal variation around patrilines rather than age- or gender-based signaling.

6. *Understanding of microprocesses generating diversi cation* and why it appears to proceed at different rates (and in different ways) in different settings. This is the theoretical counterpart to the empirical gaps delineated in Pt. 2. Labovian sociolinguistics has developed a sophisticated model of microevolution, but it is unlikely that this covers all types of microprocesses in small-scale societies.

7. *A revisiting of external selection.* The last sixty years have been characterized by a largely unexamined consensus between generativists, functionalists, cognitivists, and typologists: that "external" factors (genetic, epidemiological, sociolinguistic group size, and cultural) can be set aside when it comes to explaining the distribution of linguistic phenomena across the design space, in favor of properties of the human cognitive, perceptual and articulatory systems (all presumed to be invariant across human populations) along with universal discourse properties. Coming into the twenty- rst century, it is looking increasingly plausible that at least some of the cross-linguistic variability may re ect these external factors. Testing hypotheses of this type will be an intricate affair, requiring much more attention to the gathering of matched nonlinguistic data sets to go with our linguistic information. If they turn out to have some currency, it will show the interplay of language, culture, and biology in human populations to be a much more dynamic and interactive process than we have imagined so far. In the process, it will give quite a different sort of answer to the basic question of why languages differ so much.

Acknowledgments