The Myth of Language Universals: Language diversity and its importance for cognitive science

Nicholas Evans
Department of Linguistics, Research School of Asian and Pacific Studies
Australian National University, ACT 0200
Australia
nicholas.evans@anu.edu.au

Stephen Levinson
Max Planck Institute for Psycholinguistics
Wundtlaan 1
NL-6525 XD Nijmegen
stephen.levinson@mpi.nl
http://www.mpi.nl/Members/StephenLevinson

Abstract: Talk of linguistic universals has given cognitive scientists the impression that languages are all built to a common pattern. In fact, there are vanishingly few universals of language in the direct sense that all languages exhibit them. Instead, diversity can be found at almost every level of linguistic organization. This fundamentally changes the object of enquiry from a cognitive science perspective.

The article summarizes decades of cross-linguistic work by typologists and descriptive linguists, showing just how few and unprofound the universal characteristics of language are, once we honestly confront the diversity offered to us by the world's 6-8000 languages. After surveying the various uses of 'universal', we illustrate the ways languages vary radically in sound, meaning, and syntactic organization, then examine in more detail the core grammatical machinery of recursion, constituency, and grammatical relations. While there are significant recurrent patterns in organization, these are better explained as stable engineering solutions satisfying multiple design constraints, reflecting both cultural-historical factors and the constraints of human cognition.

Linguistic diversity then becomes the crucial datum for cognitive science: we are the only species with a communication system which is fundamentally variable at all levels. Recognising the true extent of structural diversity in human language opens up exciting new research directions for cognitive scientists, offering thousands of different natural experiments given by different languages, with new opportunities for dialogue with biological paradigms concerned with change and diversity, and confronting us with the extraordinary plasticity of the highest human skills.

Keywords; Chomsky; Coevolution; Constituency; Culture; Dependency; Evolutionary theory; Greenberg; Linguistic diversity; Linguistic typology; Recursion; Universal grammar

1. Introduction[1]

"According to Chomsky, a visiting Martian scientist would surely conclude that aside from their mutually unintelligible vocabularies, Earthlings speak a single language" (Pinker 1994, p.232)

Languages are much more diverse in structure than cognitive scientists generally appreciate. A widespread assumption among cognitive scientists, growing out of the generative tradition in linguistics, is that all languages are English-like, but with different sound systems and vocabularies. The true picture is very different: languages differ so fundamentally from one another at every level of description (sound, grammar, lexicon, meaning) that it is very hard to find any single structural property they share. The claims of Universal Grammar, we will argue, are either empirically false, unfalsifiable, or misleading in that they refer to tendencies rather than strict universals. Structural differences should instead be accepted for what they are, and integrated into a new approach to language and cognition that places diversity at centre stage.

The misconception that the differences between languages are merely superficial, and that they can be resolved by postulating a more abstract formal level at which individual language differences disappear, is serious: it now pervades a great deal of work done in psycholinguistics, in theories of language evolution, language acquisition, neurocognition, parsing and speech recognition, and just about every branch of the cognitive sciences. Even scholars like Christiansen & Chater (2008), concerned to demonstrate the evolutionary impossibility of pre-evolved constraints, employ the term 'Universal Grammar' as if it were an empirically verified construct. A great deal of theoretical work within the cognitive sciences thus risks being vitiated, at least if it purports to be investigating a fixed human language processing capacity, rather than just the particular form this takes in some well-known languages like English and Japanese.

How did this widespread misconception of language uniformity come about? In part, this can be attributed simply to ethnocentrism – most cognitive scientists, linguists included, speak only the familiar European languages, all close cousins in structure. But in part it can be attributed to misleading advertising copy issued by linguists themselves. Unfortunate

sociological splits in the field have left generative and typological linguists with completely different views of what is proven science, without shared rules of argumentation that would allow them to resolve the issue – and in dialogue with cognitive scientists it has been the generativists who have been taken as representing the dominant view. As a result, Chomsky's notion of Universal Grammar (UG) has been mistaken, not for what is – namely the programmatic label for whatever it turns out to be that all children bring to learning a language – but for a set of substantial research findings about what all languages have in common. For the substantial findings about universals across languages one must turn to the field of linguistic typology, which has laid bare a bewildering range of diverse languages, where the generalizations are really quite hard to extract. Chomsky's views, filtered through various commentators, have been hugely influential in the cognitive sciences, because they combine philosophically sophisticated ideas and mathematical approaches to structure with claims about the innate endowment for language that are immediately relevant to learning theorists, cognitive psychologists, and brain scientists. Even though psychologists learned from the Linguistic Wars of the 1970s (Newmeyer 1986) to steer clear from too close an association with any specific linguistic theory, the underlying idea that all languages share the same structure at some abstract level has remained pervasive, tying in nicely to the modularity arguments of recent decades (Fodor 1983).

It will take a historian of science to unravel the causes of this ongoing presumption of underlying language uniformity. But a major reason is simply that there is a lack of communication between theorists in the cognitive sciences and those linguists most in the know about linguistic diversity. This is partly because of the reluctance by most descriptive and typological linguists to look up from their fascinating particularistic worlds and engage with the larger theoretical issues in the cognitive sciences. Outsiders have instead taken the articulate envoys from the universalising generativist camp to represent the consensus view within linguistics. But there are other reasons as well: the relevant literature is forbiddingly opaque to outsiders, bristling with arcane phonetic symbols and esoteric terminologies. Our first goal (§2) in this paper, then, is to survey some of the linguistic diversity that has been largely ignored in the cognitive sciences, which shows how differently languages can be structured at every level: phonetic, phonological, morphological, syntactic and semantic . We critically evaluate (§3) the kind of descriptive generalizations (again, misleadingly called "universals") that have emerged from careful cross-linguistic comparisons, and survey the

treacherously different senses of "universal" that have allowed the term to survive a massive accumulation of counterevidence.

We then turn to three syntactic features which have recently figured large in debates about the origin of language: grammatical relations (§4), constituency (§5), and recursion (§6). How universal are these features? We conclude that there are plenty of languages that do not exhibit them in their syntax. What does it mean for an alleged universal to not apply in a given case? We will consider the idea of "parameters" and the idea of UG as a "toolkit" (Jackendoff 2002).

We then turn (§7) to the question of how all this diversity is to be accounted for. We suggest, first, that linguistic diversity patterns just like biological diversity, and should be understood in the same sorts of ways, with functional pressures and systems constraints engineering constant small changes. Finally (§8), we advance seven theses about the nature of language as a recently-evolved bio-cultural hybrid. We suggest that refocusing on a unique property of our communication system, namely its diversity, is essential to understanding its role in human cognition.

2. Language diversity

A review of leading publications suggests that cognitive scientists are not aware of the real range of linguistic diversity. In Box 1, for example, is a list of features, taken from a BBS publication on the evolution of language, that all languages are supposed to have – "uncontroversial facts about substantive universals" (Pinker & Bloom 1990; a similar list is found in Pinker 1994). But none of these "uncontroversial facts" are true of all languages, as noted in the box.

The crucial fact for understanding the place of language in human cognition is its diversity. For example, languages may have less than a dozen distinctive sounds, or they may have twelve dozen, and sign languages do not use sounds at all. Languages may or may not have derivational morphology (to make words from other words, e.g. *run > runner*), or inflectional morphology for an obligatory set of syntactically consequential choices (e.g. plural *the girls are* vs. singular *the girl is*). They may or may not have constituent structure (building blocks

of words that form phrases), may or may not have fixed orders of elements, and their semantic systems may carve the world at quite different joints. We detail all these dimensions of variation below, but the point here is this: *We are the only known species whose communication system varies fundamentally in both form and content*. Speculations about the evolution of language that do not take this properly into account thus overlook the criterial feature distinctive of the species. The diversity of language points to the general importance of cultural and technological adaptation in our species: language is a biocultural hybrid, a product of the intensive gene:culture coevolution over perhaps the last 2-300,000 years (Boyd & Richerson 1985; Levinson & Jaisson 2006; Enfield & Levinson 2006; Laland et al. 2000).

Why should the cognitive sciences care about language diversity, apart from their stake in evolutionary questions? First, a proper appreciation of the diversity completely alters the psycholinguistic picture: What kind of language processing machine can handle all this variation? Not the conventional one, built to handle the parsing of European sound systems and the limited morphological and syntactic structures of familiar languages. Imagine a language where instead of saying 'This woman caught that huge butterfly' one says something like 'That$_{object}$ this$_{subject}$ huge$_{object}$ caught woman$_{subject}$ butterfly$_{object}$.' – such languages exist (§4). The parsing system for English can't be remotely like the one for such a language: what then is constant about the neural implementation of language processing across speakers of two such different languages? Second, how do children learn languages of such different structure, indeed languages that vary in every possible dimension? Can there really be a fixed 'Language Acquisition Device'? These are the classic questions about how language capacities are implemented in the mind and in the brain, and the ballgame is fundamentally changed when the full range of language diversity is appreciated.

The cognitive sciences have been partially immunized against the proper consideration of language diversity by two tenets of Chomskyan origin. The first is that the differences are somehow superficial, and that expert linguistic eyes can spot the underlying common constructional bedrock. This, at first a working hypothesis, became a dogma, and it's wrong, in the straightforward sense that the experts either cannot formulate it clearly, or do not agree that it is true. The second was an interesting intellectual program that proceeded on the hypothesis that linguistic variation is 'parametric', that is, that there are a restricted number of binary switches, which in different states project out the full set of possible combinations, explaining observed linguistic diversity (Chomsky 1981; Baker 2001). This hypothesis is

now known to be false as well: its predictions about language acquisition, language change, and the implicational relations between linguistic variables simply fail (Newmeyer 2004; 2005). The conclusion is that the variation has to be taken at face value – there are fundamental differences in how languages work, with long historico-cultural roots that explain the many divergences.

Once linguistic diversity is accepted for what it is, it can be seen to offer a fundamental opportunity for cognitive science. It provides a natural laboratory of variation in a fundamental skill – seven thousand natural experiments in evolving communicative systems, and as many populations of experts with exotic expertise. We can ask questions like: how much longer does it take a child to master one hundred and forty four distinctive sounds versus eleven? How do listeners actually parse a free word order language? How do speakers plan the encoding of visual stimuli if the semantic resources of the language make quite different distinctions? How do listeners break up the giant inflected words of a polysynthetic language – in Bininj Gun-wok (Evans 2003a) for instance, the single word *abanyawoihwarrgahmarneganjginjeng* can represent what, in English, would constitute an entire sentence: 'I cooked the wrong meat for them again'. These resources offered by diversity have scarcely been exploited in systematic ways by the scientific community: we have a comparative psychology across species, but not a proper comparative psychology inside our own species in the central questions that drive cognitive science.

2.1 The current representation of languages in the world

Somewhere between 5000 and 8000 distinct languages are spoken today. How come we can't be more precise? In part because there are definitional problems: when does a dialect difference become a language difference (the 'languages' Czech and Slovak are far closer in structure and mutual intelligibility than so-called 'dialects' of Chinese like Mandarin and Cantonese)?. But mostly it is because academic linguists, especially those concerned with primary language description, form a tiny community, far outnumbered by the languages they should be studying, each of which takes the best part of a lifetime to master. Less than 10% of these languages have decent descriptions (full grammars and dictionaries). Consequently, nearly all generalizations about what is possible in human languages are based on a maximal 500 language sample (in practice, usually much smaller – Greenberg's famous universals of

language were based on 30), and almost every new language description still guarantees substantial surprises.

Ethnologue, the most dependable world-wide source (http://www.ethnologue.com/), reckons that 82% of the world's 6,912 languages are spoken by populations under 100,000, 39% by populations under 10,000. These small speaker numbers indicate that much of this diversity is endangered. Ethnologue lists 8% as nearly extinct, and about one language dies a week. This loss of diversity, as with biological species, drastically narrows our scientific understanding of what makes a possible human language.

Equally important as the brute numbers are the facts of relatedness. The number of language families is crucial to the search for universals, since typologists want to test hypotheses against a sample of independent languages. The more closely two languages are related, the less independent they are as samplings of the design space. The question of how many distinct phylogenetic groupings are found across the world's languages is highly controversial, though Nichols' (1992) estimate of 300 'stocks' is reasonable, and each stock itself can have levels of divergence that make deep-time relationship hard to detect (English and Bengali within Indo-European; Hausa and Hebrew within Afroa-asiatic). In addition, there are over 100 isolates, languages with no proven affiliation whatsoever. A major problem for the field is that we currently have no way of demonstrating higher-level phylogenetic groupings that would give us a more principled way of selecting a maximally independent sample for a set smaller than these 3-400 groups. This may become more tractable with the application of modern cladistic techniques (Gray & Atkinson 2003; Dunn et al. 2005; McMahon & McMahon 2006), but such methods have yet to be fully adopted by the linguistic community.

Suppose then that we think of current linguistic diversity as represented by 7000 languages falling into 300 or 400 groups. Five hundred years ago, before the expansion of Western colonization, there were probably twice as many. Because most surviving languages are spoken by small ethnic groups, language death continues apace. If we project back through time, there have probably been at least half a million human languages (Pagel 2000), so what we now have is a non-random sample of under 2% of the full range of human linguistic diversity. It would be nice to at least be in the position to exploit that sample, but in fact, as mentioned, we have good information for only 10% of that. The fact is that at this stage of

linguistic inquiry, almost every new language that comes under the microscope reveals unanticipated new features.

2.2   Some dimensions of diversity

In this section we illustrate some of the surprising dimensions of diversity in the world's languages. We show how languages may or may not be in the articulatory-auditory channel, and if they are how their inventories of contrastive sounds vary dramatically, how they may or may not have morphologies (processes of word derivation or inflection), how varied they can be in syntactic structure or their inventory of word classes, and how varied are the semantic distinctions which they encode. We can do no more here than lightly sample the range of diversity, drawing attention to a few representative cases.

2.2.1 Sound inventories. We start by noting that some natural human languages do not have sound systems at all. These are the sign languages of the deaf. Just like spoken languages, many of these have developed independently around the world, wherever a sufficient intercommunicating population of deaf people has arisen, usually as a result of a heritable condition (Ethnologue, an online inventory of languages, lists 121 documented sign languages, but there are certainly many more). These groups can constitute both significant proportions of local populations and substantial populations in absolute terms: in India there are around 1.5 million signers. They present interesting, well circumscribed models of gene-culture co-evolution (Aoki & Feldman 1994; Durham 1991): without the strain of hereditary deafness, the cultural adaptation would not exist, while the cultural adaptation allows signers to lead normal lives, productive and reproductive, thus maintaining the genetic basis for the adaptation.

The whole evolutionary background to sign languages remains fascinating but obscure – were humans endowed, as Hauser (1997, p.245) suggests, with a capability unique in the animal world to switch their entire communication system between just two modalities, or (as the existence of touch languages of the blind-deaf suggest) is the language capacity modality-neutral? There have been two hundred years of speculation that sign languages may be the evolutionary precursors to human speech, a view recently revived by the discovery of mirror-neurons (Arbib 2005). An alternative view is that language evolved from a modality-hybrid communication system in which hand and mouth both participated, as they do today in both

spoken and signed languages (cf. Sandler 2008). Whichever evolutionary scenario you favour, the critical point here is that sign languages are an existence proof of the modality-plastic nature of our language capacity. At a stroke, therefore, they invalidate such generalizations as "all natural languages have oral vowels", although at some deeper level there may well be analogies to be drawn: signs have a basic temporal organization of 'move and hold' which parallels the rhythmic alternation of vowels and consonants.

Returning to spoken languages, the vocal tract itself is the clearest evidence for the biological basis for language – the lowering of the larynx and the right-angle in the windpipe have been optimized for speaking at the expense of running and with some concomitant danger of choking (Lenneberg 1967). Similar specializations exist in the auditory system, with acuity tuned just to the speech range, and, more controversially, specialized neural pathways for speech analysis. These adaptations of the peripheral input/output systems for spoken language have, for some unaccountable reason, been minimized in much of the discussion of language origins, in favor of an emphasis on syntax (see for example Hauser, Chomsky & Fitch 2002).

The vocal tract and the auditory system put strong constraints on what an articulatorily possible and perceptually distinguishable speech sound is. Nevertheless, the extreme range of phonemic (distinctive sound) inventories, from 11 to 144, is already a telling fact about linguistic diversity (Maddieson 1984). Jakobson's distinctive features – binary values on a limited set of (largely) acoustic parameters – were meant to capture the full set of possible speech sounds. They were the inspiration for the Chomskyan model of substantive universals, a constrained set of alternates from which any particular language will select just a few. But as we get better information from more languages, sounds that we had thought were impossible to produce or impractical to distinguish keep turning up. Take the case of double-articulations, where a consonantal closure is made in more than one place. On the basis of evidence then available, Maddieson (1983) concluded that contrastive labial-alveolar consonants (making a sound like 'b' at the same time as a sound like 'd') were not a possible segment in natural language on auditory grounds. But it was then discovered that the Papuan language Yélî Dnye makes a direct contrast between a coarticulated 'tp', and a ' p' where the   is further back towards the palate (Ladefoged & Maddieson 1996, p.344-5; Maddieson & Levinson in preparation).

As more such rarities accrue, experts on sound systems are abandoning the Jakobsonian idea of a fixed set of parameters from which languages draw their phonological inventories, in favour of a model where languages can recruit their own sound systems from fine phonetic details that vary in almost unlimited ways (see also Mielke 2007; Pierrehumbert, Beckman & Ladd 2000):

> Do phoneticians generally agree with phonologists that we will eventually arrive at a fixed inventory of possible human speech sounds? The answer is no. (Port & Leary 2005, p. 927).

> Languages can differ systematically in arbitrarily fine phonetic detail. This means we do not want to think about universal phonetic categories, but rather about universal phonetic resources, which are organized and harnessed by the cognitive system.... The vowel space – a continuous physical space rendered useful by the connection it establishes between articulation and perception – is also a physical resource. Cultures differ in the way they divide up and use this physical resource. (Pierrehumbert 2000, p.12).

2.2.2 Syllables and the "CV" universal. The default expectation of languages is that they organize their sounds into an alternating string of more vs. less sonorant segments, creating a basic rhythmic alternation of sonorous vowels (V) and less sonorous consonants (C). But beyond this, a further constraint was long believed to be universal: that there was a universal preference for CV syllables (like *law* /lɒː/ or *gnaw* /nɒː/) over VC syllables (like *awl* /ɒːl/ or *awn* /ɒːn/). The many ways in which languages organize their syllable structures allows the setting up of implicational (if/then) statements which effectively find order in the exuberant variation: no language will allow VC if it does not also allow CV, or allow V if it does not also allow CV:

(1)   CV > V > VC

This long-proclaimed conditional universal (Jakobson & Halle 1956; Jakobson 1962; Clements & Keyser 1983) has as corollary the Maximal Onset Principle (Blevins 1995, p.230): a /....VCV..../ string will universally be syllabified as /...V-CV.../. An obvious advantage such a universal principle would give the child is that it can go right in and parse strings into syllables from first exposure.

But in 1999, Breen & Pensalfini published a clear demonstration that Arrernte organizes its syllables around a VC(C) structure, and does not permit consonantal onsets. With the addition of this one language to our sample, the CV syllable gets downgraded from absolute universal, to a strong tendency, and the status of the CV assumption in any model of UG must be revised. If CV syllables really were inviolable rules of UG, Arrernte would then be unlearnable, yet children learn Arrernte without difficulty. At best, then, the child may start with the initial hypothesis of CVs, and learn to modify it when faced with Arrernte or other such languages. But in that case we are talking about initial heuristics, not about constraints on possible human languages. The example also shows, as is familiar from the history of mathematical induction (as with the Gauss-Riemann hypothesis regarding prime number densities), that an initially plausible pattern turns out not to be universal after all, once the range of induction is sufficiently extended.

2.2.3 Morphology. Morphological differences are among the most obvious divergences between languages, and linguistic science has been aware of them since the Spanish encountered Aztec and other polysynthetic languages in sixteenth-century Mexico, while half a world away the Portuguese were engaging with isolating languages in Vietnam and China. Isolating languages, of course, lack all the inflectional affixes of person, number, tense, aspect, etc., and systematic word derivation processes. Isolating languages lack even the rather rudimentary morphology of English words like *boy*-*s* or *kiss*-*ed*, using just the root and getting plural and past-tense meanings either from context or from other independent words. Polysynthetic languages go overboard in the other direction, packing whole English sentences into a single word, as in Cayuga *Ęskakheh   na'táyęthwahs* 'I will plant potatoes for them again.' (Evans & Sasse 2002). Clearly, children learning such languages face massive challenges in picking out what the 'words' are which they must learn. They must also learn a huge set of rules for morphological composition, since the number of forms that can be built from a small set of lexical stems may run into the millions (Hankamer 1989).

But if these very long words function as sentences, perhaps there's no essential difference: perhaps, for example, the Cayuga morpheme *-ho   na-* for potatoes in the word above is just a word-internal direct object as Baker (1993; 1996) has claimed. However, the parallels turn out to be at best approximate. For example, the pronominal affixes and incorporated nouns do not need to be referential. The prefix *ban-* in Bininj Gun-wok *ka-ban-dung* [she-them-scolds] is only superficially like its English free-pronoun counterpart, since *kabandung* can mean

both 'she scolds them' and 'she scolds people in general' (Evans 2002). It seems more likely, then, that much of the obvious typological difference between polysynthetic languages and more moderately synthetic languages like English or Russian needs to be taken at face value: the vast difference in morphological complexity is mirrored by differences in grammatical organization right through to the deepest levels of how meaning is organized.

2.2.4 Syntax and word-classes. Purported syntactic universals lie at the heart of most claims regarding UG, and we will hold off discussing these in detail until §4-6. As a warm-up, though, we look at one fundamental issue: word-classes, otherwise known as parts of speech. These are fundamental to grammar, because the application of grammatical rules is made general by formulating them over word classes. If we say that in English adjectives precede but cannot follow the nouns they modify (*the rich man* but not **the man rich*) we get a generalization that holds over an indefinitely large set of phrases, since both adjectives and nouns are 'open classes' that in principle are always extendable by new members. But to stop it generating **the nerd zappy* we need to know that *nerd* is a noun, not an adjective, and that *zappy* is an adjective, not a noun. To do this we need to find a clearly delimited set of distinct behaviours, in their morphology and their syntax, that allows us to distinguish noun and adjective classes, and to determine which words belong to which class.

Now it has often been assumed that, across all languages, the major classes – those that are essentially unlimited in their membership – will always be the same 'big four': nouns, verbs, adjectives, and adverbs. But we now know that this is untenable when we consider the cross-linguistic evidence. Many languages lack an open adverb class (Hengeveld 1992), making do with other forms of modification. There are also languages like Lao with no adjective class, encoding property concepts as a sub-sub-type of verbs (Enfield 2004).

If a language jettisons adjectives and adverbs, the last stockade of word-class difference is that between nouns and verbs. Could a language abolish this, and just have a single word-class of predicates (like predicate calculus)? Here controversy still rages among linguists as the bar for evidence of single-class languages keeps getting raised, with some purported cases (e.g. Mundari) falling by the wayside (Evans & Osada 2005). For many languages of the Philippines and the Pacific North-west Coast, the argument has run back and forth for nearly a century, with the relevant evidence becoming ever more subtle, but still no definitive consensus has been reached.

A feeling for what a language without a noun-verb distinction is like comes from Straits Salish. Here, on the analysis by Jelinek (1995), all major-class lexical items simply function as predicates, of the type 'run', 'be_big', or 'be_a_man'. They then slot into various clausal roles, such as argument ('the one such that he runs'), predicate ('run(s)') and modifier ('the running (one)'), according to the syntactic slots they are placed in. The single open syntactic class of predicate includes words for events, entities, and qualities. When used directly as predicates, all appear in clause-initial position, followed by subject and/or object clitics. When used as arguments, all lexical stems are effectively converted into relative clauses through the use of a determiner, which must be employed whether the predicate-word refers to an event (the [ones who] sing), an entity (the [one which is a] fish), or even a proper name (the [one which] is Eloise). The square-bracketed material shows what we need to add to the English translation to convert the reading in the way the Straits Salish structure lays out.

There are thus languages without adverbs, languages without adjectives, and perhaps even languages without a basic noun-verb distinction. In the other direction, we now know that there are other types of major word-class – ideophones, positionals and coverbs – that are unfamiliar to Indo-European languages.

'Ideophones' typically encode cross-modal perceptual properties – they holophrastically depict the sight, sound, smell or feeling of situations in which the event and its participants are all rolled together into an undissected gestalt. They are usually only loosely integrated syntactically, being added into narratives as independent units to spice up the colour. Examples from Mundari (Osada 1992) are *ribuy-tibuy* 'sound, sight or motion of a fat person's buttocks rubbing together as they walk', and *rawa-dawa* 'the sensation of suddenly realizing you can do something reprehensible, and no-one is there to witness it'. Often ideophones have special phonological characteristics, such as vowel changes to mark changes in size or intensity, special reduplication patterns, and unusual phonemes or tonal patterns. (Note that English words like *willy-nilly* or *heeby-jeebies* may seem analogous but they differ from ideophones in all being assimilated to other pre-existing word classes, here adverb and noun.)

'Positionals' describe the position and form of persons and objects (Ameka and Levinson 2007). These are widespread in Mayan languages (Brown 1994; Bohnemeyer & Brown 2007;

England 2001; 2004). Examples from Tzeltal include *latz'al* 'of flat items, arranged in vertical stack', *chepel* 'be located in bulging bag', etc. Positionals typically have special morphological and syntactic properties.

'Coverbs' are a further open class outside the 'big four'. Such languages as Kalam (PNG, Pawley 1993) or the Australian language Jaminjung (Schultze-Berndt 2000) have only around 20-30 inflecting verbs, but form detailed event-descriptors by combining inflecting verbs with an open class of coverbs. Unlike positionals or ideophones, coverbs are syntactically integrated with inflecting verbs, with which they cross-combine in ways that largely need to be learned individually. In Jaminjung, for example, the coverb *dibird* 'wound around' can combine with *yu* 'be' to mean 'be wound around', and with *angu* 'get/handle' to mean 'tangle up'. (English 'light verbs', as in *take a train* or *do lunch*, give a feel for the phenomenon, but of course *train* and *lunch* are just regular nouns.)

'Classifiers' are yet another word class unforeseen by the categories of traditional grammar – whether 'numeral classifiers' in East Asian and Mesoamerican languages that classify counted objects according to shape, or the handshape classifiers in sign languages which represent the involved entity through a schematized representation of its shape. And further unfamiliar word classes are continuously being unearthed which respect only the internal structural logic of previously undescribed languages. Even when typologists talk of 'ideophones', 'classifiers' and so forth, these are not identical in nature across the languages that exhibit them – rather we are dealing with family-resemblance phenomena: no two languages have any word classes that are exactly alike in morphosyntactic properties or range of meanings (Haspelmath 2007).

Once again, then, the great variability in how languages organize their word-classes dilutes the plausibility of the innatist UG position. Just which word classes are supposed to be there in the learning child's mind? We would need to postulate a start-up state with an ever-longer list of initial categories (adding ideophones, positionals, coverbs, classifiers and so forth), many of which will never be needed. And, since syntactic rules work by combining these word-class categories – 'projecting' word-class syntax onto the larger syntactic assemblages that they head – each word class we add to the purported universal inventory would then need its own accompanying set of syntactic constraints.

2.2.5 Semantics. There is a persistent strand of thought, articulated most forcefully by Fodor (1975), that languages directly encode the categories we think in, and moreover that these constitute an innate, universal 'language of thought' or 'mentalese'. As Pinker (1994, p. 82) put it "Knowing a language, then, is knowing how to translate mentalese into strings of words and vice versa. People without a language would still have mentalese, and babies and many nonhuman animals presumably have simpler dialects". Learning a language, then, is simply a matter of finding out what the local clothing is for universal concepts we already have (Li & Gleitman 2002).

The problem with this view is that languages differ enormously in the concepts that they provide ready-coded in grammar and lexicon. Languages may lack words or constructions corresponding to the logical connectives 'if' (Guugu Yimithirr) or 'or' (Tzeltal), or 'blue' or 'green' or 'hand' or 'leg' (Yélî Dnye). There are languages without tense, without aspect, without numerals, or without third-person pronouns (or even without pronouns at all, in the case of most sign languages). Some languages have thousands of verbs, others only have thirty (Schultze-Berndt 2000). Lack of vocabulary may sometimes merely make expression more cumbersome, but sometimes it will effectively limit expressibility, as in the case of languages without numerals (Gordon 2004).

In the other direction, many languages make semantic distinctions we certainly would never think of making. So Kiowa, instead of a plural marker on nouns, has a marker that means roughly 'of unexpected number': on an animate noun like 'man' it means 'two or more', on a word like 'leg' it means 'one or more than two', while on 'stone' it means 'just two' (Mithun 1999, p. 81). In many languages, all statements must be coded (e.g. in verbal affixes) for the sources of evidence, for example in Central Pomo, whether I saw it, perceived it in another modality (tactile, auditory), was told about it, inferred it, or know that it is an established fact (ibid., p.181). Kwakwala insists on referents being coded as visible or not (Anderson & Keenan 1985). Athabaskan languages are renowned for their classificatory verbs, forcing a speaker to decide between a dozen categories of objects (e.g liquids, rope-like objects, containers, flexible sheets) before picking one of a set of alternate verbs of location, giving, handling, etc. (Mithun 1999, p. 106ff). Australian languages force their speakers to pay attention to intricate kinship relations between participants in the discourse – in many to use a pronoun you must first work out whether the referents are in even- or odd-numbered generations with respect to one another, or related by direct links through the male line. On

top of this, many have special kin terms that triangulate the relation between speaker, hearer and referent, with meanings like 'the one who is my mother and your daughter, you being my maternal grandmother' (Evans 2003b).

Spatial concepts are an interesting domain to compare languages in, since spatial cognition is fundamental to any animal, and if Fodor is right anywhere, it should be here. But in fact we find fundamental differences in the semantic parameters languages use to code space. For example, there are numerous languages without notions of 'left of', 'right of', 'back of', 'front of' – words meaning 'right hand' or 'left hand' are normally present, but don't generalize to spatial description. How then does one express, for example, that the book you are looking for is on the table left of the window? In most of these languages by saying that it lies on the table north of the window – that is, by using geographic rather than egocentric coordinates. Research shows that speakers remember the location in terms of the coordinate system used in their language, not in terms of some fixed, innate mentalese (see Levinson 2003; Majid et al. 2004).

Linguists often distinguish between closed-class or function words (like *the, of, in,* which play a grammatical role) and open-class items or general vocabulary which can be easily augmented by new coinages or borrowing. Some researchers claim that closed-class items reveal a recurrent set of semantic distinctions, while the open-class items may be more culture-specific (Talmy 2000). Others claim effectively just the reverse, that relational vocabulary (as in prepositions) is much more abstract, and thus prone to cultural patterning, while the open-class items (like nouns) are grounded in concrete reality, and thus less cross-linguistically variable (Gentner & Boroditsky 2001). In fact, neither of these views seems correct, for both ends of the spectrum are cross-linguistically variable. Consider for example the difference between nouns and spatial prepositions. Landau & Jackendoff (1993) claimed that this difference corresponds to the nature of the so-called 'what' vs. 'where' systems in neurocognition: nouns are 'whaty' in that their meanings code detailed features of objects, while prepositions are 'wherey' in that they encode abstract, geometric properties of spatial relations. They thus felt able to confidently predict that there would be no preposition or spatial relator encoding featural properties of objects, e.g. none meaning 'through a cigar-shaped object' (ibid, p.226). But the Californian language Karuk has precisely such a spatial verbal prefix, meaning 'in through a tubular space' (Mithun 1999, p. 142)! More systematic examination of the inventories of spatial pre- and postpositions shows that there is no simple

universal inventory, and the meanings can be very specific, e.g. 'in a liquid', 'astraddle', 'fixed by spiking' (Levinson & Meira 2003) – or distinguish 'to (a location below)' vs. 'to (a location above)' vs. 'to (a location on a level with the speaker)'.

Nor do nouns always have the concrete sort of reference we expect – for example, in many languages nouns tend to have a mass or 'stuff' like reference (meaning e.g. any stuff composed of banana genotype, or anything made of wax), and don't inherently refer to bounded entities. In such languages, it takes a noun and a classifier (Lucy 1992), or a noun and a classificatory verb (Brown 1994), to construct a meaning recognizable to us as 'banana' or 'candle'.

In the light of examples like these, the view that "linguistic categories and structures are more or less straightforward mappings from a pre-existing conceptual space programmed into our biological nature" (Li & Gleitman 2002, p. 266) looks quite implausible. Instead languages reflect cultural preoccupations and ecological interests that are a direct and important part of the adaptive character of language and culture.

3. Linguistic Universals

The prior sections have illustrated the surprising range of cross-linguistic variability at every level of language, from sound to meaning. The more we discover about languages, the more diversity we find. Clearly, this ups the ante in the search for universals.

There have been two main approaches to linguistic universals. The first, already mentioned, is the Chomskyan approach, where UG denotes structural principles which are complex and implicit enough to be unlearnable from finite exposure. Chomsky thus famously once held that language universals could be extracted from the study of a single language:

> I have not hesitated to propose a general principle of linguistic structure on the basis of observation of a single language. The inference is legitimate, on the assumption that humans are not specifically adapted to learn one rather than another human language ... Assuming that the genetically determined language faculty is a common human possession, we may conclude that a principle of language is universal if we are led to postulate it as a 'precondition' for the acquisition of a single language." (Chomsky 1980, p. 48)[2]

Chomsky (1965, p. 27-30) influentially distinguished between *substantive* and *formal* universals.Substantive universals are drawn from a fixed class of items, e.g. distinctive phonological features, or word classes like noun, verb, adjective, and adverb. No particular language is required to exhibit any specific member of a class. Consequently, the claim that property X is a substantive universal cannot be falsified by finding a language without it, since the property is not required in all of them. Conversely, suppose we find a new language with property Y, hitherto unexpected: we can simply add it to the inventory of substantive universals. Jackendoff (2002, p. 263) nevertheless holds "the view of Universal Grammar as a 'toolkit'… : beyond the absolute universal bare minimum of concatenated words… languages can pick and choose which tools they use, and how extensively". But without limits on the toolkit, UG is unfalsifiable.

Formal universals specify abstract constraints on the grammar of languages, e.g. that they have specific rule types, or cannot have rules that perform specific operations. To give a sense of the kind of abstract constraints in UG, consider the proposed constraint called Subjacency (see Newmeyer 2004, p. 537ff). This is an abstract principle meant to explain the difference between the grammaticality of the sentence (6) and (7), below, vs. the ungrammaticality (marked by an asterisk) of sentence (8):

(6) Where did John say *that* we had to get off the bus?
(7) Did John say *whether* we had to get off the bus?
(8) *Where did John say *whether* we had to get off the bus?

The child somehow has to extrapolate that (6) and (7) are OK, but (8) isn't, without ever being explicitly told that (8) is ungrammatical. This induction is argued to be impossible, necessitating an underlying and innate principle that forbids the formation of *wh*-questions if a *wh*-phrase intervenes between the 'filler' (initial *wh*-word) and the 'gap' (the underlying slot for the *wh*-word). This presumes a movement rule pulling a *wh*-phrase out of its underlying position and putting it at the front of the sentence as shown in (9):

(9) **Where* did John say *whether* we had to get off the bus ____?

However, it turns out that this constraint does not work in Italian or Russian in the same way, and theorists have had to assume that children can learn the specifics of the constraint after all, although we do not know how (Newmeyer 2004; Van Valin & LaPolla 1997, p. 615ff). This shows the danger of extrapolations from a single language to unlearnable constraints. Each constraint in UG needs to be taken as no more than a working hypothesis, to be sufficiently clearly articulated that it could be falsified by cross-linguistic data.

But what counts as falsification of these often abstract principles? Consider the so-called Binding Conditions, proposed as elements of Universal Grammar in the 1980s (see Koster & May 1982). One element (Condition A) specifies that *anaphors (reflexives and reciprocals) must be bound in their governing category*, while a second (Condition B) states that *(normal nonreflexive) pronouns must be free in their governing category*. These conditions were proposed to account for the English data in (10a-c) and comparable data in many other languages (the subscripts keep track of what each term refers to). The abstract notion of 'bound' is tied to a particular type of constituent-based syntactic representation where the subject 'commands' the object (owing to its position in a syntactic tree) rather than the other way round, and reflexives are sensitive to this command. Normal pronouns pick up their reference from elsewhere and so cannot be used in a 'bound' position.

(10a) John$_x$ saw him$_y$. (disjoint reference)
(10b) John$_x$ saw himself$_x$ (conjoint reference)
(10c) *Himself$_x$ saw John$_x$ / him$_x$.

This works well for English and hundreds, perhaps thousands, of other languages, but does not generalize to languages where you get examples as in (11a-b) (to represent their structures in a pseudo-English style).

(11a) He$_x$ saw him$_{x,y}$
(11b) They$_{x,y}$ saw them$_{a,b/x,y/y,x}$.

Lots of languages (even Old English, see Levinson 2000a) allow sentences like (11a, 11b): the same pronouns can either have disjoint reference (shown as 'a,b'), conjoint reference ('x,y') or commuted conjoint reference ('y,x', corresponding to 'each other' in English). Does this falsify the Binding Principles? Not necessarily, would be a typical response in the

generativist position – it may be that there are really two distinct pronouns (a normal pronoun and a reflexive, say) which just happen to have the same form, but can arguably be teased apart in other ways (see e.g. Chung 1989 on Chamorro). But it is all too easy for such an abstract analysis to presuppose precisely what is being tested, dismissing seeming counterexamples and rendering the claims unfalsifiable. The lack of shared rules of argumentation means that the field as a whole has not kept a generally accepted running score of which putative universals are left standing.

In short it has proven extremely hard to come up with even quite abstract generalizations which don't run afoul of the cross-linguistic facts. This doesn't mean that such generalizations won't ultimately be found, nor that there are no genetic underpinnings for language – there certainly are.[3] But, to date, strikingly little progress has been made.

We turn now to the other approach to universals, stemming from the work of Greenberg (1963a,b) which directly attempts to test linguistic universals against the diversity of the world's languages. Greenberg's methods crystallized the field of *linguistic typology*, and his empirical generalizations are sometimes called Greenbergian universals.

First, importantly, he discounted features of language that are universal by definition – that is, we would not call the object in question a *language* if it lacked these properties (Greenberg et al. 1963, p. 73). Thus many of what Hockett (1963) called 'design features' of language are excluded – e.g. discreteness, arbitrariness, productivity, and the duality of patterning achieved by combining meaningless elements at one level (phonology) to construct meaningful elements (morphemes or words) at another.[4] We can add other functional features that all languages need in order to be adequately expressive instruments, e.g. the ability to indicate negative or prior states of affairs, to question, to distinguish new from old information, etc. Secondly, Greenberg distinguished the following *types* of universal statement (the terminology may differ slightly across sources):

	Absolute (exceptionless)	Statistical (tendencies)
Unconditional (unrestricted)	Type 1. "unrestricted absolute universals" *All languages have property X*	Type 2. "unrestricted tendencies" *Most languages have property X*
Conditional (restricted)	Type 3. "exceptionless implicational universals" *If a language has property X, it also has property Y*	Type 4. "statistical implicational universals" *If a language has property X, it will tend to have property Y*

Table 1 Logical types of universal statement (following Greenberg)

Though all of these types are universals in the sense that they employ universal quantification over languages, their relations to notions of 'universal grammar' differ profoundly.

Type 1 statements are true of all languages, though not tautological by being definitional of languagehood. This is the category which cognitive scientists often imagine is filled by rich empirical findings from a hundred years of scientific linguistics – indeed Greenberg (1986, p. 14) recollects how Osgood challenged him to produce such universals, saying that these would be of fundamental interest to psychologists. This started Greenberg on a search that ended elsewhere, and he rapidly came to realize "the meagreness and relative triteness of statements that were simply true of all languages" (Greenberg 1986, p. 15):

"Assuming that it was important to discover generalizations which were valid for all languages, would not such statements be few in number and on the whole quite banal? Examples would be that languages had nouns and verbs (although some linguists denied even that) or that all languages had sound systems and distinguished between phonetic vowels and consonants" (Greenberg 1986, p.14).

To this day, the reader will find no agreed list of Type 1 universals (see Box 1). This more or less empty box is why the emperor of Universal Grammar has no clothes. Textbooks such as Comrie (1989), Whaley (1997), and Croft (2003) are almost mum on the subject, and what they do provide is more or less the same two or three examples. For the longest available list

of hypotheses, see the online resources at the Konstanz Universals Archive (**http://ling.uni-konstanz.de:591/Universals**).

The most often cited absolute unrestricted universals are that all languages distinguish nouns and verbs (discussed above) and that all languages have vowels. The problem with 'all languages have vowels' is that it does not extend to sign languages (Box 2), as already mentioned. A second problem is that, for spoken languages, if the statement is taken at a phonetic level, it is true, but for trivial reasons: they would otherwise scarcely be audible. A third problem is that, if taken as a phonological claim that all languages have distinctive vowel segments, it is in fact contested: there are some languages, notably of the Northwestern Caucasus, where the quality of the vowel segments was long maintained by many linguists to be entirely predictable from the consonantal context (see Kuipers 1960; Halle 1970; Colarusso 1982), and although most scholars have now swung round to recognizing two contrasting vowels, the evidence for this hangs on the thread of a few minimal pairs, mostly loanwords from Turkish or Arabic.

This example illustrates the problems with making simple, interesting statements that are true of all languages. Most straightforward claims are simply false – see Box 1. The fact is that it is a jungle out there: languages differ in fundamental ways – in their sound systems (even whether they have one), in their grammar, and in their semantics. Thus the very type of universal that seems most interesting to psychologists was rapidly rejected as the focus of research by Greenberg.

Linguistic typologists make a virtue out of the necessity to consider other kinds of universals. Conditional or implicational universals of Types 3 and 4 (i.e. of the kind 'If a language has property X, it has (or tends to have) property Y') allow us to make claims about the interrelation of *two, logically independent parameters*. Statements of this kind, therefore, greatly restrict the space of possible languages: interpreted as logical (material) conditionals, they predict that there are no languages with X that lack Y, where X and Y may not be obviously related at all. Here again however exceptionless or absolute versions are usually somewhat trite. For example, the following seem plausible:

(12a) IF a language has nasal vowels, THEN it has oral vowels

(12b) IF a language has a trial number, THEN there is also a dual. IF there is a dual, THEN there is also a plural.

Statement (12a) essentially expresses the *markedness* (or recessive character) of nasal vowels. However, most markedness universals are statistical, not absolute. Statement (12b) is really only about one parameter, namely number, and it is not really surprising that a language that morphologically marks pairs of things, would want to be able to distinguish singular from plural or trial, i.e. more than two. Nevertheless, there's at least one language that counter-exemplifies: Basic verbs stems in Nen are dual, with non-duals indicated by a suffix meaning 'either singular or three-or-more', the singular and the plural sharing an inflection!

But the main problem with absolute conditional universals is that, again and again (as just exemplified), they too have been shown to be false. In this sense conditional universals follow the same trajectory as unconditional ones, in that hypothesized absolute universals tend to become statistical ones as we sample languages more widely. For example, it was hypothesized as an unconditional universal (Greenberg 1966, p. 50) that all languages mark the negative by adding some morpheme to a sentence, but then we find that Classical Tamil marks the negative by deleting the tense morphemes present in the positive (Master 1946; Pederson 1993). We can expect the same general story for conditional universals, except that, given the conditional restriction, it will take a larger overall database to falsify them.

Again making a virtue out of a necessity, Dryer (1998) convincingly argues that statistical universals or strong tendencies are more interesting anyway. Although at first sight it seems that absolute implications are more easily falsifiable, the relevant test set is after all not the 7000 odd languages we happen to have now, but the half million or so that have existed, not to mention those yet to come – since we never have all the data in hand, the one counterexample might never show up. In fact, he points out, since linguistic types always empirically show a clustering with outliers, the chances of catching all the outliers are vanishingly small. The Classical Tamil counterexample to negative marking strategies is a case in point: it is a real counterexample, but extremely rare. Given this distribution of phenomena, the methods have to be statistical. And as a matter of fact, nearly all work done in linguistic typology concerns Type 4 Universals, i.e. conditional tendencies. Where these tendencies are weak, they may reveal only bias in the current languages we have, or the

sampling methods employed. But where they are strong, they suggest that there is indeed a cognitive, communicative or system-internal bias towards particular solutions evolving.

With absolute universals, sampling is not an issue: just a single counter-example is needed, and linguists should follow whatever leads they need to find them. For this reason, and because many of the claimed universals we are targeting are absolute, we have not shied away in this article from hand-picking the clearest examples that illustrate our point. But with statistical universals, having the right sampling methods is crucial (Widmann and Bakker 2006), and many factors need to be controlled for. Language family (coinherited traits are not independent), language area (convergent traits are not independent), key organisational features (dominant phrase orders have knock-on effects elsewhere), other cultural aspects (speaker population size, whether there is a written language), modality (spoken vs. signed language), and quality of available descriptions all impact on the choice. Employing geographically separate areas is crucial to minimise the risk of convergent mutual influence, but even this is contingent on our current very limited understanding of the higher-level phylogenetic relationships of the world's languages: if languages in two distinct regions (say inland Canada and Central Siberia) are found to be related, we can no longer assume these two areas supply independent samples. The longterm and not unachievable goal must be to have data on all existing languages, which should be the target for the language sciences. Where do linguistic universals, of whatever type, come from? We will return to this issue in §6, but here it is vital to point out that *a property common to languages need not have its origins in a 'language faculty'*, or innate specialization for language. First, such a property could be due to properties of other mental capacities – memory, action control, sensory integration, etc. Second, it could be due to overall design requirements of communication systems. For example, most languages seem to distinguish closed-class functional elements (cf. English *the, some, should*) from open-class vocabulary (cf. *eat, dog, big*), just as logics distinguish operators from other terms, allowing constancies in composition with open-ended vocabularies and facilitating parsing.

Universals can also arise from so-called functional factors, that is to say, the machining of structure to fit the uses to which it would be put. For example, we can ask: Why are negatives usually marked in languages with a positive 'not' morpheme rather than by a gap as in Classical Tamil? Because (a) we make more positive than negative assertions, so it's more

efficient to mark the less common negatives, and (b) it's crucial to distinguish what is said from its contrary, and a non-zero morpheme is less likely to escape notice than a gap.
In addition, given human motivations, interests and sensory perception together with the shared world we live in, we can expect all sorts of convergences in, for example, vocabulary items – most if not all languages have kin terms, body part terms, words for celestial bodies. The appeal to innate concepts and structure should be a last resort (Tomasello 1995).

Finally, a word needs to be said about the metalanguage in which typological (statistical) universals are couched. The terms employed are notions like subject, adjective, inflection, syllable, pronoun, noun phrase and the like – more or less the vocabulary of 'traditional grammar'. As we have seen, these are not absolute universals of Type 1. Rather, they are descriptive labels, emerging from structural facts of particular languages, which work well in some languages but may be problematic or absent in others (cf. Croft 2001). Consequently, for the most part they do not have precise definitions shared by all researchers, or equally applicable to all languages (Haspelmath 2007). Does this vitiate such research? Not necessarily: the descriptive botanist also uses many terms ('pinnate', 'thorn', etc.) that have no precise definition. Likewise, linguists use notions like 'subject' (§4) in a prototype way: a prototypical subject has a large range of features (argument of the predication, controller of verb agreement, topic, etc.) which may not all be present in any particular case. The 'family resemblance' character of the basic metalanguage is what underlies the essential nature of typological generalizations, namely that of soft regularities of association of traits.

4. How multiple constraints drive multiple solutions: grammatical subject as a great (but not universal) idea

We can use the notion of grammatical subject to illustrate the multi-constraint engineering problems languages face, the numerous independent but convergent solutions which cluster similar properties, and at the same time the occurrence of alternative solutions in a minority of other languages that weight competing design motivations differently.

The 'grammatical relations' of subject and object apply unproblematically to enough unrelated languages that Baker (2003) regards them as part of the invariant machinery of universal grammar. Indeed, many languages around the world have grammatical relations

that map straightforwardly onto the clusterings of properties familiar from English 'subject' and 'object'. But linguists have also known for some time that the notion 'subject' is far from universal, and other languages have come up with strikingly different solutions.

The device of subject, whether in English, Warlpiri, or Malagasy, is a way of streamlining grammars to take advantage of the fact that three logically distinct tasks correlate statistically. In a sentence like 'Mary is trying to finish her book' the subject 'Mary' is:

(a) a topic – what the sentence is about;
(b) an agent – the semantic role of the instigator of an action;
(c) the 'pivot' – the syntactic broker around which many grammatical properties coalesce

Having a subject relation is an efficient way to organize a language's grammar because it bundles up different subtasks that most often need to be done together. But languages also need ways to indicate when the properties do not coalesce. For example, when the subject is not an agent this can be marked by the passive: *John was kissed by Mary*.

'Subject' is thus a fundamentally useful notion for the analysis of many, probably most, languages. But when we look further we find many languages where the above properties don't line up, and the notion 'subject' can only be applied by so weakening the definition that it is near vacuous. For example, the semantic dimension of case role (agent, patient, recipient etc.) and the discourse dimension of topic can be dissociated, with different grammatical mechanisms assigned to deal with each in a dedicated way: this is essentially how Tagalog works (Schachter 1976). Or a language may use its case system to reflect semantic roles more transparently, so that basic clause types have a plethora of different case arrays, rather than funnelling most event types down to a single transitive type, as in the Caucasian language Lezgian (Haspelmath 1993). Alternatively, a language may split the notion subject by funnelling all semantic roles into two main 'macro-roles' – 'actor' (a wider range of semantic roles than agent) and 'undergoer' (corresponding to e.g. the subject of English *John underwent heart surgery*). The syntactic privileges we normally associate with subjects then get divided between these two distinct categories (as in Acehnese, Durie 1985).

Finally, a language may plump for the advantages of rolling a wide range of syntactic properties together into a single syntactic broker or 'pivot', but go the opposite way to English, privileging the patient over the agent as the semantic role that gets the syntactic

privileges of the pivot slot. Dyirbal (Dixon 1972; 1977) is famous for such 'syntactic ergativity'. The whole of Dyirbal's grammatical organisation then revolves around this absolutive pivot – case marking, coordination, complex clause constructions. To illustrate with coordination, take the English sentence 'The woman slapped the man and ø laughed'. The 'gap' (represented here by a zero) is interpreted by linking it to the preceding subject, forcing the reading 'and she laughed'. But in the Dyirbal equivalent *yibinggu yara bunjun ø miyandanyu*, the gap is linked to the preceding absolutive pivot *yara* (corresponding to the English object, *the man*), and gets interpreted as 'and he laughed'.

Dyirbal, then, is like English in having a single syntactic 'pivot' around which a whole range of constructions are organized. But it is unlike it in linking this pivot to the patient rather than the agent. Since this system probably strikes the reader as perverse, it is worth noting that a natural source is the fact that cross-linguistically most new referents are introduced in 'absolutive' (S or O) roles (Dubois 1987), making this a natural attractor for unmarked case and thus a candidate for syntactic 'pivot' status (see also Levinson in press).
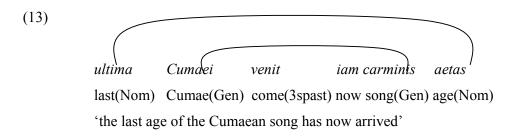
Given languages like Dyirbal, Acehnese or Tagalog, where the concepts of 'subject' and 'object' are dismembered in language specific ways, it is clear that a child pre-equipped by UG to expect its language to have a 'subject' could be sorely led astray.

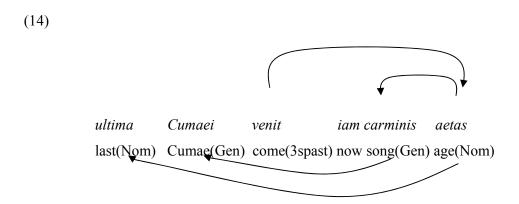## 5. The claimed universality of constituency

 In nearly all recent discussions of syntax for a general cognitive science audience, it is simply presumed that the syntax of natural languages can basically be expressed in terms of constituent structure, and thus the familiar tree diagrams for sentence structure (Pinker 1994, p. 97ff; Jackendoff 2002; 2003a; Hauser, Chomsky & Fitch 2002).

In the recent debates following Hauser et al. (2002), there is sometimes a conflation between constituent structure and recursion (see for example Pinker & Jackendoff 2005, p. 215), but they are potentially orthogonal properties of languages. There can be constituent structure without recursion, but there can also be hierarchical relations and recursion without constituency. We return to the issue of recursion in the next section, but here we focus on constituency.

Constituency is the bracketing of elements (typically words) into higher order elements (as in *[[[[the][tall [man]]] [came]]* where *[[[the][tall [man]]]* is a Noun Phrase, substitutable by a single element (*he*, or *John*). Many discussions presume that constituency is an absolute universal, exhibited by all languages. But in fact constituency is just one method, used by a subset of languages, to express constructions which in other languages may be coded as dependencies of other kinds (Matthews 1981, 2007). pThe need for this alternative perspective is that many languages show few traces of constituent structure, because they scramble the words as in the following Latin line from Virgil (Matthews 1981, p. 255):

(13)

ultima   Cumaei   venit   iam carminis   aetas
last(Nom)   Cumae(Gen) come(3spast) now song(Gen) age(Nom)
'the last age of the Cumaean song has now arrived'

Here the lines link the parts of two noun phrases, and it makes no sense to produce a bracketing of the normal sort: a tree diagram of the normal kind would have crossing lines. A better representation is in terms of *dependency* – which parts depend on which other parts, as in the following diagram where the arrowhead points to the dependent item:

(14)

ultima   Cumaei   venit   iam carminis   aetas
last(Nom)   Cumae(Gen) come(3spast) now song(Gen) age(Nom)

Classical Latin is a representative of a large class of languages, which exhibit free *word order* (not just free *phrase order*, which is much commoner still). The Australian languages are also renowned for these properties. In Jiwarli, for example, all linked nominals (part of a noun phrase if there was such a thing) are marked with case and can be separated from each other;

there is no evidence for a verb phrase, and there are no major constraints on ordering (see Austin & Bresnan 1996). (15) illustrates a discontinuous sequence of words which would correspond to a constituent in most European languages; 'the woman's dog' is grouped as a single semantic unit by sharing the accusative case.

(15) *Kupuju-lu* **kaparla-nha** *yanga-lkin* **wartirra-ku-nha**
child-ERG dog-ACC chase-PRES woman-DAT-ACC
'The child chases the woman's dog.' (Austin 1995, p. 372)

Note how possessive modifiers – coded by a special use of the dative case – additionally pick up the case of the noun they modify, as with the accusative *–nha* on 'dog' and 'woman-Dat' in (15). In this way multiple case marking (Dench & Evans 1988) allows the grouping of elements from distinct levels of structure, such as embedded possessive phrases, even when they are not contiguous. It is this case-tagging, rather than grouping of words into constituents, which forms the basic organizational principle in many Australian languages (see Nordlinger 1998 for a formalization).

It is even possible in Jiwarli to intermingle words that in English would belong to two distinct clauses, since the case suffixes function to match up the appropriate elements. These are tagged, as it were, with instructions like 'I am object of the subordinate clause verb', or 'I am a possessive modifier of an object of a main clause verb'. By fishing out these distinct cases, a hearer can discern the structure of a two-clause sentence like 'the child (ERG) is chasing the dog (ACC) of the woman (DAT-ACC) who is sitting down cooking meat (DAT)' without needing to attend to the order in which words occur (Austin & Bresnan 1996). The syntactic structure here is most elegantly represented via a dependency formalism (supplemented with appropriate morphological features) rather than a constituency one.

Though languages like Jiwarli have been increasingly well documented over the last forty years, syntactic theories developed in the English-speaking world have primarily focussed on constituency, no doubt because English fits this bill. In the Slavic world, by contrast, where languages like Russian have a structure much more like Jiwarli or Latin, models of syntactic relations have been largely based on dependency relations (Melčuk 1988). The most realistic view of the world's languages is that some yield completely to one representational system, some to the other, most to a mix. Some outgrowths of generative theory, such as Lexical-

Functional Grammar (LFG), effectively incorporate analogues of dependency representations alongside constituency-based ones, in the form of f-structures besides c-structures, with an interface system linking the two structures (see Bresnan 2001; Hudson 1993, p. 329). It is also worth emphasising, at this point, that dependency-based representations are just as capable of expressing recursive structure as constituency-based ones are.

A way of saving the claimed primacy of word order and constituency would be to impose an English-like structure on a sentence like the Latin one above (reordered, say, as
 [[[[ultima][aetas]] [[carminis][Cumae]]] [[iam] [venit]]]) and then to scramble the words with a secondary operation (see Matthews 2007 for critical review). A more sophisticated variant is to separate out the hierarchical from the ordering information and specify them separately in a differently construed version of a Phrase Structure Grammar (Gazdar & Pullum 1982). But the point is that order and constituency are playing no signalling role for the hearer – they cannot therefore play a role in the parsing of such a sentence. In all the recent applications in the cognitive sciences mentioned above, where recursion has played such an important theoretical role, the experimental evidence was from a comprehension or parsing perspective where the universality of constituency was assumed (Fitch & Hauser 2004; Friederici 2004). A further point is that there is not the slightest evidence for the *psychological reality* of any such imposed constituent structure in a language like Jiwarli. (Researchers on Australian languages have repeatedly reported the inability of speakers to repeat a sentence with the same word order: for Warlpiri 'sentences containing the same content words in different linear arrangements count as repetitions of one another' (Hale 1983, p. 5) and '[w]hen asked to repeat an utterance, speakers depart from the ordering of the original more often than not' (Hale et al. 1995, p. 1431)).

Syntactic constituency, then, is not a universal feature of languages.[5] Just like dependency relations, it is simply one possible way to mark relationships between the parts of a sentence. Just like the grammatical relation of subject (§5), employing constituency as a coding device is a common and workable solution that many languages have evolved, but it is totally absent in others, while in others again it is in the process of evolving without having yet quite crystallized (Himmelmann 1997).

It follows that any suggestion about UG which presumes the universality of constituent structure will be false. Models of the evolution of language (e.g. Bickerton 1981, etc.) that

presume the operation of phrase-structure grammar (PSG) generating sentences with surface constituency (Hauser, Chomsky & Fitch 2002, p. 1577) are also therefore aimed at a particular kind of (English-like) language as the target of evolutionary development. But it is clear that the child must be able to learn (at least) both types of system, constituency or dependency. It will not always be the case that the child needs to use constituency-detecting abilities in constructing its grammar, since constituency relations are, as shown, not universal.

6. Recursion in syntax as a non-universal

We turn now to recursion, the feature which is at the heart of recent heated discussions: indeed Hauser, Chomsky & Fitch 2002, p. 1569 hypothesize that recursion is "the only uniquely human component of the faculty of language".[6] Recent findings are said to show that "animals lack the capacity to create open-ended generative systems", whereas human "languages go beyond purely local structure by including a capacity for recursive embedding of phrases within phrases" (Hauser, Chomsky & Fitch 2002, p. 1577). Recursion, in syntax, is commonly defined as the looping back into a set of rules of its own output, so as to produce a potentially infinite set of outputs. It is sometimes assumed in the debate that recursion is defined over constituent structure, in that recursion "consists of embedding a constituent in a constituent of the same type" (Pinker & Jackendoff 2005, p. 211). However, since dependency structures are also generated by rule, it is equally possible to have recursive structures that employ dependency relations rather than constituency structures (Levelt 2008, II:134ff).

The terms of the debate were set fifty years ago by Chomsky (1955; 1957) when he introduced the hierarchy of formal languages, using methods from logic and mathematics, and applied them to constituent structure. He showed that English constituent structure could not be generated by a grammar limited to state transitions (a finite state grammar or FSG). Rather, the indefinitely embedded structures of English required at least a phrase-structure grammar or PSG, as in "If A, then B", where A itself could be of the form "X and Y", and "X" of the form "W or Z". Taken as a whole, this generates structures like *If John comes or Mary comes and Bill agrees, let's go to the movies*. Chomsky has consistently held that this recursion in constituent structure is the magic ingredient in language, which gives it its expressive power.

Since then a vast amount of work in theoretical linguistics has elaborated on the mathematical properties of abstract grammars (Partee et al. 1990; Gazdar et al. 1985), while many non-Chomskyan linguistic theories have moved beyond this syntactic focus, developing models of language that reapportion generativity to other components of grammar (see for example: Bresnan 2001; Jackendoff 2003a; 2003b). But recently this classic "syntactocentricism" (as Jackendoff (2003b) has called it), never relinquished by Chomsky, has re-emerged centrally in interdisciplinary discussions about the evolution of language, re-enlivened by Hauser, Chomsky & Fitch's (2002) proposal that the property of recursion over constituent structures represents the only key design feature of language that is unique to humans: sound systems and conceptual systems (which provide the semantics) are found in other species. Fitch & Hauser (2004) have gone on to show that despite impressive learning powers over FSGs, tamarin monkeys don't appear to be able to grasp the patterning in PSG-generated sequences, while O'Donnell, Hauser & Fitch (2005) argue that comparative psychology should focus on these formal features of language. Meanwhile Friederici (2004), on the basis of these developments, suggests different neural systems for processing FSG vs. PSGs, which she takes to be the critical juncture in the evolution of human language.

In this context where recursion has been suggested to be *the* criterial feature of the human language capacity, it is important for cognitive scientists to know that many languages show distinct limits on recursion in this sense, or even lack it altogether.

First, many languages are structured to minimize embedding. For example, polysynthetic languages – which typically have extreme levels of morphological complexity in their verb, but little in the way of syntactic organization at the clause level or beyond – show scant evidence for embedding.

In Bininj Gun-wok, for example (Evans 2003a, p. 633), the doubly-embedded English sentence '[They stood [watching us [fight]]]' is expressed, without any embedding, as 'they_stood / they_were_watching_us / we_were_fighting_each_other', where underscores link morphemes within a word. In fact the clearest cases of embedding are morphological (within the word) rather than syntactic: to a limited degree one verb can be incorporated within another, for example: *barri-kanj-ngu-nihmi-re* [they-meat-eat-ing-go] 'they go along eating meat'. But this construction has a maximum of one level of embedding – so that even

if it were claimed that polysynthetic languages simply shift the recursive apparatus out of the syntax into the morphology (Baker 1988), the limit to one degree of embedding means it can be generated by a finite state grammar. Mithun (1984) counted the percentage of subordinate clauses (embedded or otherwise) in a body of texts for three polysynthetic languages and found very low levels in all three: 7% for oral Mohawk texts, 6% in Gunwinggu (a dialect of Bininj Gun-wok), and just 2% in Kathlamet. Examples like this show how easily a language can dispense with subordination (and hence with the primary type of recursion), by adopting strategies that present a number of syntactically independent propositions whose relations are worked out pragmatically.

Kayardild is another interesting case of a language whose grammar allows recursion, but caps it at one level of nesting (see Evans 1995a; 1995b). Kayardild forms subordinate clauses in two ways: either it can nominalize the subordinate verb (something like English *-ing*), or it can use a finite clause for the subordinate clause. Either way, it makes special use of a case marking – the oblique (OBL) – which can go on all or most clausal constituents. This oblique case marker then stacks up outside any other case markers that may already be there independently. We'll illustrate with the nominalized variant, but identical arguments carry through for the finite version. For example, to say 'I will watch the man spearing the turtle', you say (16).

(16) *ngada*   *kurri-ju*   *dangka-wu*   *raa-n-ku*   *banga-wuu-nth*
      I   watch-FUT   man-OBJ   spear-NOMZR-OBJ   turtle-OBJ-OBL

The object marker on 'man' is required because it is the object of 'watch', and the object marker on 'turtle' because it is the object of 'spear'. Of particular relevance here is that 'turtle' is marked with the object case plus the Oblique case, because the verb 'spear' of which it is the object has been nominalized.

Now the interesting thing is that, even though in general Kayardild (highly unusually) allows cases to be stacked up to several levels, the oblique case has the particular limitation (found only with this one case), that it cannot be followed by any other case. This morphological restriction, combined with the fact that subordinate clauses require their objects and other non-subject NPs to be marked with an oblique for the sentence to be grammatical, means that

the morphology places a cap on the syntax: at most, one level of embedding.

In discussions of the infinitude of language, it is normally assumed that once the possibility of embedding to one level has been demonstrated, iterated recursion can then go on to generate an infinite number of levels, subject only to memory limitations. And it was arguments from the need to generate an indefinite number of embeddings that were crucial in demonstrating the inadequacy of finite-state grammars. But, as Kayardild shows, the step from one-level recursion to unbounded recursion cannot be assumed, and once recursion is quarantined to one level of nesting it is always possible to use a more limited type of grammar, such as a finite state grammar, to generate it.

The most radical case would be of a language that simply disallows recursion altogether, and an example of this has recently been given for the Amazonian language Pirahã by Everett (2005), which lacks not only subordination but even indefinitely expandable possessives like 'Ko'oi's son's daughter'. This has been widely discussed and we refer the reader to that paper for the details. Village-level sign languages of three generations depth or more also systematically show an absence of embedding (Meir et al. in press), suggesting that recursion in language is an evolved socio-cultural achievement rather than an automatic reflex of a cognitive specialism.

The languages we have reviewed, then, show that languages can employ a range of alternative strategies to render, without embedding, meanings whose English renditions normally use embedded structures. In some cases the languages do, indeed, permit embedding, but it is rare, as with Bininj Gun-wok or Kathlamet. In other cases, like Kayardild nominalized clauses, embedding is allowed, but to a maximum of one iteration. Moreover, since this is governed by clear grammatical constraints, it is not simply a matter of performance or frequency. Finally, there is at least one language, Pirahã, where embedding is impossible, both syntactically and morphologically. The clear conclusion that these languages point to is that recursion is not a necessary or defining feature of every language. It is a well-developed feature of some languages, like English or Japanese, rare but allowed in others (like Bininj Gun-wok), capped at a single level of nesting in others (Kayardild), and in others, like Pirahã, it is completely absent. Recursion, then, is a capacity languages may exhibit, not a universally present feature.

The example of Pirahã has already been raised in debate with Chomsky, Hauser & Fitch, by Pinker & Jackendoff (2005). Fitch, Hauser & Chomsky (2005) replied that "the putative absence of obvious recursion in one of these languages is no more relevant to the human ability to master recursion than the existence of three vowel languages calls into doubt the human ability to master a five- or ten-vowel language". That is, despite the fact that recursion is the "only uniquely human component of the language faculty", recursion is not an absolute universal, but just one of the design features provided by UG from which languages may draw: "as Jackendoff (2002) correctly notes, our language faculty provides us with a toolkit for building languages, but not all languages use all the tools" (2002, p. 204).

But we have already noted that the argument from *capacity* is weak. By parity of argument, every feature of every language that has ever been spoken must then be part of the language faculty or UG. This seems no more plausible than claiming that, because we can learn to ride a bicycle or read music, these abilities are part of our innate endowment. Rather, it is the ability to *learn* bicycle riding by putting together other, more basic abilities, which has to be within our capacities, not the trick itself. Besides, if syntactic *recursion* is the single core feature of language, one would surely expect it to have the strong form of a 'formal universal', a positive constraint on possible rule systems, not just an optional part of the toolkit, in the style of one of Chomsky's 'substantive universals'.

No one doubts that humans have the ability to create utterances of indefinite complexity, but there can be serious doubt about where exactly this recursive property resides, in the syntax or elsewhere. Consider that instead of saying "If the dog barks, the postman may run away" we could say "The dog might bark. The postman might run away". In the former case we have syntactic embedding. In the latter the same message is conveyed, but the "embedding" is in the discourse understanding – the semantics and the pragmatics, not the syntax. It is because pragmatic inference can deliver embedded understandings of non-embedded clauses that languages often differ in what syntactic embeddings they allow. For example, in Guugu Yimithirr there is no overt conditional – and conditionals are expressed in the way just outlined (Haviland 1979).

In these cases, the expressive power of language lies outside syntax. It is a property of conceptual structure, that is, of the semantics and pragmatics of language. This is a central problem for the "syntactocentric" models associated with Chomsky and his followers, but

less so of course for the kind of view championed by Jackendoff in these pages (2003a), where semantics or conceptual structure is also argued to have generative capacity. More specifically, the generative power would seem to lie in the semantics/pragmatics or the conceptual structure in all languages, but only in some is it also a property of the syntax.

To recapitulate:

1. Many languages do not have syntactic constituent structure. As such, they cannot have embedded structures of the kind indicated by a labelled bracketing like $[_A[_A]]$. Most of the suggestions for rule constraints (like Subjacency) in UG falsely presume the universality of constituency. The Chomsky, Hauser & Fitch vs. Pinker & Jackendoff controversy simply ignores the existence of this wide class of languages.

2. Many languages have no, or very circumscribed recursion in their syntax. That is, they do not allow embedding of indefinite complexity, and in some languages there is no syntactic embedding at all. Fitch, Hauser & Chomsky's (2005) response that this is of no relevance to their selection of syntactic recursion as the single unique design feature of human language reveals their choice to be empirically arbitrary.

3. The cross-linguistic evidence shows that although recursion may not be found in the syntax of languages, it is always found in the conceptual structure, i.e. the semantics or pragmatics – in the sense that it is always possible in any language to express complex propositions. This argues against the syntacticocentrism of the Chomskyan paradigm. It also points to a different kind of possible evidence for the evolutionary background to language, namely the investigation of embedded reasoning across our nearest phylogenetic cousins, as is required, for example, in theory of mind tasks, or spatial perspective taking. Even simple tool making can require recursive action patterning (Greenfield 1991).

7. The new synthesis: evolutionary approaches to language

A linguist who asks "Why?" must be a historian

(Haspelmath 1999, p.205)

Our message has been that the languages of the world offer a real challenge to current theory and analysis about the place of language in human cognition. From the perspective of some approaches, the message of diversity may suggest that there is no clear way forward. In fact, however, there is a growing body of work that shows exactly where the language sciences are headed, which is to tame the diversity with theories and methods that stem ultimately from the biological sciences. Evolutionary approaches, in the broadest sense, are transforming the theoretical terrain.

This work is of different kinds. In the first instance, there is a great deal of speculation, elegant theory and mathematical modelling aimed at the problem of language origins (Christiansen and Kirby 2003). Some of this is devoted to the preconditions – for example, the origin of human cooperation (Boyd & Richerson 2005; Tomasello 2000), or the properties of human interaction or theory of mind (Enfield & Levinson 2006). Other work is centrally concerned with the co-evolution of cognition and culture generally, arguing for a twin-track model in which biological and cultural evolution run partially independently, but with reciprocal interaction (Durham 1991; Levinson & Jaisson 2006). This provides a mechanism for the biological evolution of traits adaptive to cultural environments, for which the neuroanatomical foundations for language must be a prime example.[7]

Language diversity can best be understood in terms of such a twin-track model, with the diversity being largely accounted for in terms of diversification in the cultural track, in which traits evolve under similar processes to those in population genetics, by drift, lineal inheritance, recombination and hybridization. These create the population conditions in which new variants arise in separate social groups. A range of selectors – characteristics of the brain and vocal tract, constraints on the communicative channel, internal constraints within the system, and transition constraints on what can turn into what – then shape the chances of different variants catching on.

Historical linguistics is the oldest branch of scientific linguistics, with long-standing interests in lineal inheritance vs. horizontal transfer through contact and borrowing. Its greatest achievement was the development of rigorous methods for the tracking of vocabulary through descendant languages. But brand new is the application of bioinformatic techniques to linguistic material, allowing the quantitation of inheritance vs. borrowing in vocabulary (McMahon & McMahon 2006); see also Pagel et al. (2007) and Atkinson et al. (2008) for further examples of statistical phylogenetic approaches to understanding language evolution. Application of cladistics and Bayesian phylogenetics to vocabulary allow much firmer inferences about the date of divergences between languages (Gray & Atkinson 2003; Atkinson & Gray 2005). These methods can also be applied to the structural (phonological and grammatical) features of language, and this can be shown to replicate the findings based on traditional vocabulary methods (Dunn et al. 2005; Dunn et al. 2008), while potentially reaching much further back in time. These explicit methods allow the comparison between, for example, linguistic phylogeny, human genetics, and the diversification of cultural traits in some area of the world. A stunning result is that, in at least some parts of the world, linguistic traits are the most tree-like, least hybridized properties of any human population (Hunley et al. 2007), with stable linguistic groupings solidly maintained across thousands of years despite enormous flows of genes and cultural exchange across groups.

These bioinformatic methods throw new light on the nature of Greenbergian universals (Dunn et al. 2008). Using Bayesian phylogenetics, we can reconstruct family trees with the structural properties at each node, right back to the ancestral proto-language of a family. We can then ask how much these structural features are, over millennia, co-dependent, i.e. changing together, or instead evolving independently. First impressions from these new methods show that the great majority of structural features are relatively independent; only a few resemble the Greenbergian word-order conditional universals, with closely correlated state-transitions. The emerging picture, then, confirms the view that most linguistic diversity is the product of historical cultural evolution operating on relatively independent traits. On the other hand, some derived states are inherently instable and unleash chains of changes till a more stable overall state is reached. As Greenberg once put it, 'a speaker is like a lousy auto mechanic: every time he fixes something in the language, he screws up something else'.

In short, there are evolutionarily stable strategies, local minima as it were, that are recurrent solutions across time and space, such as the tendency to distinguish noun and verb roots, to

have a subject role, or mutually consistent approaches to the ordering of head and modifier, which underlie the Greenbergian statistical universals linking different features. These tendencies cannot plausibly be attributed to UG, since changes from one stable strategy to another take generations (sometimes millennia) to work through. Instead they result from myriad interactions between communicative, cognitive and processing constraints which reshape existing structures through use. A major achievement of functionalist linguistics has been to map out, under the rubric of grammaticalization, the complex temporal sub-processes by which grammar emerges as frequently-used patterns sediment into conventionalized patterns (Bybee 2000). Cultural preoccupations may push some of these changes in particular directions, such as the evolution of kinship-specific pronouns in Australia (Evans 2003b). And social factors, most importantly the urge to identify with some groups by speaking like them, and to maximize distance from others by speaking differently (studied in fine-grained detail by Labov 1980), act as an amplifier on minor changes that have arisen in the reshaping process (Nettle 1999).

Gaps in the theoretically possible design space can be explained partly by the nature of the sample (we have 7000 survivors from an estimated half million historical languages), partly by chance, partly because the biased state changes above make arriving at, or staying in, some states rather unlikely (Evans 1995b; Dunn et al. 2008).

An advantage to this evolutionary and population biology perspective is that it more readily accounts for the cluster-and-outlier pattern found with so many phenomena when a broad sample of the world's languages is plumbed. We know that 'rara' are not cognitively impossible, because there are speech communities that learn them, but it may be that the immediately preceding springboard state requires such specific and improbable collocations of rare features that there is a low statistical likelihood of such systems arising (Harris 2008).

It also accounts for common but not universal clusterings, such as grammatical subject, through the convergent functional economies outlined in §4, making an all-purpose syntactic pivot an efficient means of dealing with the statistically correlated roles of agent and discourse topic in one fell swoop. And it explains why conditional universals, as well, almost always turn out to be mere tendencies rather than absolute universals: Greenberg's word-order correlations – e.g. prepositions where verb precedes object, postpositions where verb follows object – are functionally economical. They allow the language-user to consistently

stick to just one parsing strategy, right- or left-branching as appropriate, and channel state transitions in particular directions that tend to maintain the integrity of the system. For example, where adpositions derive from verbs, if the verb follows its object it only has to stay where it is to become a post- rather than a pre-position.

The fertile research program, briefly summarized in the above paragraphs, allows us to move our explanations for the recurrent regularities in language out of the prewired mind and into the processes that shape languages into intricate social artefacts. Cognitive constraints and abilities now play a different role to what they did in the generative program (cf. Christiansen & Chater 2008): their primary role is now as stochastical selective agents that drive along the emergence and constant resculpting of language structure.

We emphasise that this view does not, of itself, provide a solution to the other great Chomskyan question: What cognitive tools must children bring to the task of language learning? If anything, this question has become more challenging in two vital respects, and here we part company with Christiansen & Chater (2008), who assume a much narrower spectrum of structural possibilities in language than we do: first, because the extraordinary structural variation sketched in this paper presents a far greater range of problems for the child to solve than we were aware of fifty years ago, and second because the child can bring practically no specific hypotheses, of the UG variety, to the task. But, however great it is, this learning challenge is not peculiar to language – it was set up as the crucial human cognitive property when we moved into a coevolutionary mode, using culture as our main means of adaptation to environmental challenges, well over a million years ago.

8. Conclusion: Seven theses about the nature of language and mind

The new and more complex emerging picture that we have sketched here, however uncomfortable it may be for models of learning that minimize the challenge by postulating some form of universal grammar, in fact promises us a much better understanding of the nature of language and the cognition that makes it possible.

On the one hand, there are thousands of diverse languages, with the organizing principles that sort them being largely similar to the radiation and diversification of species. In other words,

language diversification and hybridization works just like the evolution of biological species – it is an historical process, following the laws of population biology. Consider the fact that linguistic diversity patterns just like biological diversity generally, so that the density of languages closely matches the density of biological species across the globe, correlating with rainfall and latitude (Mace & Pagel 1995; Nettle 1999; Pagel 2000; Collard & Foley 2002). Minor genetic differences between human populations may act as 'attractors' for certain linguistic properties which are then easier to acquire and propagate (Dediu & Ladd 2007).

On the other hand, the human cognition and physiology that has produced and maintained this diversity is a single system, late evolved and shared across all members of the species. It is a system that is designed to deal with the following shared Hockettian design features of spoken languages: the use of the auditory-vocal channel with its specialized (neuro)anatomy, fast transmission with output-input asymmetries (with a production-comprehension rate in the proportion 1:4, see Levinson 2000a, p. 28), multiple levels of structure (phonological, morphosyntactic, semantic) bridging sound to meaning, linearity combined with non-linear structure (constituency and dependency), and massive reliance on inference. The learning system has to be able to cope with an amazing diversity of linguistic structures, as detailed in this article. Despite this, the hemispherical lateralization and neurocognitive pathways are largely shared across speakers of even the most different languages, to judge from comparative studies of European spoken and signed languages (Emmorey 2002). Yet there is increasing evidence that few areas of the brain are specialized just for language processing (see for example recent work on 'Broca's area' in Fink et al. (2005), and Hagoort (2005)).

How are we to reconcile diverse linguistic systems as the product of one cognitive system? Once the full diversity is taken into account, the UG approach becomes quite implausible – we would need to stuff the child's mind with principles appropriate to thousands of languages working on distinct structural principles. That leaves just two possible models of the cognitive system. Either the innate cognitive system has a narrow core, which is then augmented by general cognition and general learning principles to accommodate the additional structures of a specific language (as in e.g. Elman, Bates et al. 1996), or it is actually a 'machine tool', prebuilt to specialize, and construct a machine appropriate to indefinitely variable local conditions – much the picture assumed in crosslinguistic psycholinguistics of sound systems (Kuhl 1991)). Either way, when we look at adult

language processing, we find a hybrid: a biological system tuned to a specific linguistic system, itself a cultural historical product.

The fact that language is a bio-cultural hybrid is its most important property, and a key to understanding our own place in nature. For human success in colonizing virtually every ecological niche on the planet is due to adaptation through culture and technology, made possible by brains gradually evolved specifically to do that. The rapidly expanding theory of co-evolution explores the twin-track descent mechanisms of culture and biology, and the feedback loops between them (Durham 1991; Boyd & Richerson 2005; Laland et al. 2000; Odling-Smee et al. 2003). Language is one of the best exemplars of such co-evolution, with evolved biological underpinnings for culturally variable practices, where the biology constrains and canalizes but does not dictate linguistic structures.

We may summarize this emerging general picture in the following seven theses, each linked to a specifically implicated research initiative. Some of these initiatives are already under way across a range of subliteratures (linguistic typology, cognitivist and functionalist treatments, optimality theory), others not. But in either case the initiatives need to be linked across schools into an integrated general theory with hypothesis-testing procedures accepted by the whole field.

1. The diversity of language is, from a biological point of view, its most remarkable property – there is no other animal whose communication system varies both in form and content. It presupposes an extraordinary plasticity and powerful learning abilities able to cope with variation at every level of the language system. This has to be the central explicandum for a theory of human communication. [It seems inevitable that part of the explicans will be that language has coevolved with culture, which itself evolved to give rapid adaptation to fast changing environments and migration across niches - see thesis 5]. Research initiative: a principled and exhaustive global mapping of the world's linguistic diversity.

2. Linguistic diversity is structured very largely in phylogenetic (cultural-historical) and geographical patterns. Understanding these patterns basically involves the methods of population biology and cladistics, together with the principles that generate change and diversity. To the extent that there are striking similarities across languages, they

have their origin in two sources: historical common origin or mutual influence on the one hand, and on the other from selective pressures on what systems can evolve. The relevant selectors are the brain and speech apparatus, functional and cognitive constraints on communication systems, including conceptual constraints on the semantics, and internal organisational properties of viable semiotic systems. Research initiatives: First , a global assessment of structural variability comparable to that geneticists have produced for human populations, assembled in accessible syntheses like the World Atlas of Linguistics Structures (WALS) (www.eva.mpg.de/lingua/files/maps.html) and the structural phylogenetics database (Reesink, Dunn and Singer, under review). Second, we need a full and integrated account of how selectors generate structures.

3. Language diversity is characterized not by sharp boundaries between possible and impossible languages, between sharply parameterized variables, or by selection from a finite set of types. Instead it is characterized by clusters around alternative architectural solutions, by prototypes (like 'subject') with unexpected outliers, and by family-resemblance relations between structures ('words', 'noun phrases') and inventories ('adjectives'). Hypothesis: there are cross-linguistically robust system-preferences and functions, with recurrent solutions (e.g. subject) satisfying several highly ranked preferences, and outliers either satisfying only one preference, or having low-probability evolutionary steps leading to their states.

4. This kind of statistical distribution of typological variation suggests an evolutionary model with attractors (e.g. the CV syllable, a colour term 'red', a word for 'arm'), 'canals' and numerous local peaks or troughs in an adaptive landscape. Some of the attractors are cognitive, some functional (communicational), some cultural-historical in nature. Some of the canalization is due to systems-biases, as when one sound change sparks off a chain of further changes to maintain signalling discreteness. Research initiative: Each preference of this kind calls for its own focussed research in terms of which selectors are at work, along with a modelling of system interactions – e.g. computational simulations of the importance of these distinct factors along the lines reported by Steels & Belpaeme (2005).

5. The dual role of biological and cultural-historical attractors underlines the need for a co-evolutionary model of human language, where there is interaction between entities of completely different orders – biological constraints and cultural-historical traditions. A coevolutionary model explains how complex socially-shared structures emerge from the crucible of biological and cognitive constraints, functional constraints and historically inherited material. Such a model unburdens the neonate mind, reapportioning a great deal of the patterning to culture, which itself has evolved to be learnt. Inititiative: Coevolutionary models need to work for two distinct phases: one for the intensely coevolutionary period leading to modern humans, where innovations in hardware (human physiology) and software (language as cultural institution) egged each other on, and a second phase where the full variety of modern languages mutate regularly between radically different variants against a relatively constant biophysical backdrop, though population genetics may nevertheless predispose to specific linguistic variants (Dediu & Ladd 2007).

6. The biological underpinnings for language are so recently evolved, that they cannot be remotely compared to, for example, echo-location in bats (pace Jackendoff 2002, p. 79). Echolocation is an ancient adaptation shared by 17 families (the Microchiroptera) with nearly a thousand species and over 50 million years of evolution (Teeling et al. 2005), while language is an ability very recently acquired along with spiralling culture in perhaps the last 2-300,000 years by a single species.[8] Language therefore must exploit pre-existing brain machinery, which continues to do other things to this day. Language processing relies crucially on plasticity, as evidenced by the modality switch in sign languages. The major biological adaptation may prove to be the obvious anatomical one, the vocal tract itself. The null hypothesis here is that all needed brain mechanisms, outside the vocal-tract adaptation for speech, were co-opted from preexisting adaptations not specific to language (though perhaps specific to communication and sociality in a more general sense).

7. The two central challenges that language diversity poses are, firstly, to show how the full range of attested language systems can evolve and diversify as sociocultural products constrained by cognitive constraints on learning, and secondly, to show how the child's mind can learn and the adult's mind can use, with approximately equal ease, any one of this vast range of alternative systems. The first of these challenges

returns language histories to centre stage in the research program: 'why state X?' is recast as 'how does state X arise?' The second calls for a diversified and strategic harnessing of linguistic diversity as the independent variable in studying language acquisition and language processing (Box 3): can different systems be acquired by the same learning strategies, are learning rates really equivalent, and are some types of structure in fact easier to use ?

This picture may seem to contrast starkly with the assumption that was the starting point for classic cognitive science, namely the presumption of an invariant mental machinery, both in terms of its psychological architecture and neurocognitive underpinnings, underlying the common cognitive capacities of the species as a whole.

Two points need to be made here. First, there is no logical incompatibility with the classic assumption, it is simply a matter of the level at which relative cognitive uniformity is to be sought. On this new view, cognition is less like the proverbial toolbox of ready-made tools than a machine tool, capable of manufacturing special tools for special jobs. The wider the variety of tools that can be made, the more powerful the underlying mechanisms have to be. Culture provides the impetus for new tools of many different kinds – whether calculating, playing the piano, reading right to left or speaking Arabic.

Second, the classic picture is anyway in flux, under pressure from increasing evidence of individual differences. Old ideas about expertise effects are now complemented with startling evidence for plasticity in the brain – behavioural adaptation is reflected directly in the underlying wetware (as when taxi drivers' spatial expertise is correlated with growth in the hippocampal area; Maguire et al. 2000). Conversely, studies of individual variance show that uniform behaviour in language and elsewhere can be generated using distinct underlying neural machinery, as shown for example in the differing degrees of lateralization of language in individuals (see e.g. Baynes & Gazzaniga 2005; Knecht et al. 2000).

Thus the cognitive sciences are faced with a double challenge: culturally variable behaviour running on what are, at a 'zoomed out' level of granularity, closely related biological machines, and intra-cultural uniformity of behaviour running on what are, from a zoomed-in perspective, individually variable, distinct machines. But that is the human cognitive specialty that makes language and culture possible – to produce diversity out of biological

similarity, and uniformity out of biological diversity. Embedding cognitive science into what is, in a broad sense including cultural and behavioural variation, a population biology perspective, is going to be the key to understanding these central puzzles.

References

Ameka, F. & Levinson, S. C., eds. (2007) The typology and semantics of locative predicates. Special issue of *Linguistics* 45(5/6). Mouton de Gruyter.

Anderson, S. & Keenan, E. L. (1985) Deixis. In: *Language typology and syntactic description, Vol. III: Grammatical categories and the lexicon*, ed. T. Shopen, pp. 259-308. Cambridge University Press.

Aoki, K. & Feldman, M. W. (1989) K. Aoki and M.W. Feldman, Pleiotropy and preadaptation in the evolution of human language capacity, *Theoretical Population Biology* 35:181–194.

Aoki, K. & Feldman, M. W. (1994) Cultural transmission of a sign language when deafness is caused by recessive alleles at two independent loci. *Theoretical Population Biology* 45:253–61.

Arbib, M. A. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2):105-124.

Aronoff M., Meir, I., Padden, C., & Sandler, W. (2008) The roots of linguistic organization in a new language. In: *Holophrasis, Compositionality And Protolanguage, Special Issue of Interaction Studies*, ed. D. Bickerton & M. Arbib, pp. 133-149. Benjamins.

Atkinson, Q. D. & Gray, R. D. (2005) Are accurate dates an intractable problem for historical linguistics? In: *Mapping Our Ancestry: Phylogenetic Methods in Anthropology and Prehistory*, ed. C. Lipo, M. O'Brien, S. Shennan & M. Collard, pp. 193-219. Aldine.

Atkinson, Q. D., Meade, A., Venditti, C., Greenhill, S. & Pagel, M. (2008) Languages evolve in punctuational bursts. *Science*: 319(5863):588.

Austin, P. (1995) Double case marking in Kanyara and Mantharta languages, Western Australia. In: *Double Case: Agreement by Suffixaufnahme*, ed. F. Plank, pp. 363-79. Oxford University Press.

Austin, P. & Bresnan, J. (1996) Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory* 14:215-268.

Baker, M. C. (1988) *Incorporation: A theory of grammatical function changing.* University of Chicago Press.

Baker, M. C. (1993) Noun incorporation and the nature of linguistic representation. In: *The role of theory in language description*, ed. W. Foley, pp. 13-44. Mouton de Gruyter.

Baker, M. C. (1996) *The Polysynthesis Parameter*. Oxford University Press.

Baker, M. C. (2001) *Atoms of Language: The Mind's Hidden Rules of Grammar*. Basic Books.

Baker, M. C. (2003) Linguistic Differences and Language Design. *Trends in Cognitive Science* 7:349-353.

Barwise, J. & Cooper, R. (1981) Generalized Quantifiers and Natural Language. *Linguistics and Philosophy* 4:159-219.

Bates, E., Devescovi, A. & Wulfeck, B. (2001) Psycholinguistics: A Cross-Language Perspective. *Annual Review of Psychology* 52:369-96.

Baynes, K. & Gazzaniga, M. (2005) Lateralization of language: Toward a biologically based model of language. *The Linguistic Review* 22:303-26.

Bickerton, D. (1981) *Roots of Language.* Karoma.

Blevins, J. (1995) The syllable in phonological theory. In: *Handbook of Phonological Theory*, ed. J. Goldsmith, pp. 206-244. Blackwell.

Bohnemeyer, J. & Brown P. (2007) Standing divided: Dispositional verbs and locative predications in two Mayan languages. *Linguistics* 45(5/6):1105-1151.

Boroditsky, L. (2001) Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology* 43(1):1-22.

Boroditsky, L., Schmidt, L. & Phillips, W. (2003) Sex, syntax, and semantics. In: *Language in Mind: Advances in the study of Language and Cognition*, ed. D. Gentner & S. Goldin-Meadow, pp. 61-80. Cambridge University Press.

Bowerman, M. & Levinson, S. (eds.) (2001) *Language acquisition and conceptual development*. Cambridge University Press.

Boyd, R. & Richerson, P. J. (1985) *Culture and The Evolutionary Process.* University of Chicago Press.

Boyd, R. & Richerson, P. J. (2005) Solving the Puzzle of Human Cooperation in Evolution and Culture. In Levinson & Jaisson (eds):105–132.

Breen, G. & Pensalfini, R. (1999) Arrernte: A Language with no syllable onsets. *Linguistic Inquiry* 30:1–25.

Bresnan, J. (2001) *Lexical functional syntax*. Blackwell.

Brown, P. (1994) The INs and ONs of Tzeltal locative expressions: the semantics of static descriptions of location. In: Special Issue of *Linguistics* 32(4/5): *Space in Mayan languages*, ed. J.B. Haviland & S.C. Levinson, pp. 743–90.

Bybee, J. (2000) Lexicalization of sound change and alternating environments. In: *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, ed. M. D. Broe & J. B. Pierrehumbert, pp. 250–268. Cambridge University Press.

Casasanto, D. & Boroditsky, L. (2007) Time in the mind: Using space to think about time. *Cognition* 106(2):579–593.

Chomsky, N. (1955) Logical syntax and semantics: Their linguistic relevance. *Language* 31(1):36–45.

Chomsky, N. (1957) *Syntactic structures*. De Gruyter.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.* The MIT Press.

Chomsky, N. (1980) On cognitive structures and their development: A reply to Piaget. In: *Language and learning: The debate between Jean Piaget and Noam Chomsky*, ed. M. Piattelli-Palmarini, pp. 35–52. Harvard University Press.

Chomsky, N. (1981) *Lectures on Government and Binding.* Foris.

Chung, S. (1989) On the notion 'null anaphor' in Chamorro. In: *The null subject parameter*, ed. O. Jaeggli & K. Safir, pp. 143–184. Kluwer.

Christiansen, M. H. & Kirby, S. (2003) *Language evolution*. Oxford University Press.

Christiansen, M. H. & Chater, N. (2008) Language as shaped by the brain. *Behavioral and Brain Sciences* 31:489–558.

Christianson, K. & Ferreira, F. (2005) Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition* 98:105–135.

Clements, N. & Keyser, S. J. (1983) *CV Phonology: A Generative Theory of the Syllable.* MIT Press.

Colarusso, J. (1982) Western Circassian Vocalism. *Folia Slavica* 5:89–114.

Collard, I. F. & Foley, R. A. (2002) Latitudinal patterns and environmental determinants of recent cultural diversity: Do humans follow biogeographical rules? *Evolutionary Ecology Research* 4:371–83.

Comrie, B. (1976) *Aspect.* Cambridge University Press.

Comrie, B. (1985) *Tense.* Cambridge University Press.

Comrie, B. (1989) Language universals and linguistic typology. Second Edition. Blackwell.

Croft, W. (2001) *Radical construction grammar. Syntactic theory in typological perspective.* Oxford University Press.

Croft, W. (2003) *Typology and Universals.* (2nd Edition). Cambridge University Press.

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1983) A language-specific comprehension strategy. *Nature* 304:159–160.

Cutler, A., Mehler, J., Norris, D. & Segui, J. (1989) Limits on bilingualism**.** *Nature* 340:229–230.

Cysouw, M. 2001. *The paradigmatic structure of person marking*. Ph.D. thesis, Radboud University Nijmegen.

Davidoff, J., Davies, I. & Roberson, D. (1999) Color categories of a strone-age tribe. *Nature* 398:203–204.

Dediu, D. & Ladd, D. R. (2007) Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the USA* 104:10944–10949.

Dench, A. & Evans, N. (1988) Multiple case–marking in Australian languages. *Australian Journal of Linguistics* 8(1):1–47.

Dixon, R. M. W. (1972) *The Dyirbal language of North Queensland*. Cambridge University Press.

Dixon, R. M. W. (1977) The syntactic development of Australian languages. In: *Mechanisms of syntactic change*, ed. C. Li, pp. 365–418. University of Texas Press.

Dryer, M. S. (1998) Why statistical universals are better than absolute Universals. *Chicago Linguistic Society: The Panels* 33:123–145.

Dubois, J. W. (1987) The discourse basis of ergativity. *Language* 63(4):805–855.

Dunn, M., Terrill, A., Reesink, G., Foley, R. & Levinson, S. C. (2005) Structural Phylogenetics and the Reconstruction of Ancient Language History. *Science* 309:2072–2075.

Dunn, M., Levinson, S. C., Lindström, E., Reesink, G. & Terrill, A. (2008) Structural Phylogeny in historical linguistics: methodological explorations applied in Island Melanesia. *Language* 84(4):710–759.

Durie, M. (1985) *A grammar of Acehnese, on the basis of a dialect of North Aceh*. Foris.

Durham, W. (1991) *Coevolution: Genes, culture and human diversity*. Stanford University Press.

Elman, J. L., Bates, E., Johnson, M. H. & Karmiloff–Smith, A. (1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press.

Emmorey, K. (2002) *Language, cognition, and the brain: Insights from sign language research*. Erlbaum.

Enfield, N. J. (2004) Adjectives in Lao. In: *Adjective Classes: A Cross-Linguistic Typology*, ed. R. M. W. Dixon, & A. Aikhenvald, pp. 323-347. Oxford University Press.

Enfield, N. & Levinson, S. C. (2006) *Roots of human sociality: Cognition, culture and interaction.* Berg.

England, N. (2001) *Introducción a la Gramática de los idiomas Mayas.* Guatemala. Cholsomaj.

England, N. (2004) Adjectives in Mam. In: *Adjective classes: A cross-linguistic typology*, ed. R. M. W. Dixon, pp. 125-146. Oxford University Press.

Evans, N. (1995a) Multiple case in Kayardild: Anti-iconic suffix ordering and the diachronic filter. In: *Double Case: Agreement by Suffixaufnahme*, ed. F. Plank, pp. 396-430. Oxford University Press.

Evans, N. (1995b) *A Grammar of Kayardild.* Mouton de Gruyter.

Evans, N. (2002) The true status of grammatical object affixes: Evidence from Bininj Gun-wok. In: *Problems of polysynthesis*, ed. N. Evans & H.-J. Sasse, pp. 15–50. Akademie Verlag.

Evans, N. (2003a) *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune*. Pacific Linguistics.

Evans, N. (2003b) Context, culture and structuration in the languages of Australia. *Annual Review of Anthropology* 32:13–40.

Evans, N. & Sasse, H. –J., eds. (2002) *Problems of polysynthesis*. Akademie Verlag.

Evans, N. & Osada, T. (2005) Mundari: The myth of a language without word classes. *Linguistic Typology* 9.3:351-390.

Everett, D. (2005) Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46:621–646.

Fink, G. R., Manjaly, Z. M., Stephen, K. E., Gurd, J. M., Zilles, K., Amunts, K. & Marshall, J. C. (2005) A role for Broca's area beyond language processing: Evidence from neuropsychology and fMRI. In: *Broca's Area*, ed. K. Amunts & Y. Grodzinsky, pp. 254–271. Oxford University Press.

Fisher, S. & Marcus, G. (2006) The eloquent ape: genes, brains and the evolution of language. *Nature Reviews Genetics* 7:9–20.

Fitch, W. T. & Hauser, M. D. (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303(5656):377–380.

Fitch, W. T., Hauser, M. D. & Chomsky, N. (2005) The Evolution of the language faculty: Clarifications and implications. *Cognition* 97:179–210.

Fodor, J. A. (1975) *The language of thought*. Thomas Y. Crowell.

Fodor, J. A. (1983) *The modularity of mind*. MIT Press.

Friederici, A. (2004) Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Science* 8:245–247.

Gazdar, G. & Pullum, G. K. (1982) *Generalized phrase structure grammar: A theoretical synopsis.* Indiana Linguistics Club.

Gazdar, G., Klein, E., Pullum, G. & Sag, I. (1985) *Generalized phrase structure grammar.* Blackwell.

Gentner, D. & Boroditsky, L. (2001) Individuation, relativity, and early word learning. In: *Language acquisition and conceptual development*, ed. M. Bowerman & S. C. Levinson, pp. 215–256. Cambridge University Press.

Gil, D. (2001) Escaping eurocentrism. In: *Linguistic Fieldwork*, ed. P. Newman & M. Ratliff, pp. 102–132. Cambridge University Press.

Gleitman, L. (1990) The structural sources of verb meanings. *Language Acquisition* 1:3–55.

Goldin-Meadow, S. (2003) *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. Psychology Press.

Gordon, P. (2004) Numerical cognition without words: Evidence from Amazonia. *Science* 306(5695):496–499.

Guo, J., Lieven, E., Budwig, N., Ervin-Tripp, S., Nakamura, K. & Ozcaliskan, S., eds. (2008) *Cross-linguistic approaches to the psychology of language.* Psychology Press.

Gray, R. D. & Atkinson, Q. D. (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426:435–439.

Greenberg, J. H. (1963a) Some universals of grammar with particular reference to the order of meaningful elements. In: *Universals of Language*, ed. J. H.Greenberg, pp. 72–113. MIT Press.

Greenberg, J. H., ed. (1963b) *Universals of Language.* MIT Press.

Greenberg, J. H. (1966) *Language Universals*. Mouton de Gruyter.

Greenberg, J. H. (1978) Generalizations about numeral systems. In: Universals of human language, ed. J. H. Greenberg, pp. 249–296. Stanford University Press.

Greenberg, J. H. (1986) On being a linguistic anthropologist. *Annual Review of Anthropology* 15:1–24.

Greenberg, J. H., Osgood, C. E. & Jenkins, J. J. (1963) Memorandum concerning language universals. In: *Universals of Language,* ed. J. H. Greenberg, pp. xv–xxvii. MIT Press.

Greenfield, P. M. (1991) Language, tools, and brain: The ontogeny and phylogeny of

hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14(4): 531–551.

Hagoort, P. (2005) On Broca, Brain and Binding: A New Framework. *Trends in Cognitive Science* 9:416–23.

Hale, K. (1983) Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1:5–47.

Hale, K., Laughren, M. & Simpson, J. (1995) Warlpiri. In: *Syntax. Ein internationales Handbuch zeitgenössischer Forschung. An International Handbook of Contemporary Research*, ed. J. Jacobs, A. von Stechow, W. Sternefeld & T. Vennemann, pp. 1430–1451. Walter de Gruyter.

Halle, M. (1970) Is Kabardian a Vowel-Less Language? *Foundations of Language* 6:95–103.

Hankamer, J. (1989) Morphological parsing and the lexicon. In: *Lexical Representation and Process*, ed. W. Marslen-Wilson, pp. 392–408. MIT Press.

Harris, A. (2008) On the explanation of typologically unusual structures. In: *Language Universals and Language Change*, ed. J. Good, pp. 54–76. Oxford University Press.

Haspelmath, M. (1993) *A grammar of Lezgian.* (Mouton Grammar Library, 9). Mouton de Gruyter.

Haspelmath, M. (1999) Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18.2:180-205.

Haspelmath, M. (2007) Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11.1:119–132

Hauser, M. D. (1997) *The evolution of communication.* MIT Press/BradfordBooks.

Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) The faculty of language: What is it, who has it, and how does it evolve? *Science* 298:1569–1579.

Haviland, J. B. (1979) Guugu Yimidhirr. In: *Handbook of Australian Languages,* vol. I, ed. B. Blake & R. M.W. Dixon, pp. 27–182. ANU Press.

Hengeveld, K. (1992) Parts of speech. In: *Layered Structure and Reference in a Functional Perspective*, ed. M. Fortescue, P. Harder & L. Kristofferson, pp. 29–56. Benjamins.

Himmelmann, N. (1997) *Deiktikon, Artikel, Nominalphrase, Zur Emergenz Syntaktischer Struktur.* Niemeyer.

Hockett, C. F. (1963) The problem of universals in language. In: *Universals of Language*, ed. J. H. Greenberg, pp. 1-29. MIT Press.

Hornstein, N., Nunes, J. & Grohmann, K. (2005) *Understanding minimalism.* Cambridge University Press.

Hudson, R. (1993) Recent developments in dependency theory. In: *Syntax: An International Handbook of Contemporary Research*, ed. J. Jacobs, A. von Stechow, W. Sternefelt & T. Vennemann, pp. 329-338. Walter de Gruyter.

Hunley, K., Dunn, M., Lindström, E., Reesink, G., Terrill, A., Norton, H., Scheinfeldt, L., Friedlaender, F., Merriwether, D.A., Koki, G. & Friedlaender, J. (2007) Inferring prehistory from genetic, linguistic, and geographic variation. In: *Genes, Language, and Culture History in the Southwest Pacific*, ed. J.S. Friedlaender, pp. 141–154. Oxford University Press.

Jackendoff, R. (2002) *Foundations of Language.* Oxford University Press.

Jackendoff, R. (2003a) Reintegrating generative grammar (Précis of *Foundations of Language). Behavioral and Brain Sciences* 26:651–665.

Jackendoff, R. (2003b) Toward better mutual understanding (response to peer commentaries on 2003a. *Behavorial and Brain sciences* 26:695–702.

Jakobson, R. (1962) *Selected Writings I: Phonological Studies.* Mouton de Gruyter.

Jakobson, R. & Halle, M. (1956) *Fundamentals of Language.* Mouton de Gruyter.

Jelinek, E. (1995) Quantification in Straits Salish. In: *Quantification in Natural Languages*, ed. E. Bach, E. Jelinek, A. Kratzer & B. Partee, pp. 487-540. Kluwer.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65–79.

Klima, E. S. & Bellugi, U. (1979) *The signs of language*. Harvard University Press.

Knecht S., Deppe M., Dräger B., Bobe L., Lohmann H., Ringelstein E-B & Henningsen, H. (2000) Language lateralization in healthy right-handers. *Brain* 123(1):74–81.

Koster, J. & May, R. (1982) On the constituency of infinitives. *Language* 58:116–143.

Kuhl, P. (1991) Perception, cognition and the ontogenetic and phylogenetic emergence of human speech. In: *Plasticity of development*, ed. S. E. Brauth, W. S. Hall & R. J. Dooling, pp. 73–106. MIT Press.

Kuhl, P. K. (2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5:831–843.

Kuipers, A. H. (1960) *Phoneme and morpheme in Kabardian*. Mouton de Gruyter.

Labov, W. (1980) The social origins of sound change. In: *Locating language in time and space*, ed. W. Labov, pp. 251–66. Academic Press.

Ladefoged, P. & Maddieson, I. (1996) *The Sounds of the World's Languages.* Blackwell.

Laland, K. N., Odling-Smee, J. & Feldman, M. W. (2000) Niche Construction, Biological Evolution and Cultural Change. *Behavioral and Brain Science* 23:131–175.

Landau, B. & Jackendoff, R. (1993) 'What' and 'Where' in Spatial Language and Spatial Cognition, *Behavioral and Brain Sciences* 16:217–265.

Lenneberg, E. H. (1967) *Biological Foundations of Language.* John Wiley.

Levelt, W. J. M. (2008) *Formal Grammars In Linguistics And Psycholinguistics*. Benjamins.

Levinson, S. C. (2000a) *Presumptive meanings: The theory of generalized conversational implicature.* MIT Press.

Levinson, S. C. (2003) *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press.

Levinson, S. C. & Meira, S. (2003) 'Natural concepts' in the spatial topological domain - adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79(3):485–516.

Levinson, S. C. & Jaisson, P., eds. (2006) *Evolution and Culture. A Fyssen Foundation Symposium*. The MIT Press.

Levinson, S. C. & Wilkins, D., eds. (2006) *Grammars of space*. Cambridge University Press.

Levinson, S. C. (in press) Syntactic ergativity in Yélî Dnye, the Papuan language of Rossel Island, and its implications for typology. Linguistic Typology.

Li, P. W. & Gleitman, L. (2002) Turning the tables: language and spatial reasoning. *Cognition* 83(3):265–294.

Li, P., Hai Tan, L., Bates, E. & Tzeng, O. (2006) *The Handbook of East Asian Psycholinguistics*, Vol. 1: *Chinese*. Cambridge University Press.

Lucy, J. (1992) *Grammatical categories and thought: A case study of the linguistic relativity hypothesis*. Cambridge University Press.

Mace, R. & Pagel, M. (1995) A latitudinal gradient in the density of human languages in North America. *Proceedings of the Royal Society of London (B)* 261:117–121.

MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S. et al. (2002) Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain* 125:1583–1593.

MacWhinney, B. & Bates, E., eds. (1989) *The cross-linguistic study of sentence processing*. Cambridge University Press.

Maddieson, I. (1983) The analysis of complex phonetic elements in Bura and the syllable. *Studies in African Linguistics* 14:285–310.

Maddieson, I. (1984) *Patterns of Sounds.* Cambridge University Press.

Maddieson, I. & Levinson, S. C. (in prep.) The phonetics of Yélî Dnye, the language of Rossel Island.

Maguire, E., Gadian, D., Johnsrude, I, Good, D., Ashburner, J., Frackowiak, R. & Frith, C. (2000) Navigation-related structural changes in the hippocampi of taxi drivers. *PNAS* 97(8):4398–4403.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M. & Levinson, S. C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8(3):108–114.

Margetts, A. (2007) Learning verbs without boots and straps? The problem of 'give' in Saliba. In *Cross-linguistic perspectives on argument structure: Implications for learnability*, ed. M. Bowerman & P. Brown. pp 111–141. Erlbaum.

Marsaja, I. G. (2008) *Desa kolok - A deaf village and its sign language in Bali, Indonesia*. Nijmegen: Ishara Press.

Master, A. (1946) The zero negative in Dravidian. *Transactions of the Philological Society* 1946:137–155.

Matthews, P. H. (1981) *Syntax*. Cambridge University Press.

Matthews, P. H. (2007) *Syntactic relations: A critical survey.* Cambridge University Press.

MacLarnon, A. & Hewitt, G. (2004) Increased Breathing Control: Another factor in the Evolution of Human Language. *Evolutionary Anthropology* 13:181–197.

McMahon, A. & McMahon, R. (2006) *Language classification by numbers.* Oxford University Press.

Meir, I., Sandler, W., Padden, C. & Aronoff, M. (in press) Emerging sign languages. In: *Oxford Handbook of Deaf Studies, Language, and Education*, Volume 2, ed. M. Marschark & P. Spencer. Oxford University Press.

Melčuk, I. (1988) *Dependency syntax: Theory and Practice.* The SUNY Press.

Mielke, J. (2007) *The emergence of distinctive features*. Oxford University Press.

Mithun, M. (1984) How to avoid subordination. *Berkeley Linguistic Society* 10:493–509.

Mithun, M. (1999) *The Languages of North America.* Cambridge University Press.

Nakayama, R., Mazuka, R. & Shirai, Y. (2006) *Handbook of East Asian Psycholinguistics* Vol 2: *Japanese.* Cambridge University Press.

Nettle, D. (1999) *Linguistic Diversity*. Oxford University Press.

Newmeyer, F. J. (1986) *Linguistic theory in America* (2nd ed.). Academic Press.

Newmeyer, F. J. (2004) Typological evidence and Universal Grammar. *Studies in Language* 28:527–548.

Newmeyer, F. J. (2005) *Possible and probable languages: A generative perspective on linguistic typology*. Oxford University Press.

Nichols, J. (1992) *Language diversity in space and time*. University of Chicago Press.

Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., Nishimura, T. & Kubo, T. (1999) Sign language 'heard' in the auditory cortex. *Nature* 397 (6715): 116.

Nordlinger, R. (1998) *Constructive case. Evidence from Australian languages*. CSLI.

Nordlinger, R. & Sadler, L. (2004) Nominal tense in crosslinguistic perspective. *Language* 80:776–806.

Norman, J. 1988. *Chinese*. Cambridge University Press.

O'Donnell, T., Hauser, M. & Fitch, W. T. (2005) Using mathematical models of language experimentally. *Trends in Cognitive Science* 9:284–289.

Odling-Smee, F. J., Laland, K. N. & Feldman, M.W. (2003) *Niche construction: The neglected process in evolution.* Princeton University Press.

Osada, T. (1992) *A reference grammar of Mundari*. Institute for the Languages and Cultures of Asia and Africa.

Padden, C. & Perlmutter, D. (1987) American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory* 5:335–375.

Pagel, M. (2000) The history, rate and pattern of world linguistic evolution. In: *The evolutionary emergence of language*, ed. C. Knight, M. Studdert-Kennedy & J. Hurford, pp. 391-416. Cambridge University Press.

Pagel, M., Atkinson, Q. D. & Meade, A. (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449:717–720.

Parker, A. (2006) Evolving the narrow language faculty: was recursion the pivotal step? In: *The Evolution of Language: Proceedings of the 6th International Conference on the Evolution of Language*, ed. A. Cangelosi, A. D. M. Smith & K. Smith, pp.239–246. World Scientific Press.

Partee, B. H. (1995) Quantificational structures and compositionality. In: *Quantification in Natural Languages*, ed. E. Bach, E. Jelinek, A. Kratzer & B. H. Partee, pp. 541–602. Kluwer.

Partee, B., ter Meulen, A. & Wall, R. (1990) *Mathematical methods in linguistics*. Kluwer.

Pawley, A. (1993) A language which defies description by ordinary means. In: *The role of theory in language description*, ed. W. Foley, pp. 87–129. Mouton de Gruyter.

Pederson, E. (1993) Zero negation in South Dravidian. In: *CLS 27: The parasession on negation*, ed. L. M. Dobrin, L. Nicholas & R. M. Rodriguez, pp. 233–245. Chicago Linguistic Society.

Perniss, P. M., Pfau, R. & Steinbach, M., eds. (2008) Visible Variation: Comparative Studies on Sign Language Structure. *Trends in Linguistics* 188. Mouton de Gruyter.

Perniss, P. M. & Zeshan, U., eds. (2008) *Possessive and existential constructions in sign languages*. Sign Language Typology Series No. 2. Nijmegen: Ishara Press.

Pierrehumbert, J. (2000) What people know about the sounds of language. *Linguistic Sciences* 29:111–120.

Pierrehumbert, J., Beckman, M. E. & Ladd, D. R. (2000) Conceptual foundations of phonology as a laboratory science. In: *Phonological Knowledge: Its Nature and Status*, ed. N. Burton-Roberts, P. Carr & G. Docherty, pp. 273–303. Cambridge University Press.

Pinker, S. (1994) *The Language Instinct.* Morrow.

Pinker, S. & Bloom, P. (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–726.

Pinker, S. & Jackendoff, R. (2005) The Faculty of Language: What's Special about it? *Cognition* 95:201–236.

Port, R. & Leary, A. (2005) Against formal phonology. *Language* 81:927–964.

Postal, P. (1970) The method of universal grammar. In: *On Method in Linguistics*, ed. P. Garvin, pp. 113–131. Mouton de Gruyter.

Pye, C., Pfeiler, B., de León, L., Brown, P. & Mateo, P. (2007) Roots or edges? Explaining variation in children's early verb forms across five Mayan languages. In: *Learning indigenous languages: Child language acquisition in Mesoamerica and among the Basques*, ed. B. Blaha Pfeiler, pp.15-46. Mouton de Gruyter.

Reesink, G., Dunn, M. & Singer, R. (under review) Explaining the linguistic diversity of Sahul using population models. PLOS.

Sandler, W. (2008) *Symbiotic symbolization by hand and mouth in sign language.* (PDF) ms. University of Haifa.

Sandler, W., Meir, I., Padden, C. & Aronoff, M. (2005) The Emergence of Grammar in a New Sign Language. *Proceedings of the National Academy of Sciences.* Vol 102, No. 7:2661-2665.

Schachter, P. (1976) The subject in Philippine languages: Topic–actor, actor–topic, or none of the above. In: *Subject and topic*, ed. C. Li, pp. 491–518. Academic Press.

Schultze-Berndt, E. (2000) *Simple and Complex Verbs in Jaminjung: A Study of Event Categorization in an Australian Language.* PhD dissertation, MPI Series in Psycholinguistics.

Schwager, W. & Zeshan, U. (2008) Word classes in sign languages – Criteria and classifications. *Studies in Language* 32(3):509-545.

Senghas, A., Kita, S. & Özyürek, A. (2004) Children Creating Core Properties of Language: Evidence from an Emerging Sign Language in Nicaragua. *Science* 305(5691):1779–1782.

Slobin, D., ed. (1997) *The crosslinguistic study of language acquisition*. Erlbaum.

Steels, L. & Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: a case study for colour. *Behavioral and Brain Sciences* 28:469–529.

Talmy, L. (2000) *Towards a cognitive semantics*. MIT Press.

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'Brien, S. J. & Murphy, W. J. (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307: 580–584.

Tomasello, M. (1995) Language is not an instinct. *Cognitive Development* 10:131–156.

Tomasello, M. (2000) *The Cultural Origins of Human Cognition.* Harvard University Press.

Van Valin, R. D. & LaPolla, R. (1997) *Syntax. Structure, meaning and function*. Cambridge University Press.

Vigliocco, G., Vinson, D. P., Paganelli F. & Dworzynski, K. (2005) Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General* 134:501–520.

Wall, J. D., & Kim S. K. (2007) Inconsistencies in Neanderthal Genomic DNA Sequences. PLoS Genetics 3(10): e175 doi:10.1371/journal.pgen.0030175

Werker, J. F. & Tees, R. C. (2005) Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* 46(3):233–251.

Whaley, L. (1997) *Introduction to Typology.* Sage Publications.

Widmann, T. & Bakker, D. (2006) Does sampling matter: a test in replicability, concerning numerals. *Linguistic Typology* 10(1):83–95.

Zeshan, U. (2002) Indo-Pakistani sign language grammar: a typological outline. *Sign Language Studies* 3.2:157–212.

Zeshan, U., ed. (2006a) Interrogative and negative constructions in sign languages. *Sign Language Typology. Series No. 1.* Nijmegen: Ishara Press.

Zeshan, U. (2006b) Sign Languages of the World. In: *Encyclopedia of Language and Linguistics* (2nd ed), ed. K. Brown, pp. 358-365. Elsevier Publishers.

Boxes

Box 1: "Every language has X, doesn't it?"
- Proposed substantive universals (from Pinker & Bloom 1990) supposedly common to all languages:

1. "major lexical categories (noun, verb, adjective, preposition)" ($\rightarrow$ §2.2.4)
2. "major phrasal categories (noun phrase, verb phrase, etc.)" ($\rightarrow$ §5)
3. "phrase structure rules (e.g. "X-bar theory" or "immediate dominance rules")" ($\rightarrow$ §5)
4. "rules of linear order" to distinguish e.g. subject from object, or "case affixes" which "can take over these functions" ($\rightarrow$ §5)
5. "verb affixes" signalling "aspect" and "tense" (including pluperfects) ($\rightarrow$ §2.2.3)
6. "auxiliaries"
7. "anaphoric elements" including pronouns and reflexives
8. "Wh-movement"

There are clear counterexamples to each of these claims. Problems with the first three are discussed in §2.2.4 and §5; here are counterexamples to the others:
(4) Some languages (e.g. Riau Malay) exhibit neither fixed word-order nor case-marking (Gil 2001)
(5) Many languages (e.g. Chinese, Malay) do not mark tense (Norman 1988, p.163; Comrie 1985, p. 50-5) and many (e.g. spoken German) lack aspect (Comrie 1976, p. 8)
(6) Many languages lack auxiliaries (e.g. Kayardild, Bininj Gun-wok)
(7) Many languages (e.g. Fijian) lack reflexives or reciprocals of any form (Levinson 2000a, p. 334ff). Some SE Asian languages lack clear personal pronouns, using titles (of the kind 'honourable sir') instead, and many languages lack 3rd person pronouns (Cysouw 2001). Sign languages like ASL also lack pronouns, using pointing instead.
(8) Not all languages (e.g. Chinese, Japanese, Lakhota) move their wh-forms, saying in effect "You came to see who?" instead of "Who did you come to see _" (Van Valin & Lapolla 1997, p. 424-5)

Some further universalizing claims with counterevidence:
(9) verbs for 'give' always have three arguments (Gleitman 1990) – Saliba is a counter-example (Margetts 2007)
(10) no recursion of case (Pinker & Bloom 1990). Kayardild has up to 4 layers (Evans 1995a,b)
(11) no languages have nominal tense (Pinker & Bloom 1990) – Nordlinger & Sadler (2004) give numerous counterexamples, such as Guarani 'my house-FUTURE-FUTURE' 'it will be my future house'
(12) All languages have numerals (Greenberg 1978 – Konstanz #527) - see Everett (2005; Gordon 2004) for counterexample.
(13) All languages have syntactic constituents, specifically NPs, whose semantic function is to express generalized quantifiers over the domain of discourse (Barwise & Cooper 1981, Konstanz#1203) – see Partee (1995). See also collection of "rara" at http://typo.uni-konstanz.de/rara/intro/index.php

Box 2 The challenge of sign languages

Many proposed universals of language ignore the existence of sign languages – the languages of the deaf, now recognized to be full-blown languages of independent origin (Klima & Bellugi 1979). Studies of e.g. American Sign Language, British Sign Language, and Indo-Pakistani Sign Language (Zeshan 2002) show that these are unrelated, complex systems of their own. They can even be said to have 'phonologies' – patterns of hand shape, facial expression, etc., which although individually meaningless can be combined to make morphemes or words (Padden & Perlmutter 1987).

The typology of sign languages is in its infancy (see e.g. Zeshan 2006a,b; Perniss et al. 2008; Perniss & Zeshan 2008; Schwager & Zeshan 2008). The Ethnologue lists 121 sign languages, but there are certainly many not yet listed. The major sign languages show some typological similarities, but the smaller ones, only now coming under scrutiny, are typologically diverse (see e.g. Meir et al. in press).

Sign languages offer a model 'organism' for understanding the relation between biological and cultural aspects of language (Aoki & Feldman 1994). They also offer unique opportunities to study the emergence of new languages under different conditions: (a) where home-signers (Goldin-Meadow 2003) are congregated and a sign language emerges among themselves, as in Nicaragua (Senghas et al. 2004), (b) where a localized hereditary deaf population lives among hearers who also sign, as in Bali (Marsaja 2008) or in a Bedouin group in Israel (Sandler et al. 2005). These studies show that although word order constraints may show early, it may take three generations or more to evolve syntactic embedding and morphology.

When due allowance is made for the manual-visual interface, sign languages seem to be handled by the same specialized brain structures as spoken ones, with parallel aphasias, similar developmental trajectories (e.g. infants 'babble' in sign), and similar processing strategies as spoken languages (see Emmorey 2002). The neurocognition of sign does not look, e.g. like the neurocognition of gesture, but instead recruits, e.g. auditory cortex (Nishimura et al. 1999; MacSweeney et al. 2002). These results show that our biological endowment for language is by no means restricted to the input/output systems of a particular modality.

Box 3   Using linguistic diversity as a 'natural laboratory' in cognitive science
Instead of yearning for simple language universals, cognitive scientists should embrace the diversity and use it for what it is good for: supplying rich independent variables for experimental purposes. Listed here are some of the existing outstanding uses of this natural laboratory of variation in the communication system of our species (see also Bates et al. 2001; Nakayama et al. 2006; Li et al. 2006; Slobin 1997; Guo et al. 2008).

Modality – sign vs. speech
The modality transfer in sign vs. spoken language can be exploited to explore the nature of language processing when the input/output systems are switched, thus allowing glimpses into language-specific cognition beyond the vocal-auditory specializations (Emmorey (2002), MacSweeney et al. (2002)) .

Sound systems and how we process them
Kuhl (1991; 2004); Werker & Tees (2005) have exploited cross-linguistic differences in sound systems, showing that infants from 6 months old are already 'tuning' their acoustic space to the sound system of the language they are learning.
Cutler et al. (1983) have shown that parsing the sound stream so for word recognition crucially involves the rhythmic structure of the language, which can be . of at least three distinct kinds: based on syllable-timing (e.g. Spanish), stress-timing (e.g. English) and on the mora (e.g. Japanese). Even the best bilinguals tend to use only one system (Cutler et al. 1989).

Morphology
Psycholinguists have exploited structural differences between languages to explore , e.g., the cognitive effects of gender systems (Boroditsky et al. 2003) or the effects on processing (Vigliocco et al. 2005).

Syntax
The full variety of syntax types has hardly begun to be exploited by psycholinguists (but see MacWhinney & Bates 1989). For some first psycholinguistic investigations of a free word order language (Odawa) see Christianson & Ferreira (2005).

Semantics and conceptual structure
Differences in the linguistic coding of e.g. space can be shown to correlate with differences in the non-linguistic conceptual coding (Levinson 2003; Majid et al. 2004; Levinson & Wilkins 2006), suggesting that linguistic distinctions affect how we think (see also Lucy 1992; 2001; Boroditsky 2001).

 Lexicon
Lexcical gaps can impact perception (e.g. of color: Kay & Kempton 1984; Davidoff et al. 1999) and cognition (e.g. of number: Gordon 2004). Meaning diversity has obvious implications for language acquisition: do we name pre-existing concepts or construct the concepts during learning (Bowerman & Levinson 2001)?

 Acquisition of morphosyntax
The diversity of the child's target has been the focus of a great deal of acquisition research (see e.g. Slobin 1997). Controlled comparison across languages can be very revealing of children's learning strategies (Pye et al. 2007).

Endnotes

[1] We thank Jos van Berkum, Morten Christiansen, Peter Hagoort, Martin Haspelmath, Fritz Newmeyer, Wolfgang Klein, Asifa Majid, Karl-Magnus Petersson for helpful comments, Barbara Finlay, Adele Goldberg, Shimon Edelman and three further anonymous BBS referees for their suggestions on improvement, and Edith Sjoerdsma for her work on the MS. Evans acknowledges the Australian Research Council (Grant: Social Cognition and Language, DP878126) for relevant grant support.

[2] See also Postal (1970, p. 130), responding to a challenge: 'I would be willing to postulate universals on the basis of an even weaker study, namely of one language'.

[3] There will certainly prove to be a genetic basis for the input/output mechanisms for language, and very probably for vocal learning – the FOXP2 story points in these directions (Fisher & Marcus 2006). Whether the higher, more abstract aspects of language reside in language-specific abilities, however, remains quite unclear (but see Box 2).

[4] Such "design features" may, it is thought, offer us at least a bedrock of exceptionless universals. Even here, though, exceptions are reported – e.g. not all sign languages appear to show duality of patterning (Aronoff et al 2008).

[5] e.g., Hornstein et al. 2005, p.7 describe constituent structure as one of the uncontentious 'big facts' about language. See Matthews 2007 for a careful demolition of this position.

[6] There is some unclarity over what is meant here by recursion: in the Chomsky hierarchy all formal languages are recursively enumerable, but only some are recursive in the sense of decidable (Levelt 2008:(I)106). As far as structures are concerned, tail (or peripheral) recursion (sometimes hard to distinguish from mere iteration) is usefully differentiated from nested recursion or centre-embedding, which requires push-down stack memory (see Pinker & Jackendoff 2005:203, Parker 2006).

[7] For examples of ongoing co-evolution in the language domain see Dediu & Ladd 2007; Aoki & Feldman 1989.

[8] Any date for the origin of a graduated capacity will be arbitary, but consider for example the breathing control necessary for extended speech (MacLarnon & Hewitt 2004): the relevant thoracic vertebral enervation dates after Homo erectus and before the split between modern humans and Neanderthals at perhaps 325,000 years ago (Wall & Kim 2007).