BEHAVIORAL AND BRAIN SCIENCES (2009) **32**, 429–492
doi:10.1017/S0140525X0999094X

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
http://rspas.anu.edu.au/people/personal/evann_ling.php

Stephen C. Levinson
Max Planck Institute for Psycholinguistics, Wundtlaan 1, NL-6525 XD Nijmegen, The Netherlands; and Radboud University, Department of Linguistics, Nijmegen, The Netherlands
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. This target 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 6,000 to 8,000 languages. After surveying the various uses of "universal," we illustrate the ways languages vary radically in sound, meaning, and syntactic organization, and then we examine in more detail the core grammatical machinery of recursion, constituency, and grammatical relations. Although 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 that is fundamentally variable at all levels. Recognizing 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

> *According to Chomsky, a visiting Martian scientist would surely conclude that aside from their mutually unintelligible vocabularies, Earthlings speak a single language.*
>
> — Steven 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 argue here, 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 and 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

© Cambridge University Press, 2009 0140-525X/09 $40.00

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 advertizing 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 it 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 universalizing 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 (sect. 2) in this article, 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 (sect. 3) the kind of descriptive generalizations (again, misleadingly called "universals") that have emerged from careful cross-linguistic comparisons, and we 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 that have recently figured large in debates about the origin of language: grammatical relations (sect. 4), constituency (sect. 5), and recursion (sect. 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 (sect. 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 (sect. 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.

NICHOLAS EVANS is Professor of Linguistics at the Australian National University. His more than 120 linguistic publications include grammars of Kayardild and Bininj Gun-wok; dictionaries of Kayardild and Dalabon; edited books on polysynthesis, linguistic prehistory, and grammar-writing; and the recent *Dying Words: Endangered Languages and What They Have To Tell Us* (Wiley Blackwell, 2009). He has carried out intensive fieldwork on a number of languages of Australia and Papua New Guinea. Current research projects focus on the encoding of psychosocial cognition in grammar, song language traditions of Arnhem Land, and languages of South Coast New Guinea. Evans is a fellow of the Australian Academy of the Humanities.

STEPHEN C. LEVINSON is co-director of the Max Planck Institute for Psycholinguistics, and Professor of Comparative Linguistics at Radboud University, Nijmegen, The Netherlands. He is the author of more than 150 publications on language and cognition, including the books *Pragmatics* (Cambridge University Press [CUP], 1983), *Politeness* (CUP, 1987), *Presumptive Meanings* (MIT, 2000), *Space in Language and Cognition* (CUP, 2003). In addition, he has co-edited the following collections: *Language Acquisition and Conceptual Development* (CUP, 2001) with M. Bowerman; *Grammars of Space* (CUP, 2006) with D. Wilkins; *Evolution and Culture* (MIT, 2006) with P. Jaisson; and *Roots of Sociality* (Berg, 2006) with N. Enfield. Levinson has done extensive fieldwork on languages in India, Australia, Mexico, and Papua New Guinea, and coordinated research on the typology of languages in New Guinea and Australia. He is a fellow of the British Academy and the Academia Europaea.

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.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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)" (→ sect. 2.2.4)
2. "Major phrasal categories (noun phrase, verb phrase, etc.)" (→ sect. 5)
3. "Phrase structure rules (e.g., "X-bar theory" or "immediate dominance rules")" (→ sect. 5)
4. "Rules of linear order" to distinguish, for example, subject from object, or "case affixes" which "can take over these functions" (→ sect. 5)
5. "Verb affixes" signaling "aspect" and "tense" (including pluperfects) (→ sect. 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 section 2.2.4 and section 5; here are counterexamples to the others:

(4) Some languages (e.g., Riau Indonesian) exhibit neither fixed word-order nor case-marking (Gil 2001).
(5) Many languages (e.g., Chinese, Malay) do not mark tense (Comrie 1985, pp. 50–55; Norman 1988, p. 163), 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. Mwotlap; François 2005, p. 119) lack dedicated reflexive or reciprocal constructions altogether, so that "they hit them dead" can mean "they killed them," "they killed themselves," or "they killed each other" (Levinson 2000, p. 334 ff.). Some Southeast Asian languages lack clear personal pronouns, using titles (of the kind "honorable sir") instead, and many languages lack third-person pronouns (Cysouw 2001). Sign languages like ASL (American Sign Language) 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, pp. 424–25).

Some further universalizing claims with counterevidence:

(9) Verbs for "give" always have three arguments (Gleitman 1990); Saliba is a counterexample (Margetts 2007).
(10) No recursion of case (Pinker & Bloom 1990). Kayardild has up to four layers (Evans 1995a; 1995c).
(11) No languages have nominal tense (Pinker & Bloom 1990) – Nordlinger and Sadler (2004) give numerous counterexamples, such as Guarani "my house-FUTURE-FUTURE" "it will be my future house."
(12) All languages have numerals (Greenberg 1978b – 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) and sect. 5.

See also collection of "rara" at: http://typo.uni-konstanz.de/rara/intro/index.php

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 12 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 later, 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 bio-cultural hybrid, a product of intensive gene:culture coevolution over perhaps the last 200,000 to 400,000 years (Boyd & Richerson 1985; Enfield & Levinson 2006; Laland et al. 2000; Levinson & Jaisson 2006).

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 (sect. 4). The parsing system for English cannot 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 is 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; see also 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 – 7,000 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 144 distinctive sounds versus 11? 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 5,000 and 8,000 distinct languages are spoken today. How come we cannot 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 languages 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 worldwide 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 a language dies every two weeks. 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, because 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, although 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 more than 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 300 to 400 groups. This may become more tractable with the application of modern cladistic techniques (Dunn et al. 2005; Gray & Atkinson 2003; 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 7,000 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 less than 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 coevolution (Aoki & Feldman 1994; Durham 1991): Without the strain of hereditary deafness, the cultural adaptation would not exist, whereas 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 2009). Whichever evolutionary scenario you favor, 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 et al. 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 "t̪p" where the t̪ is further back towards the palate (Ladefoged & Maddieson 1996, pp. 344–45; 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 favor 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 et al. 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)

And,

> 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 versus 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* /lɔ:/ or *gnaw* /nɔ:/) over VC syllables (like *awl* /ɔ:l/ or *awn* /ɔ:n/). 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:

$$CV > V > VC$$

This long-proclaimed conditional universal (Jakobson & Halle 1956; cf. Clements & Keyser 1983; Jakobson 1962) 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 and Pensalfini published a clear demonstration that Arrernte organizes its syllables around a VC(C) structure and does not permit consonantal onsets (Breen & Pensalfini 1999). 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, and aspect, as well as systematic word derivation processes. They even lack 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 *Ęskakhehǫna'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 that 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 *-hǫna-* 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 hold off discussing these in detail until sections 4 through 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, because 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 behaviors, 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 Northwest 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 – e.g., *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 color. 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 & Levinson 2007). These are widespread in Mayan languages (Bohnemeyer & Brown 2007; Brown 1994; England 2001; 2004). Examples from Tzeltal include *latz'al*, "of flat items, arranged in vertical stack"; *chepel*, "be located in bulging bag," and so on. 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 to 30 inflecting verbs, but they 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 hand-shape classifiers in sign languages that represent the involved entity through a schematized representation of its shape. And further unfamiliar word classes are continuously being unearthed that 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, etc.), many of which will never be needed. And, because 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 effectively limits 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," and 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 (Mithun 1999, 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, and so on (Mithun 1999, p. 106 ff.). 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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, because 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, whereas 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, whereas 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 and Jackendoff (1993) claimed that this difference corresponds to the nature of the so-called *what* versus *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. These researchers thus felt able to confidently predict that there would be no preposition or spatial relator encoding featural properties of objects, for example, none meaning "through a cigar-shaped object" (Landau & Jackendoff 1993, 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 post-positions shows that there is no simple universal inventory, and the meanings can be very specific; for example, "in a liquid," "astraddle," "fixed by spiking" (Levinson & Meira 2003) – or distinguish "to (a location below)" versus "to (a location above)" versus "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 do not 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)[1]

Chomsky (1965, pp. 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, because 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. 537 ff.). This is an abstract principle meant to explain the difference between the grammaticality of the sentence (6) and (7), below, versus 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?

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

The child somehow has to extrapolate that (6) and (7) are okay, but (8) is not, 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. 615 ff.). 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, hopefully 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*, whereas 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 $\text{John}_x/\text{him}_x$.

This works well for English and hundreds, perhaps thousands, of other languages, but it 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 $\text{him}_{x,y}$
(11b) $\text{They}_{x,y}$ saw $\text{them}_{a,b/x,y/y,x}$.

Many languages (even Old English; see Levinson 2000) allow sentences like (11a) and (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 that 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.[2] 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; 1963b), 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, Greenberg 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 the "design features" of language are excluded – for example, 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.[3] 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.).

Second, Greenberg (1960, see also Comrie 1989: 17–23) distinguished the different *types* of universal statement laid out in Table 1 (the terminology may differ slightly across sources):

Although 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

Table 1. *Logical types of universal statement (following Greenberg)*

	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*

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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, for example, 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, and so on, 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., Perniss et al. 2008; Perniss & Zeshan 2008; Schwager & Zeshan 2008; Zeshan 2006a; 2006b). 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); and (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, for example, like the neurocognition of gesture, but instead recruits, for example, auditory cortex (MacSweeney et al. 2002; Nishimura et al. 1999). These results show that our biological endowment for language is by no means restricted to the input/output systems of a particular modality.

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 those by 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 earlier) and that all languages have vowels. The problem with the notion "all languages have vowels" is that it does not extend to sign languages (see 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 Colarusso 1982; Halle 1970; Kuipers 1960), 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. Hence, 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 is 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 7,000 odd languages we happen to have now, but the half million or so that have existed, not to mention those yet to come – because we never have all the data in hand, the one counterexample might never show up. In fact, Dryer points out, because 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 in 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 counterexample 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 & 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 organizational 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 minimize 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 long-term 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 return to this issue in section 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, and so on. 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 is more efficient to mark the less common negatives, and (b) it is 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 so on – 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" (sect. 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.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 that 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 properties just described do not 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 organization 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. Because 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, under review).

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

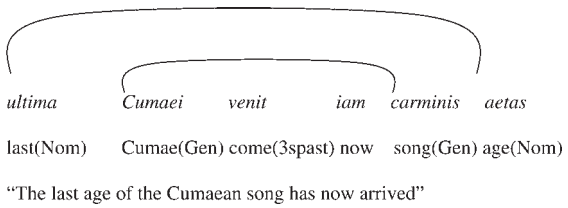
In nearly all recent discussions of syntax for a general cognitive science audience, it is simply presumed that the syntax of natural languages can basically be expressed in terms of constituent structure, and thus the familiar tree diagrams for sentence structure (Hauser et al. 2002; Jackendoff 2002;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

In the recent debates following Hauser et al. (2002), there is sometimes a conflation between constituent structure and recursion (see, e.g., 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). The 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):

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

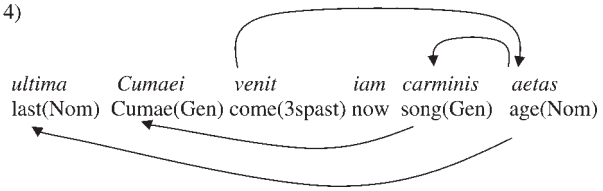
(13)



"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)



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 and Thalanyji, 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). Example (15) illustrates a discontinuous sequence of words in Thalanyji, 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.

Although 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 emphasizing, 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 given earlier (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 (PSG) (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 earlier, 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.[4] 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 (sect. 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 that presumes the universality of constituent structure will be false. Models of the evolution of language (e.g., Bickerton

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

1981) that presume the operation of phrase-structure grammar (PSG) generating sentences with surface constituency (Hauser et al., 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, because 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 et al. (p. 1569) hypothesize that recursion is "the only uniquely human component of the faculty of language."[5] 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 et al. 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, because 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, Vol 2, pp. 134 ff.).

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 (Gazdar et al. 1985; Partee et al. 1990), 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, e.g., 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 et al.'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 and 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 et al. (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 versus 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 in Bininj Gun-wok. 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; 1995c). 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 shall 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)	*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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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 Everett's 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, and Fitch, by Pinker and Jackendoff (2005). Fitch et al. (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" (p. 203). 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 (see Jackendoff 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, and Fitch versus Pinker and 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 et al's (2005) response that this is of no relevance to their selection of

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

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, that is, 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.
—Martin 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 & Kirby 2003). Some of this is devoted to the preconditions – for example, the origin of human cooperation (Boyd & Richerson 2005; Tomasello 2000, 2008), or the properties of human interaction or theory of mind (Enfield & Levinson 2006). Other work is centrally concerned with the coevolution 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.[6]

Language diversity can best be understood in terms of such a twin-track model, with the diversity 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 versus 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 quantification of inheritance versus 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 (Atkinson & Gray 2005; Gray & Atkinson 2003). 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, codependent (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" (quoted in Croft 2002, p. 5).

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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

emerges as frequently used patterns sediment into conventionalized patterns (Bybee 2000; Givón 2008). 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 7,000 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 (Dunn et al. 2008; Evans 1995c).

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 section 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 – for example, 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 preceding 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 emphasize 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 and 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 article 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 a 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 (Collard & Foley 2002; Mace & Pagel 1995; Nettle 1999; Pagel 2000). 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 2000, p. 28), multiple levels of structure (phonological, morphosyntactic, semantic) bridging sound to meaning, linearity combined with nonlinear 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, e.g., 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

augmented by general cognition and general learning principles to accommodate the additional structures of a specific language (as in, e.g., Elman 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 cross-linguistic 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 coevolution explores the twin-track descent mechanisms of culture and biology, and the feedback loops between them (Boyd & Richerson 2005; Durham 1991; Laland et al. 2000; Odling-Smee et al. 2003). Language is one of the best exemplars of such coevolution, 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 convergent 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 organizational 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 synthesis like the World Atlas of Language Structures (WALS) (http://WALS.info/) and the structural phylogenetics database (Reesink et al. 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 color 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 – for example, computational simulations of the importance of these distinct factors along the lines reported by Steels and Belpaeme (2005).

5. The dual role of biological and cultural-historical attractors underlines the need for a coevolutionary 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. Initiative: 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, although 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, for example, to echolocation in bats (pace Jackendoff 2002, p. 79). Echolocation is an ancient adaptation shared by 17 families (the Microchiroptera) with nearly 1,000 species and over 50 million years of evolution (Teeling et al. 2005), whereas language is an ability very recently acquired along with spiralling culture in perhaps the last 200,000 to 300,000 years by a single species.[7] 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

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

vocal tract itself. The null hypothesis here is that all needed brain mechanisms, outside the vocal-tract adaptation for speech, were co-opted from pre-existing 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, first, to show how the full range of attested language systems can evolve and diversify as sociocultural products constrained by cognitive constraints on learning, and second, 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

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; Guo et al. 2008; Li et al. 2006; Nakayama et al. 2006; Slobin 1997a).

Modality: Sign versus Speech

The modality transfer in sign versus 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) and Werker and Tees (2005) have exploited cross-linguistic differences in sound systems, showing that infants from 6 months of age 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 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, for example, 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 and Ferreira (2005).

Semantics and Conceptual Structure

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

Lexicon

Lexical gaps can impact perception (e.g., of color: Davidoff et al. 1999; Kay & Kempton 1984) 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 1997a). Controlled comparison across languages can be very revealing of children's learning strategies (Pye et al. 2007).

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

the brain – behavioral 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 behavior 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 behavior running on what are, at a "zoomed-out" level of granularity, closely related biological machines, and intra-cultural uniformity of behavior 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 behavioral variation, a population biology perspective, is going to be the key to understanding these central puzzles.

ACKNOWLEDGMENTS

This work was partially supported by the Australian Research Council (Social Cognition and Language Grant DP878126) to Nicholas Evans, and by the Max Planck Society to Stephen Levinson. Author order is alphabetical. We thank Jos van Berkum, Lera Boroditsky, Morten Christiansen, Peter Hagoort, Martin Haspelmath, Wolfgang Klein, Asifa Majid, Fritz Newmeyer, and Karl-Magnus Petersson for helpful comments; Barbara Finlay, Adele Goldberg, Russell Gray, Shimon Edelman, Mark Pagel, and two further anonymous BBS referees for their suggestions on improvement; and Edith Sjoerdsma for her work on the manuscript.

NOTES

1. 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."

2. 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, Vernes et al. 2008). Whether the higher, more abstract aspects of language reside in language-specific abilities, however, remains quite unclear (but see Box 2).

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

4. For example, 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.

5. There is some lack of clarity 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, vol. 1, p. 113). 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 Parker 2006; Pinker & Jackendoff 2005, p. 203).

6. For examples of ongoing coevolution in the language domain, see Dediu and Ladd (2007) and Aoki and Feldman (1989).

7. Any date for the origin of a graduated capacity will be arbitrary, 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).

Open Peer Commentary

Language universals: Abstract but not mythological

doi:10.1017/S0140525X09990604

Mark C. Baker

Department of Linguistics, Rutgers University, New Brunswick, NJ 08901.
mabaker@ruccs.rutgers.edu
http://www.rci.rutgers.edu/~mabaker/

Abstract: I present the so-called Verb-Object Constraint as a serious proposal for a true linguistic universal. It provides an example of the kind of abstraction in linguistic analysis that seems warranted, of how different languages can confirm such a universal in different ways, and why approaches that avoid all abstractness miss important linguistic generalizations.

I agree with Evans & Levinson (E&L) that natural languages display an amazing diversity, and that this has great significance for cognitive science. Like them, I have devoted myself to studying this diversity. But I do not agree that the descriptive data becoming available refute the Chomskian notion of a rich Universal Grammar (UG). On the contrary, each new language I have studied presents *both* fascinating new examples of diversity *and* important new evidence that human languages are all variations on the same theme. E&L are looking at only one side of this paradox.

The root of the paradox is that – according to generative linguistics – natural languages are properly described at several distinct levels of analysis (Chomsky 1957; 1965). Given this, it is perfectly possible for new languages to demonstrate diversity at one level and uniformity at another. It is easy to imagine how "comparative architecture" might find that the façade designs of buildings vary much more than expected, whereas fundamental structural support systems vary very little. This is parallel to what I believe comparative linguistics shows us about language.

As an example (synopsized from Baker 2001; Baker, in press), consider (1) as a possible linguistic universal:

(1) *The Verb-Object Constraint (VOC)*: A nominal that expresses the theme/patient of an event combines with the event-denoting verb before a nominal that expresses the agent/cause does.

One obvious reflection of the VOC in English is that (normally) objects immediately follow the verb, the two forming a verb phrase, whereas subjects do not. Comparison with other languages shows that many (35%; Dryer 2005) are like English in this respect – but also that many have a different word order, or no fixed order at all. At this level of description, there is substantial diversity. Nevertheless, we find other reflections of the VOC in language after language. For example, some

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

40% of languages (Turkish, Japanese, etc.) have Subject-Object-Verb order, with the object left-adjacent to the verb rather than right-adjacent to it. This is different from English but equally consistent with the VOC.

Languages with Verb-Subject-Object order (7%) look like violations of (1), but careful research has shown that many of them (at least) are not. For example, Irish clearly has a (Subject + [Verb + Object]) structure when there is no tense marking; Verb-Subject-Object order arises only when the verb needs to combine with tense (McCloskey 1991); see also Chung 1998 on Chamorro. These languages thus support the VOC, understood as holding at an abstract level.

Mohawk is a language in which word order gives no evidence of constituency. But Mohawk also has noun incorporation, where an argument can combine with the verb to form a compound. Crucially, the theme-object of the verb can incorporate into the verb, but the agent-subject argument cannot: Mohawk (*and* Bininj Gun-wok) allows "The baby meat-ate" but not "*Baby-ate the meat." This is another manifestation of the VOC, the compositional asymmetry showing up at the level of compounding rather than phrase formation (Baker 1988; 1996).

Unusual features of newly described languages can reveal universal properties in new ways. Kayardild has the very rare property of copying tense marking onto every constituent related to the verb phrase (Evans 1995a, pp. 399–401). Strikingly, this marking shows up on the theme-object but not on the agent-subject (see E&L's example [16]) – which is new support for the VOC. Including minor types (like Verb-Object-Subject languages), we now have solid leads that the VOC is valid for well over 90% of the known linguistic diversity. And I know of no counterexamples that have been investigated directly by mutually correcting research communities that include some researchers open to using abstractness.

The descriptive and typological research that E&L draw on summarily rejects most abstractness in linguistic analysis. E&L say it is a misconception that the differences among languages can be resolved by postulating a more abstract formal level, declaring this to be a false dogma. But the only support they give for this declaration is saying that "the experts either cannot formulate it clearly or do not agree that it is true" (target article, sect. 2). There is a real issue underlying this: It is a serious intellectual challenge to find exactly the right formulations of principles like the VOC (or Subjacency, or the Binding Principles). Generativists thus offer different formulations, and they do not claim to have found the definitive ones yet. But these formulations share a common core. Saying that UG is false on these grounds is thus like saying (as some do) that evolution is false because experts disagree about the details.

The challenge and opportunity of finding the right statement of universals can be seen in the putative contrast between "dependency" languages and "constituency" languages, which E&L emphasize. I find it striking that the dependency relations they identify for Latin in their example (14) are *exactly* the same as the dominance relations in the phrase structure of the English equivalent in example (13). We can thus isolate something substantive that these allegedly different language types have in common by finding a neutral mode of representation that expresses this important topological equivalence.

E&L also voice the widespread concern that abstractness allows generative theories to immunize themselves from counterexamples. No doubt this happens. But the VOC is not an unfalsifiable dogma for generativists. On the contrary, they have seriously considered alternative possibilities. For example, Marantz (1984) proposed that the VOC is a feature of language that varies parametrically, to account for "deep ergative" languages like Dyirbal. This hypothesis was investigated, but the preponderance of evidence showed it to be false, as more data came in from languages like Inuit (Bok-Bennema 1991). That the VOC is universal is simply the hypothesis that has fared better empirically than any well-articulated alternative.

Note that if the VOC is universal, this is certainly of great interest to cognitive science. *Why* should verbs combine with their theme arguments before their agent arguments? It is easy to write formal languages that do it the other way around. Presumably this tells us something contingent and potentially profound about how humans mentally represent events.

Linguistic universals are thus not myths, but hypotheses – hypotheses that gain new support from much of the same research that E&L cite. This research shows that we cannot be superficial in our approach to language, not that we cannot be universalist.

Widening the field: The process of language acquisition

doi:10.1017/S0140525X09990616

Edith L. Bavin

School of Psychological Science, La Trobe University, Bundoora, Victoria 3083, Australia.

e.bavin@latrobe.edu.au

http://www.latrobe.edu.au/psy/staff/bavine.html

Abstract: Evans & Levinson (E&L) argue against Universal Grammar on the basis of language diversity. A related and fundamental issue is whether the language input provides sufficient information for a child to acquire it. I briefly discuss the more integrated approaches to language acquisition which focus on the mechanisms, and research showing that input cues provide valuable information for the language learner.

Evans & Levinson (E&L) argue against the notion of Universal Grammar (UG), illustrating the huge diversity in human language – a diversity that exists at all levels. A justification for UG was the assumption that language is too complex for a child to acquire and that the input does not contain sufficient evidence for the child to learn the system; prior knowledge of language was therefore assumed. Hence, information about the process of language acquisition is fundamental in any discussion of UG and the domain specificity of language, something that E&L do not elaborate on.

Acquisition data are available from a range of languages, including those represented in the cross-linguistic studies pioneered by Dan Slobin (1985a; 1985b; 1992; 1997a; 1997b) and from more recent publications. The data indicate that young children quickly attune to the input language at all levels. Young children are adept at identifying the recurrent patterns of organisation within the input language, whether these are related to case systems, tense and aspect systems, word order, syntactic alternations, complex sentences, pronominal systems, the encoding of spatial relations, or other features. Based on the findings, the proposed UG principles and parameters are not adequate to explain the acquisition process. Nor do they convince all researchers that the language input lacks sufficient evidence for acquisition without prior knowledge of language structures, or that children from different language environments follow the same path in acquisition.

The crucial question is: What does the child bring to the task? A related question is: What factors influence language development? In the UG approach, the focus is on the end state, the mature grammar. Abstract linguistic concepts are available to the child, and language forms in the input are mapped onto these concepts. Both absolute and relative universals have been proposed, candidates for possible innate syntactic content (Valian 2009, p. 18). The relative universals allow for variation across languages including syntactic features and categories from which languages are built and from which they can select, as well as proposed binary parameters of syntactic variation

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

(Valian 2009, p. 19). Given the diversity of syntactic features and categories in languages discussed by E&L, there would need to be numerous relative universals to accommodate them. However, the child will need to process the input to determine which were relevant to the target language.

Many researchers in the field reject the formal UG approach; they do not assume domain (language)-specific constraints to facilitate language acquisition. Rather domain-specific knowledge emerges as a product of development. Their concern is with the mechanisms, processes, and strategies involved in acquiring a language or languages. The assumption is that language development cannot be isolated from the child's brain development or social and cognitive development. In the emergentist approaches, language structures are not innate; they emerge from known processes linking "a growing understanding of the brain with new theories of cognition" (MacWhinney 1999, p. xvii). The child uses the cues available in the input to identify the language specific patterns (Bates & MacWhinney 1987), with some cues more reliable than others. In constructivist usage-based approaches, children are assumed to build up syntactic categories and structures of their language gradually, using cues such as frequency and regularity of specific constructions (e.g., Lieven et al. 2003; Tomasello 2003a; 2009). Instead of assuming that the input lacks sufficient cues for the child to acquire the language, the research focuses on which cues it *does* provide and the cognitive and perceptual tools brought by the child to the task of acquisition.

A large proportion of the research designed to test proposed UG principles has focused on complex syntax. However, by the time children are processing complex structures, they have vast experience with their language and the contexts of use. In developing a language, new knowledge is built on existing knowledge. In the initial stages, perceptual biases, attentional mechanisms, and cognitive abilities are involved in processing the rich information provided in the input language. Rhythmic and distributional information provide cues to segmentation (Jusczyk 1997; Werker & Curtin 2005; Werker & Tees 1984). Research on statistical learning (e.g., Saffran et al. 1996) shows that young infants are sensitive to language-specific transitional probabilities, correlational probabilities, and distributional features of the input (Höhle et al. 2004; Mintz 2006; Thiessen 2009). The developing sensitivities allow for segmentation of syllables, words, and other grammatical units of the input language, segmentation that is an essential precursor to acquiring the system. As shown by Kuhl (2004), as infants become attuned to the sound contrasts of their environmental language, reorganisation of their perceptual abilities takes place; similarly, infants' developing statistical knowledge influences what they later perceive from the input. Thus, knowledge is advanced as they map sound sequences to meaning and retain these mappings in memory (e.g., Hollich et al. 2000) and as they identify category membership, for example, by linking the language context to properties of referents in particular domains (Smith 1999). Similarly, in the later stages of acquisition, knowledge of language structures gained facilitates the acquisition of new knowledge.

Phonological memory is important in forming mental representations of new words (Gathercole & Baddeley 1989), and vocabulary development is a precursor of vocabulary development (Bates & Goodman 1999). By assuming that language acquisition is guided by universal principles specific to the language domain, as in the UG approach, the role of cognitive skills and the influence of individual cognitive abilities on language development are not adequately considered. However, the link is clear from typical language development as well as atypical, an example of which is specific language impairment (SLI). It was first proposed that SLI supports separation of language from other cognitive domains, and explanations for language problems in SLI were related to principles of UG (e.g., Rice & Wexler 1996). Although children identified as having SLI are judged to have nonverbal abilities in the normal range, a significant body of research has revealed memory and information processing deficits (e.g., Archibald & Gathercole 2007; Bavin et al. 2005; Montgomery et al. 2009). In addition, significantly lower scores on standardised cognitive assessments are typically reported for SLI groups compared to age-matched, non-impaired children. Thus, an alternative explanation is that cognitive deficits lead to difficulty in processing information from the input, information required in acquiring the language (Leonard et al. 2007).

Theorists need to understand more about the diversity of languages, such as discussed by E&L, and the impact that such typological features have on the acquisition process; and, in addition, develop a greater understanding of language in atypical situations. Such understanding can only advance discussion about constraints on human language.

Unveiling phonological universals: A linguist who asks "why" is (inter alia) an experimental psychologist

doi:10.1017/S0140525X09990628

Iris Berent

Department of Psychology, Northeastern University, Boston, MA 02115-5000.
i.berent@neu.edu
http://www.psych.neu.edu/people/faculty/berent.html

Abstract: Evans & Levinson (E&L) are right to hold theories of language accountable for language diversity, but typological data alone cannot determine the structure of mental phonological grammars. Grammatical universals are nonetheless testable by formal and experimental methods, and the growing research in experimental phonology demonstrates the viability of a comparative experimental evaluation of the Universal Grammar (UG) hypothesis.

There is little doubt that the twin challenges of language universals and language diversity are critical for understanding the architecture of the language faculty, its domain-specificity and evolutionary origins. Despite their crucial import, these questions remain unaddressed in most existing psycholinguistic research. Evans & Levinson (E&L) should be commended for reminding the cognitive science community of its outstanding intellectual debt in this area. Nonetheless, E&L's own conclusion – that the hypothesis of universal grammar is false – does not follow from the evidence they present. Here, I specifically consider E&L's analysis of phonological universals – the role of syntactic and semantic universals falls beyond the scope of this commentary.

In its bare minimum, the hypothesis of Universal Grammar (UG) states that the brains of all speakers represent a shared set of grammatical constraints. Although this hypothesis is often associated with the claims that UG constraints are innate, and domain- and species-specific, these additional claims are not logically linked to the basic hypothesis of grammatical universals. E&L appear to reject all four claims on the grounds that language typology exhibits no absolute, exceptionless regularities. Typological universals, however, are distinct from grammatical universals, and the link between them is complex. Grammatical universals – the object of cognitive inquiry – are mental representations (I-language), whereas typological universals are statistical generalizations concerning external linguistic outputs (E-language). Such outputs are shaped by multiple factors, of which putative grammatical universals are only one force – the restrictions on perception, motor control, conceptual structure and memory, coupled with cultural and social factors, are equally strong determinants.

Consider, for example, the typological prevalence of CV syllables (discussed by E&L). One theory of UG, Optimality Theory

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

(Prince & Smolensky 1993/2004), attributes this fact to a universal, but violable, well-formedness constraint that requires all syllables to begin with an onset. Such a constraint, however, does not guarantee that CV syllables are most frequent typologically (typological frequency is also determined by extra-grammatical factors); nor does it preclude the existence of onsetless syllables (e.g., V – such syllables can be protected by other constraints enforcing faithfulness to grammatical inputs). Instead, the Onset constraint predicts that no grammatical process will actively transform syllables with an onset to onsetless ones (de Lacy, 2008). Whether the case of Arrernte (cited by E&L) counters this prediction is debatable (Berry 1998; Smith 2005), but typological frequency alone clearly cannot decide this matter.

Although this conclusion calls for a more careful interpretation of the typological diversity, it does not render the UG hypothesis unfalsifiable: optimality theory asserts that universal well-formedness constraints are active in the grammars of all speakers, irrespective of whether the relevant structures are present or absent in their linguistic experience. This strong hypothesis has sparked a productive research program that uses experimental tools to test the role of grammatical language universals – an enterprise that has unfortunately gone unnoticed by E&L. The available findings suggest that speakers are sensitive to putatively UG restrictions unattested in their language while ignoring other regularities that are equally motivated on statistical and phonetic grounds (Becker et al., submitted; Davidson 2006; Hayes et al., submitted; Moreton 2008; Wilson 2006).

Consider, for example, the restrictions on onset clusters (e.g., *bl* in *blocks*). It is well known that onsets such as *bl* are typologically more frequent than *lb*, and languages that tolerate syllables like *lba* tend to allow *bla*. This fact is attributed to sonority – a scalar property that correlates with the intensity of consonants: least sonorous (softest) on the scale are stops (e.g., *b,d*), followed by nasals (e.g., *n*) and liquids (e.g., *l*). Accordingly, *bla* rises in sonority, whereas *lba* manifests a sonority fall. The typological preference for onsets like *bl* is captured by a scalar UG constraint that favors onsets with large sonority distances (e.g., *bl*>*bn*> *bd*>*lb*, where>indicates preference; Clements 1990; Smolensky 2006).

Although sonority restrictions are widely documented, the typological evidence reflects only implicational tendencies, and many languages manifest outright reversals (e.g., Russian allows sonority falls, e.g., *lb*). Such observations might lead E&L to conclude that sonority restrictions are not grammatical universals, but rather, artifacts of modality-specific acoustic and articulatory preferences. However, sonority-based restrictions have been documented in sign languages (Corina 1990; Sandler & Lillo-Martin 2006), and recent experimental work suggests that they are active in the brains of individual speakers even when the relevant structures are absent in their language. English speakers, for example, favor syllables that rise in sonority (e.g., *bnif*) compared to sonority plateaus (e.g., *bdif*), which, in turn, are preferred to sonority falls (e.g., *lbif*), and this preference shapes the perception of these syllables (Berent 2008; Berent et al. 2007): the worst-formed onsets of falling sonority (e.g., *lbif*) are more likely to be misperceived (as *lebif*) compared to sonority plateaus (e.g., *bdif*), which, in turn, are misperceived relative to sonority rises (e.g., *bnif*). The misperception of ill-formed onsets is not due to an inability to extract their surface form from the acoustic input (e.g., it obtains with printed materials; Berent et al. 2009); nor is it explained by the statistical properties of English – similar results have been reported among speakers of Korean, a language that arguably lacks onset clusters altogether (Berent et al. 2008). Instead, these systematic misperceptions imply broad grammatical knowledge that triggers the active recoding of ill-formed structures. The convergence of sonority preferences across phonological systems, both spoken and signed (Sandler 1993), is consistent with a domain-specific phonological mechanism.

Nonetheless, the present results cannot determine whether phonology preferences are, in fact, universal or innate. Markedness (i.e., well-formedness) hierarchies, such as sonority, could vary in detail due to both predictable grammatical processes (e.g., conflation; de Lacy 2006) and variation in fine-grained phonetic properties that could inform their inference (Hayes & Steriade 2004). Whether phonological markedness hierarchies are experience-independent or learned is unknown, and there is vanishingly little information on their domain- and species-specificity. Far from being untestable, however, these questions call for a comparative cross-linguistic research program that combines typological, formal, and experimental methods. The emerging field of experimental phonology demonstrates the viability of this approach in evaluating the UG hypothesis.

ACKNOWLEDGMENT

This research was supported by NIDCD grant DC003277.

Language evolution: Two tracks are not enough

doi:10.1017/S0140525X0999063X

A. Charles Catania

Department of Psychology, University of Maryland, Baltimore County (UMBC), Baltimore, MD 21250.

catania@umbc.edu

http://www.umbc.edu/psyc/personal/catania/catanias.html

Abstract: This commentary argues that Evans & Levinson (E&L) should expand their two-track model to a three-track model in which biological and cultural evolution interact with the evolution of an individual's language repertories in ontogeny. It also comments on the relevance of the argument from the poverty of the stimulus and offers a caveat, based on analogous issues in biology, on the metaphor of language as a container, whether of meanings or of other content.

I welcome the arguments of Evans & Levinson (E&L) and offer only some supplementary remarks perhaps relevant to the future elaboration of their case. I strongly agree that, given the shared properties of the world we live in, our appeals to innate concepts and structures should be a last resort. My primary recommendation is that they expand their model from two tracks to three, so that they can deal with the ways in which biological and cultural evolution must necessarily interact with the evolution of an individual's language repertories in ontogeny. I also comment on the relevance of the argument from the poverty of the stimulus, and I offer a caveat, based on analogous issues in biology, on the metaphor of language as a container of meanings or other content. I have elsewhere addressed related issues, including the primacy of behavior in shaping brain structure (Catania 2008), the interpretation of learning in terms of selection rather than in terms of associations (Catania 1995; 2003; Catania & Shimoff 1998), and the distinction between language structure and function and its implications (Catania 1973; 1990; 2001; 2004; Catania & Cerutti 1986).

E&L make a persuasive case for the interaction of biological and cultural evolution. But languages can survive only if acquired and maintained by individuals who then pass them on. This acquisition and maintenance involves a third variety of selection operating on individual behavior, at every level: from the shaping of complex vocalizations as the child masters phonetic structures, through the semantic and syntactic details of those languages, to the sophisticated social interactions of mature speakers. Without selection at the level of individual language repertories, those linguistic and social practices cannot last long enough to be transmitted across successive generations. For this to work, by the way, ontogenetic and cultural selection must operate upon similar kinds of linguistic units; what is acquired in learning must share dimensions with what is culturally transmitted.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Adequate scenarios for language evolution must take into account these three varieties of selection. Each requires variations among the units upon which it operates; each entails different mechanisms by which environments select surviving variants. Biological contingencies must select requisite physiological attributes (e.g., vocal tract structure, neural organization). Ontogenetic contingencies must maintain the features of languages acquired by individuals. Cultural selection must perpetuate language practices across generations. E&L's discussion of language diversity nicely accommodates these processes, because it allows for substantial ranges of variation while still showing how languages can be maintained within and across generations.

These different varieties of selection often complement one another, but selection at one level may oppose selection at another. For example, behavior selected through ontogenetic processes during an individual's lifetime may reduce the individual's reproductive fitness at the phylogenetic level (consider substance abuse or sexual behavior involving partners at risk for AIDS).

Although they do so sparingly, and typically only implicitly, E&L sometimes discuss language as a container of meanings or of other content expressed in words and sentences. The conduit metaphor of language is pervasive but too often misleading (Reddy 1979). Assuming that language expresses or contains something transmitted from one individual to another raises some of the same difficulties as the assumption that genetic materials transmit information. Dawkins (1982) made that point when distinguishing between blueprints and recipes in characterizing genetic materials, which do not contain body plans or other information about the organism that will be their product. Rather than blueprints, they are recipes for development; the relevant information is not about past selective environments or about the way mature organisms will interact with their environments. A parallel distinction is relevant to language and has implications for language evolution. Words are not carriers of meanings or schemas or other information relevant to the selective environments that engendered them; we should look at words not in terms of what they contain but rather in terms of what they do (Catania 1987). Of all that they can do, one irreducible function of language is providing a mechanism by which one individual can change the behavior of others. This alone may be a sufficient basis for the origins of language. Other functions of language (e.g., communication, calculation, narrative, truth) may then be its derivatives.

I was surprised that E&L did not mention Chomsky's argument from the poverty of the stimulus (Chomsky 1965), because claims that the relevant structures cannot be found in the environmental contingencies justify the reification and internalization of language universals. The argument that children learn grammatical structures even in the absence of examples of negative instances seems to assume that learning any distinction requires exposure to all possible negative exemplars. But, for example, a bird raised in monochromatic light does not discriminate among wavelengths. Yet after learning to discriminate between the original wavelength and just a single new one not seen before, the bird responds in an orderly way to wavelengths across the entire visual spectrum (Terrace 1975). Similarly, any verbal environment including contrasts between several grammatical structures provides the differences on which learning can be based: for example, both active and passive voice are grammatical but provide a contrast that may enable later discrimination of grammatical versus ungrammatical.

Of course, these outcomes tell us about the structure of the underlying continua (the orderliness of the color continuum could be described as its universal grammar). But with regard to grammatical dimensions, those structures are so highly variable across languages, as E&L so aptly argue, that it would be as reasonable to assume that they are driven by language cultures within lifetimes in social environments as that they are driven by brain structures evolved over biological time (Catania 2008; Moerk 1992). This is consistent with evidence that rich verbal environments in which parents spend lots of time interacting verbally with their children make vast and lasting differences in their verbal competence, as measured by vocabulary growth rate, school performance, and IQ scores (Hart & Risley 1995). Verbal environments matter.

The myth of language universals and the myth of universal grammar

doi:10.1017/S0140525X09990641

Morten H. Christiansen[a] and Nick Chater[b]

[a]*Department of Psychology, Cornell University, Ithaca, NY 14853, and Santa Fe Institute, Santa Fe, NM 87501;* [b]*Division of Psychology and Language Sciences, University College London, London, WC1E 6BT, United Kingdom.*

christiansen@cornell.edu

http://www.psych.cornell.edu/people/Faculty/mhc27.htm

n.chater@ucl.ac.uk

http://www.psychol.ucl.ac.uk/people/profiles/chater_nick.htm

Abstract: Evans & Levinson (E&L) argue that language universals are a myth. Christiansen and Chater (2008) have recently suggested that innate universal grammar is also a myth. This commentary explores the connection between these two theses, and draws wider implications for the cognitive science of language.

It has been widely argued that an innate Universal Grammar (UG) must be postulated to explain two key observations: first, that languages share putatively "universal" patterns, which appear arbitrary from a functional, communicative point of view; and second, that children acquire language so readily from an apparently impoverished linguistic input (the "poverty of the stimulus" argument).

The second point has been the subject of considerable recent interest, with many theorists arguing that linguistic input is richer than has previously been suspected (e.g., Pullum & Scholz 2002; Reali & Christiansen 2005) or that modern learning methods are richer than is often presumed (e.g., Chater & Vitányi 2007; Harman & Kulkarni 2007). The first argument, based on language universals, has gone relatively unchallenged in the cognitive science literature – but no longer. Evans & Levinson (E&L) provide powerful evidence that language universals are myth rather than reality, and hence, that this line of defense of UG is swept aside. It remains to be explained, though, how languages came to display such stunning diversity, and this is where research on language evolution may offer some insight.

We have recently argued (Christiansen & Chater 2008) that an innate UG is not merely poorly evidenced, but indefensible on evolutionary grounds. Specifically, we argue that the cultural variability of language provides a "moving target," which changes too rapidly to support the biological adaptation that would be required to lead to an innate UG (Chater et al. 2009). Thus, language is best viewed as the product of cultural evolution, not biological evolution (Christiansen et al., in press).

The cultural evolution of language does not, of course, take place in a biological or social vacuum, but rather, is shaped by multiple constraints. One type of constraint derives from the nature of the thoughts that language expresses. For example, whatever the nature of our mental representations, they apparently afford an infinite range of different thoughts, promoting the likely emergence of compositionality in language (Kirby 2007). Linguistic structure is also shaped by socio-pragmatic principles relating to the communicative function of language;

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

for example, as embodied by Gricean implicatures (Grice 1967). A further source of constraints on language evolution derives from the operation of our perceptuo-motor apparatus, which, for example, enforces substantial seriality in both spoken and signed languages. Similarly, cognitive limitations on learning, processing, and memory also provide strong constraints on linguistic structure; for example, our limited working memory promotes a general tendency to resolve ambiguities as quickly as possible in both linguistic (Clark 1975) and perceptual input (Pomerantz & Kubovy 1986).

Individual languages are seen as evolving under the pressures from these constraints, as well as cultural-historical factors (including language contact and sociolinguistic influences), resulting over time in the kind of linguistic diversity described by E&L. Cross-linguistically recurring patterns do emerge due to similarity in constraints and culture/history, but such patterns should be expected to be probabilistic tendencies, not the rigid properties of UG (Christiansen & Chater 2008). Thus, we construe recurring patterns of language along the lines of Wittgenstein's (1953) notion of "family resemblance": although there may be similarities between pairs of individual languages, there is no single set of features common to all languages.

This perspective on language evolution and universals has important implications for language acquisition and processing. The ready learnability of language is explicable not because language fits an innate UG within each child; but rather, because language itself embodies patterns that are most naturally acquired from past generations of learners. We have argued, more generally, that learning cultural forms (C-induction) is very much easier than learning aspects of the natural world (N-induction) – because learning merely requires that each new generations agrees with the previous generation. For example, the number sequence 1, 2, 3... could be continued in any numbers of ways using repetition (1, 2, 3, 1, 2, 3, 1...), oscillation (1, 2, 3, 2, 1, 2...), as a Fibonacci sequence (1, 2, 3, 5, 8...), or some entirely irregular pattern (1, 2, 3, 73, 0, 18...). In the context of N-induction, the learner faces real difficulties: In encountering the sequence in some aspect of the natural world (e.g., the number of planets observed on successive nights), it is very difficult to know how the sequence will continue. However, C-induction is much more reliable and straightforward. The vast majority of people would find it most natural to continue the sequence with "...4, 5, 6... ." Thus, predicting how *other people* will continue the sequence is relatively easy, at least if people have the same inductive biases. Similarly, children must extrapolate a language from the sample of language they encounter; but such extrapolation is likely to be correct, given that it is the result of prior extrapolations by previous generations of learners. Again, the learning problem is dramatically easier if the objective is to mirror a cultural form that has been learned by others. Indeed, through prior generations of cultural selection, the form itself will have been optimized to embody whatever inductive biases the learner may have (Chater & Christiansen, in press).

Our emphasis on C-induction in language *acquisition* dovetails with a usage-based approach to language *processing*. Connectionist cognitive science has for some time explored the computational implications of a usage-based approach to language in which constituency and recursion are not built into the architecture of the language system but rather emerge through learning as probabilistic generalizations (Christiansen & Chater 2003). Importantly, we have developed usage-based models of recursive sentence processing that are capable of constituent-like generalization and have quasi-recursive abilities comparable to human performance on a variety of complex recursive constructions (Christiansen & Chater 1999; Christiansen & MacDonald, in press). Thus, at least some aspects of cognitive science do fit with the picture of language outlined by E&L, in which recursion and constituency are not innately defined universals.

E&L's important paper will substantially shift the debate in the cognitive science of language. Cognitive scientists have often taken rigid language universals as a "given," to be explained by theories of language acquisition and processing (e.g., by the postulation of an innate UG). E&L make clear that this is a mistake – and that a much more nuanced view of the patterns in the world's languages is required. We argue that this perspective is consistent with the view that an innate UG is as much a myth as language universals; and that language should be viewed as primarily a product of cultural, rather than biological, evolution.

Syntax is more diverse, and evolutionary linguistics is already here

doi:10.1017/S0140525X09990653

William Croft

Department of Linguistics, University of New Mexico, Albuquerque, NM 87131-0001.

wcroft@unm.edu

http://www.unm.edu/~wcroft

Abstract: Evans & Levinson (E&L) perform a major service for cognitive science. The assumption of Chomskyan generative linguistics – that there are absolute unrestricted universals of grammatical structure – is empirically untenable. However, E&L are too reluctant to abandon word classes and grammatical relations in syntax. Also, a cognitive scientist can already draw on a substantial linguistics literature on variationist, evolutionary models of language.

Evans & Levinson (E&L) make an important point to cognitive scientists about language universals. The nature of language is seriously misrepresented by Chomskyan generative linguistics, which is all too often taken as the standard theory of language that must be accommodated by any interdisciplinary research in cognitive science. In particular, the Chomskyan emphasis on absolute unrestricted universals of grammatical structure flies in the face of overwhelming empirical evidence of the structural diversity of languages. This structural diversity has to be the starting point for any theory that attempts to come up with universals of language. I have only two additions to make to E&L's argument: First, their critique of syntax, the heartland of Chomskyan linguistics, is too timid; and second, the dynamic evolutionary approach to language that E&L advocate already has a lengthy pedigree in linguistics.

E&L's approach to syntax contrasts with their assertions about sound structure. E&L state that phonologists are abandoning the idea of a fixed inventory of possible human speech sounds in favor of a model where languages construct sound systems from fine phonetic details that vary substantially. But in syntax, they assume that word classes of the traditional sort (noun, verb, adjective, etc.) exist, and they respond to diversity by arguing that some languages do not take advantage of some of these classes or add others. This is very much like Jackendoff's notion of a syntactic toolkit, which the authors otherwise reject. In fact, E&L should abandon the assumption that nouns, verbs, and so on, have theoretical, cross-linguistic validity as part of a fixed inventory of word classes, just as they have abandoned the parallel assumption for sound inventories.

Linguists identify word classes by using syntactic criteria or tests. These criteria are the possible occurrence of words of a particular class in a set of grammatical constructions (including morphological inflectional contexts), and their inability to occur in other constructions. This is called the distribution of the words, or word class. But the fact of all languages, including English, is that the distribution of words is highly variable, so that there

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

are no large coherent word classes in any language (see Gross 1979). Thus, it makes no sense to say (as E&L do) that English words like *willy-nilly* and *heeby-jeebies* differ from ideophones in other languages in kind rather than degree, or that *lunch* and *train* in *do lunch* or *take a train* are just regular nouns; they are not (Wierzbicka 1982). Instead, the only cross-linguistically valid – and language-internally valid – conclusion is that each construction defines its own word classes; words do not enter into a small number of large classes; and noun, verb, and so on, should be abandoned as word classes (Croft 1991; 2001; 2007; 2009).

This is not to deny that there are language universals constraining syntax, just as one cannot deny universals governing sound inventories; but they are not in the form of universal word classes found in traditional grammar and generative grammar. Instead, there are typological prototypes for parts of speech reflecting the preference of particular propositional acts (reference, predication, modification) for particular semantic classes; these are basically the same as E&L's clusters or attractors.

A similar argument applies to E&L's analysis of so-called subjects. A division of functions into topic, agent, and pivot is insufficient to reflect the diversity of encoding participants in clauses and the categorization of participant roles across constructions with different functions. Unlike word classes and propositional acts, there is no simple typological prototype to account for the diversity of participant encoding in clauses. Yet there are universal patterns that constrain the types of grammatical relations used for different construction types (Croft 2001; Kazenin 1994) and that constrain grammatical relations relative to each other in a clause (Croft 1991).

The chief sense in which the dynamic, complex adaptive systems approach to language that E&L advocate is now emerging (as they put it) is that it is finally emerging from under the shadow of generative linguistics. Cognitive science must embrace language diversity and the consequences it entails for our understanding of the nature of language. But cognitive scientists need not wait for linguists to draw those consequences; there is much existing research to guide a nonlinguist cognitive scientist interested in the approach to language presented by E&L. A few examples are given here, linked to E&L's seven theses in their conclusion.

Linguistic diversity (thesis 1) is the basic starting point of typological analysis. An important omission in an article asserting the centrality of linguistic diversity is the existence of diversity (variation) within speech communities, not just across speech communities (see Labov 1994; 2001; Milroy 1992). Numerous databases and inventories of typological diversity (thesis 2) already exist. Although there are clusters (thesis 3) or attractors (thesis 4) governing the distribution of some linguistic diversity (typological prototypes), these are only some of the patterns of typological distribution that have been observed, as mentioned earlier. Greenberg (1969; 1978a; 1979) presents a model of language that shifts away from synchronic language states to language history (thesis 7). A detailed evolutionary theory of language (theses 2 and 5) which integrates variation across languages, within speech communities, and in language use is presented in Croft (2000a) and in Mufwene (2001; 2008). Although this work is not coevolutionary, it balances the extensive recent work on the biological evolutionary origins of the language capacity by providing a cultural evolutionary framework for understanding language. Cognitive and functional constraints have been investigated in cognitive linguistics for a quarter of a century (theses 2 and 6; see Croft & Cruse [2004] and Langacker [2008] for recent overviews, and Tomasello [1998; 2003b] for an introduction oriented to cognitive scientists).

Much of course remains to be done in evolutionary linguistics. Cognitive science can draw on these strands of linguistic theory and empirical discovery, and evolutionary linguistics can draw on the rest of cognitive science for a better understanding of the factors that motivate speakers' creation and use of the linguistic structures that are found in the world's languages – before they disappear.

A note on methodology in linguistics

Robert Freidin

doi:10.1017/S0140525X09990665

Program in Linguistics, Princeton University, Princeton, NJ 08544-5264.
freidin@princeton.edu

Abstract: Evans & Levinson's (E&L's) critique of Universal Grammar fails because their methodology is flawed, as illustrated in their discussion of the Subjacency Condition. The lack of explicit analysis leads the authors to a false conclusion that is refuted by work published in this journal twenty years ago. They miss the point that unanalyzed data cannot disprove grammatical hypotheses.

From the outset, modern generative grammar has been concerned with the construction of precise models for linguistic structure. As Chomsky noted in the preface to *Syntactic Structures*, such models

> can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. (Chomsky 1957, p. 5)

One can evaluate competing proposals in linguistics to the extent that they are formulated precisely. In the case of the target article, this is not possible because there are no such proposals.

In general, the target article eschews precise analysis and offers instead a number of general descriptive statements about a range of languages that are taken to be refutations of proposed universals. Often these are offered without citation of data or, more importantly, the explicit analysis that would support the claims. Such refutations are based on unstated and unverified assumptions about grammatical analysis. In general, such discussion is not useful.

Consider the discussion of the Subjacency Condition in section 3, one of only two UG principles in generative grammar discussed.[1] Evans & Levinson (E&L) give no formulation of this condition; nor do they give explicit syntactic analyses of the examples cited in (6–8). Example (6) is actually ambiguous (i.e., 'where-said' versus 'where-got off') and Example (8), which is marked deviant, is grammatical under the interpretation of 'where-said'. The inadequate analysis of (8) offered in Example (9) does not distinguish between the possible and impossible interpretations because a place adverbial modifying the main clause VP occurs in the same linear position as one modifying the subordinate clause VP. Only the analysis of hierarchical structure disambiguates the two interpretations. Either *where* starts out as a constituent of the subordinate clause (the impossible interpretation) or of only the main clause (the possible interpretation). Moreover, if *whether* is analyzed as a complementizer on a par with *that* in Example (6) rather than a *wh*-phrase, then the Subjacency Condition does not account for the deviance of (8) on the 'where-get off' interpretation.

To understand E&L's criticism, one needs to know that the Subjacency Condition is a locality constraint that restricts the scope of syntactic movement, thereby contributing to an account of so-called island phenomena in natural language, a

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

central topic with an extensive literature.[2] E&L attempt to dismiss the Subjacency proposal on the grounds that it "does not work in Italian or Russian in the same way, and theorists have had to assume that children learn the specifics of the constraint after all" (sect. 3, para. 6). This is simply wrong, as illustrated for Italian in Rizzi's famous 1978 paper, which demonstrates that Subjacency holds in Italian but is parameterized in terms of which clause boundary type counts as a bounding domain for the constraint (CP for Italian, but TP for English)[3] (Rizzi 1978). The parametric analysis (including Russian, which apparently has no interclausal wh-movement) has been spelled out in detail in this journal; see Freidin and Quicoli (1989), which analyzes the poverty of stimulus issues for parameters, arguing that the choice of parameter settings for Subjacency cannot be made solely from presentation of data.[4] Thus, some parameters that account for variation across languages are also part of UG.

E&L conclude their discussion of Subjacency by saying, "Each constraint in UG needs to be taken as no more than a working hypothesis, hopefully sufficiently clearly articulated that it could be falsified by cross-linguistic data" (sect. 3, para. 6). The first part is a truism. However, the assumption in the second part that data alone can falsify any working hypothesis (especially one that works for a significant range of data) is clearly false. As Chomsky noted about Subjacency more than thirty years ago,

> It is easy enough to find phenomena that appear to violate the subjacency condition. Consider, e.g., the sentences (9), (10), where there is a relation between the phrase in bold face and the position marked by *t*, "violating" subjacency under the assumption that the rule in question is a movement rule:
>
> (9) ***John*** *seems* [$_{S'}$ *to be certain* [$_{S'}$ *t to win*]]
>
> (10) ***who*** *did Mary hope* [$_{S'}$ *that Tom would tell Bill* [$_{S'}$ *that he should visit t*]]
>
> Putting the matter more carefully, a proposed condition on rules, such as subjacency, cannot be confirmed or refuted directly by phenomena of this (or any other) sort. A condition on rules can be confirmed or refuted only by rules, which observe or violate it, respectively. If the rule of NP movement that yields (9) applies successive cyclically, as often assumed, then the rule will observe subjacency. If, as I have argued in the references cited, the rule of *wh*-movement applies successive cyclically, then it too will observe subjacency, giving (10). To find evidence to support or to refute a proposed condition on rules, it does not suffice to list unexplained phenomena; rather, it is necessary to present rules, i.e., to present a fragment of a grammar. The confirmation or refutation will be as convincing as the fragment of grammar presented. This is a simple point of logic, occasionally overlooked in the literature. The status of conditions on rules is empirical, but evidence can only be indirect and the argument, one way or another, is necessarily rather abstract and "theory bound." (Chomsky 1977, p. 74)

Data alone cannot speak to the validity of explicit proposals about the content of UG. What is required is an explicit analysis of data that follows from a precisely formulated fragment of a grammar. This is a comment about methodology, independent of any particular linguistic theory. In science there is no alternative to providing explicit analysis of data. The discussion of UG in this article misses the mark entirely.

NOTES

1. The other is Principle B of binding theory, although not identified as such nor explicitly formulated.

2. This is well understood within generative linguistics but possibly not to the wider readership of *Brain and Behavioral Sciences.* For recent accounts, see Chomsky (2008) and Boeckx and Hornstein (2008). The original work goes back to Chomsky (1964) and Ross (1967). See Chomsky (1973; 1977) for the original Subjacency proposal and subsequent modifications. Note that Newmeyer (2004), which is the only citation E&L give, also provides no explicit formulation for Subjacency.

3. The core principle is the same for all languages that manifest syntactic movement. Subjacency can be formulated as in (i):

(i) No movement rule can extract a constituent out of more than one bounding domain per movement operation.

NP and the clausal constituents CP and TP have been proposed as binding domains. See Note 2 for references.

4. For a similar parametric analysis of binding conditions, see Freidin (1992, Chs. 7 and 8), and the works cited there.

Essentialism gives way to motivation

doi:10.1017/S0140525X09990677

Adele E. Goldberg

Department of Linguistics, Princeton University, Princeton, NJ 08540.
adele@princeton.edu
http://www.princeton.edu/~adele

Abstract: The recognition that contentful universals are rare and often "banal" does not undermine the fact that most non-universal but recurring patterns of language are amenable to explanation. These patterns are sensical or *motivated* solutions to interacting and often conflicting factors. As implied by the Evans & Levinson's (E&L's) article, linguistics would be well served to move beyond the essentialist bias that seeks universal, innate, unchanging categories with rigid boundaries.

The Evans & Levinson (E&L) target article is a welcome concise summary of the problems and confusions created by facile claims of universality (see also Croft 2001). Given the authors' own extensive fieldwork on understudied languages, they clearly know the typological terrain. The article provides another reminder that the 50-year-old promise of a contentful "Universal Grammar" has gone unfulfilled.

And yet, many sympathetic linguists may be concerned that the picture portrayed appears a bit nihilistic. Surely the regularities of languages that have caught smart linguists' attention cannot be pure chance or complete accidents of language contact. In fact, as E&L note, recurrent patterns are solutions that satisfy "multiple design constraints, reflecting both cultural-historical factors and the constraints of human cognition" (target article, Abstract). This point is clearly central to the alternative approaches the authors and others espouse, and it is worth emphasizing.

A key idea is that most patterns of language are neither strictly predictable nor arbitrary: they are sensical or *motivated* solutions to interacting and often conflicting demands (cf. Lakoff 1987). We cannot predict that every language should have a given option, and we should not assume that because one language has the option, all languages do so "underlyingly." But neither are all the logical combinations of options attested, as we might expect if the options were arbitrary and simply learned from whatever unsystematic input a child might be exposed to. Importantly, we should not require that *all* languages show the same recurrent pattern in order to recognize the pattern as being motivated.

Haiman (1985) wrote that polysemy (one word form with more than one *related* meaning) can be defined as recurrent homonymy: that is, if the same two meanings are named by one label in unrelated languages, then it is reasonable to conclude that there is a motivated relationship between the two meanings. We do not require that the same polysemy exists in every language in order to understand that it is motivated by semantic and pragmatic factors.

The recognition of multiple motivating factors should make languages more, not less interesting to cognitive scientists. Instead of decreeing that a certain recurrent pattern exists just because it exists – because it is part of a "Universal Grammar" – we ask, how does the recurrent pattern develop and what independently needed factors conspire to motivate it? Explanation often comes from an understanding of historical processes, processing and developmental constraints, and a clear understanding of the functions of the constructions involved.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Recognition of conflicting motivations can reveal explanations as to why exceptions, as well as the generalizations themselves, exist. For example, regular morphological forms are compositional, but at times they license violations of general phonotactic constraints. For example, the consonant cluster /-lkd/ does not occur in root English words, and yet we tolerate it in *walked* and *talked*. Compositional morphology is motivated because it allows interpretations to be transparent, and it allows us to use old words in new ways. Irregular forms often satisfy one or more other demands, while violating compositionality. Irregular *made* allows for phonotactic regularity and reduction that would have been lacking if we used the word *maked* (Burzio 2002).

Unfortunately, providing "motivation" may fail to satisfy the way that prediction does. A normal reaction could be: "Surely necessary and sufficient criteria and simple predictive causes exist, even if we have not managed to discover them yet!" An understanding of this reaction may ultimately lead to a better understanding of why Universal Grammar has held such great appeal to so many for so long.

People, even young children, prefer explanations that appeal to *essentialist* categories (e.g., Gelman 2003). An essentialist category, C, combines several key ideas: (1) Though surface reality may differ, there exists an underlying invariant essence that defines C; (2) C has clear boundaries: one can determine categorically whether an entity is or is not a member of C; (3) the essence of C does not change over time; and (4) if C is construed as a biological category, the essence of C is assumed to be innately determined.

Essentialist explanations were sought throughout the biological and social sciences, perhaps on the model of mathematics or physics where such programs are often successful (but see Anderson 1972). And yet, most biological and social sciences have moved beyond the essentialist perspective. For example, evolutionary biologists in the nineteenth century thought of species as having an "essence" characterized by certain necessary, abstract features; but by the middle of the twentieth century, essentialism was replaced by "population thinking" (Mayr 1975) in which species were conceptualized as populations of individuals related to one another closely but in myriad complex ways. Similarly, within genetics, researchers had sought single genes that might be responsible for all kinds of syndromes and phenotypes, but more recently this view has given way to systems biology that rests on the recognition that most phenotypes are influenced by multiple interacting factors that include subtle environmental effects (Karmiloff-Smith 2006; Lander 1994).

Much of linguistics has remained in a stranglehold of the essentialist mindset, although important work in laboratory phonology and computational, experimental, and corpus linguistics has successfully moved beyond it. From definitions of word meanings to feature assignments for grammatical categories or constructions, the quest for simple, universal, unchanging, necessary, and sufficient conditions continues.

Yet such essentialist conditions are hard to come by, even for categories within a single language. For example, we like to think English adjectives form a clear-cut set. Definitions might appeal to their semantic property of noun modification, their prenominal attributive distribution, or their appearance after copular verbs such as *seem*. And yet there are exceptions to each of these criteria. For instance, a subclass of adjectives beginning with an unstressed schwa resists prenominal position (*??the asleep/alive boy*), the adjective use of quantifiers like *occasional* do not semantically modify the following noun (as in, *the occasional cigarette*), and adjectives do not all readily appear after *seem* (?*The idiot seems blithering*). Motivations for these exceptions exist (see e.g., Coppock 2008; Goldberg & Boyd 2009), but the fact remains that the category of English adjectives requires a more nuanced, multifaceted characterization.

E&L wonder aloud why the notion of a Universal Grammar has remained such a compelling notion within our field, despite the lack of consensus about what exactly it is. The lure of essentialism may provide a part of the explanation: "Universal Grammar" (UG) provides an essentialist answer to the question, *What is human language?* As with all essentialist explanations: (a) UG is assumed to offer an underlying invariant essence of language, although surface realities are recognized to differ; (b) UG is assumed to have clear boundaries: all human languages, and no other communication systems, are assumed to share UG; (c) UG is viewed as unchanging over time; and (d) UG is assumed to be innately determined.

Each of these assumptions is controversial at best and nonsensical at worst (e.g., Christiansen & Chater 2008; Goldberg 2006; Tomasello 2004). Emergent phenomena, widespread statistical generalizations, and multi-source interacting causes undermine the essentialists' program, as does the sort of widespread variation documented in the E&L article. It is time we let go of the essentialist mindset and embraced motivation as linguistic explanation.

The universal basis of local linguistic exceptionality

doi:10.1017/S0140525X09991130

Daniel Harbour

Department of Linguistics, Queen Mary University of London, London, E1 4NS, United Kingdom.

harbour@alum.mit.edu

http://webspace.qmul.ac.uk/dharbour

Abstract: Evans & Levinson (E&L) claim Kiowa number as a prime example of the semantically unexpected, threatening both Universal Grammar and Linguistic Universals. This commentary, besides correcting factual errors, shows that the primitives required for Kiowa also explain two unrelated semantically unexpected patterns and derive two robust Linguistic Universals. Consequently, such apparent exceptionality argues strongly for Universal Grammar and against E&L.

According to Evans & Levinson (E&L), the search for semantic universals is misguided because "many languages make semantic distinctions we . . . would never think of" (sect. 2.2.5, para. 3): no theory of universals can capture local linguistic exceptionality. This commentary shows that the exact opposite is true: apparent exceptionality provides highly robust evidence for Universal Grammar (UG). Here I summarize E&L's semantic argument, expose its fundamental flaw, then empirically refute their claim: two primitives are proposed that explain not only their primary example of "semantic distinctions we . . . would never think of," but derive two Universals and two further semantic Surprises.

Consider E&L's argument from semantics. For limited counting, languages may use noun/verb marking instead of numerals: for example, English *king~king-s*; Sanskrit *rājan~rājan-au~rājan-as* "one~two~many kings." Faced with singular~plural versus singular~dual~plural, accounting for number variation hardly appears challenging. However, Kiowa, say E&L (mis-citing Mithun), possesses "a marker that means roughly 'of unexpected number': on . . . 'man' it means 'two or more,' on . . . 'leg' . . . 'one or more than two, . . . on 'stone' . . . 'just two'" (sect. 2.2.5, para. 3). Thus, number systems present semantic distinctions too diverse for any universals to subsume. So, number variation refutes UG.

To reveal the flaw, consider chemistry. Common oxygen-carbon compounds (CO, CO_2) are odorless, colorless gases; but C_3O_2 is a colorless, odorous gas and $C_{12}O_9$, a white solid. Does such diversity mean these are not all constructed from the same two basic building blocks? No. Chemists simply recognize that atoms combined in surprising ways yield surprising properties. Analogously, linguists seek universal primitives and aim to explain cross-linguistic differences through the different ways these may combine. For Kiowa to refute UG, one must (a) look beyond surface appearances, (b) discover, by analysis, the

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

building blocks it uses, and (c) show that these are irreconcilable with any universal (non-Kiowa-specific) set of primitives. E&L do none of these: Like early chemists, their interest lies only in listing differences between substances, not in discovering how their components explain those properties.

Such criticism is more than methodological. It empirically undermines E&L's position. To explain Kiowa requires just two primitives (rooted in Thomas 1955), and these same primitives derive two robust Universals ($U_{1/2}$) and two semantic Surprises ($S_{1/2}$) from unrelated languages. Thus, local linguistic exceptionality, properly analyzed, argues strongly for UG.

UG number theory is basic set theory, so the two primitives concern sets. The first asserts ([+atomic]) or denies ([−atomic]) that a set is a singleton. The second asserts ([+augmented]) or denies ([−augmented]) that some given property of the whole set also holds of a subset.[1] English uses just [±atomic]: *king* is [+atomic], *kings* [−atomic]. Sanskrit uses both (momentarily ignore typefaces):

singular = [+*atomic* −***augmented***] — singletons without singleton subsets
dual = [−***atomic*** −***augmented***] — nonsingletons without nonsingleton subsets
plural = [−***atomic*** +*augmented*] — nonsingletons with nonsingleton subsets

For example, {*Henry*, *Richard*} is [−atomic] (nonsingleton) and [−augmented] (lacking [−atomic] subsets).

UG aims to explain constraints on cross-linguistic variation. The primitives just cited do precisely this, deriving robust two universals concerning the expression of number in natural languages. In broad outline:

(U_1) *Three is the highest number expressible grammatically in any language, without specialized numeral vocabulary* (contra quadrals, see Corbett 2000). Derivation: *three kings* (e.g., {*Alfred*, *Henry*, *Richard*}) is a collection (hence [−atomic]); it has nonatomic subsets ({*Alfred*, *Henry*}) (hence [−atomic + augmented]); but no such subset itself has nonatomic subsets (hence [−atomic +-augmented −augmented]). Therefore, [−augmented +augmented −atomic] = 3. Moreover, by basic set theory, addition of extra primitives is either contradictory or redundant.[2] So, trial (3) is the highest the number primitives of UG can reach.

(U_2) *Languages with dual invariably have singular/plural* (Corbett 2000), because the primitives required for dual ([±atomic], [±augmented]) also generate singular/plural (see Sanskrit above).[3] Likewise, *languages with trial invariably have singular/dual/plural* because the resources for trial ([±atomic], [±augmented], opposing +/− specifications) include [±atomic]/[±augmented], which generate dual, and so on.[4] (Harbour, in press, derives further analogous results.)

Important as such Universals are, the true test of these primitives, and UG, is whether they capture systems "we ... would never think of." In fact, these capture two semantic surprises E&L do not discuss, in addition to Kiowa.

(S_1) Although $2 \neq 1 + 3$, *languages may create dual from singular + plural*. Consider Hopi "this (one) ran/these (ones) ran" (Hale 1997). Plural *puma*, "these," and singular *wari*, "ran," combine to create dual *puma wari*, "these two ran," as follows:

singular	= ***pam***	***wari***
Dual	= ***puma***	***wari***
Plural	= ***puma***	***yùutu***

The semantic surprise that $1 + 3 = 2$ follows from the primitives. The boldfacing in both tables is identical: dual/plural share [−atomic]/*puma*; singular/dual share [−augmented]/*wari*.

(S_2) Although "we" is plural, *some languages treat the meaning {me, you} as singular*. Winnebago *-wĩ* pluralizes both *sh-'ũ*, "you do," and *hĩ-'ũ*, "you and I do," giving *sh-'ũ-**wĩ***, "you and other(s) do," and *hĩ-'ũ-**wĩ***, "you, I and other(s) do" (Lipkind 1945). Anomalously, singular "we" arises because Winnebago uses only [±augmented] (Noyer 1992). Consequently, "we" = {*me*, *you*, *other(s)*} is [+augmented] relative to "contains *me*, *you*": it and its subset {*me*, *you*} both contain *me*, *you*. However, "we" = {*me*, *you*} is [−augmented]: it contains *me*, *you* but has no proper subset that does. Thus, [±augmented] explains the semantically surprising singularity of some senses of "we."

Finally, Kiowa. First, correcting E&L's errors, unsuffixed, "man"/"leg" mean "one or two"; "stick," "more than one"; "hair," "just two"; "stone," "any number." Suffixation yields "more than two men/legs," "one stick," "one, or three or more, hair(s)" (e.g., *aa*, "sticks," *aa-**dau***, "one stick"; *aul*, "two hairs," *au-**dau***, "one, or three or more, hair[s]").

Nobody thinks groups of men odd, but pairs of hairs and solo legs normal. So, *-dau* does not mean "unexpected number of." Actually, Kiowa is like French, with two minor differences. Both languages organize their nouns into classes, but, where French uses a property of animates [±feminine], Kiowa uses number primitives: for example, the "gender" of *stick* is [−atomic]. This cleverly avoids the French problem of giving sticks anatomical gender. However, it encounters a different difficulty: talking about "one stick" (where "one" is [+atomic] but "stick" is [−atomic]). All *-dau* does is register when such +/− conflicts occur (French faces similar mismatches – when [+feminine] *la sentinelle* references a [−feminine] guard – but lacks an analogous suffix). Nothing more exotic need be said (see Harbour 2006 for Jemez and full details in Harbour 2007).

E&L claim that there are no Universals and that semantic Surprises vitiate attempts to discover UG. Both claims have been refuted here: two primitives suffice to derive two Universals, E&L's prime example of something "we ... would never think of," and two further Surprises besides. Linguists, like chemists, must look beyond surface appearances to discover universal building blocks and their properties. E&L's critique of this program succeeds only in highlighting some of its compelling results.

NOTES

1. Semi-formally: *S* satisfies [+atomic] iff $|S| = 1$; *S* satisfies [+augmented], relative to predicate *P*, iff *S*, and $S' \subset S$, satisfy *P*.

2. [+atomic −atomic] means $1 \neq |S| = 1$. [+augmented −augmented +augmented ...] = [+augmented −augmented ...] by Axiom: $\{a, a\} = \{a\}$.

3. Arbitrarily narrow focus misleads: Kiowa 'hair' (below) contrasts just dual/nondual, but Kiowa globally does not (cf. Nen, E&L's purported, unreferenced counterexample).

4. Occurrence of opposing +/− specifications varies cross-linguistically: Sanskrit does not permit them, Kiowa (below) does but not for counting. Both consequently lack trial.

The best-supported language universals refer to scalar patterns deriving from processing cost

doi:10.1017/S0140525X09990689

Martin Haspelmath

Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany.
haspelmath@eva.mpg.de
http://www.eva.mpg.de/lingua/staff/haspelmath/home.php

Abstract: Conditional universals have always interested linguists more than unrestricted universals, which are often impossible to demonstrate empirically because categories cannot be defined in a cross-linguistically meaningful way. But deep dependencies have not been confirmed by more recent empirical research, and those universals with

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

solid empirical support mostly relate to scalar patterns that can plausibly be related to processing cost.

For a long time, linguists have been more interested in general dependencies between different aspects of language structure than in categories shared by all languages, or, in other words, they have found conditional universals ("if a language has X, it also has Y") more interesting than unrestricted universals ("all languages have X") (cf. Table 1 in Evans & Levinson's [E&L's] target article). Although there are obvious similarities between the kinds of categories languages have, and linguists have found it possible to carry over category labels from one language to another ever since Greek grammatical terminology was adopted for Latin, there is something unsatisfactory about claiming, for instance, that "all languages have nouns and verbs." Clearly, all languages allow their speakers to talk about people, things, and events (nobody disputes this), and all languages have specialized words for meanings such as "mother," "tree," "sleep," and "hit." But how do we know whether the "tree" word in a language is a *noun*? The basic problem is that there is no way of defining structural categories such as *noun*, *article*, *dative case*, or *future tense* in such a way that the definition applies to different languages, other than via meanings that they share – but that these meanings are cross-linguistically comparable is not at issue. Thus, whether the Straits Salish words for "run" and "(be a) man" are verbs and nouns, or whether they are merely subclasses of a more general word-class "major word," is not decidable (Croft 2000b). There are no agreed upon criteria for nounhood, which makes debates about this rather pointless; and the same applies to all kinds of other structural categories (even *word*; cf. Haspelmath 2009).

Not all comparative linguists share this view, and E&L (sect. 2.2.4) seem to regard it as an empirical question whether the noun-verb distinction will be found to be universal; but in any event the literature has put much more emphasis on dependencies (or correlations) between structural patterns. Since the eighteenth century (cf. Plank 2001), linguists have often felt that given a certain property *X*, a language simply had to have property *Y* as well. Language structures were often seen as coherent systems in which a change in one place would necessarily entail changes elsewhere, perhaps multiple changes throughout the system. This view is particularly associated with the structuralist movement of circa 1920 to 1970, and a famous dictum is A. Meillet's *"une langue est un système où tout se tient"* ('a language is a system where everything hangs together'). Although Noam A. Chomsky's original interest in universals came from the claim that languages are unlearnable without rich innate structures, and he dissociated himself from structuralism, the 1980s and 1990s saw a strong movement of generative research into deep dependencies between different structures, due to a few "macroparameters," whose setting during language acquisition fixes a whole range of patterns (Baker 2001). In Pinker's (1994) words, children "are not acquiring dozens or hundreds of rules; they are just setting a few mental switches" (p. 112).

However, this work on macroparameters has not been a success, and it has largely been abandoned by mainstream generative linguists (cf. Haspelmath 2008b). Similarly, a large number of the proposals for deep connections between different parts of grammar that were made outside of the generative paradigm have not been confirmed by subsequent research. What has happened again and again is that a particular proposal for a correlation was based on a few languages (e.g., "languages with vowel harmony tend to have object-verb order," based on Turkish and Hungarian), but subsequent research on more languages showed that the generalization was too rash. Languages were found to cohere less than the intuitions of linguists indicated.

This does not mean, however, that universal dependencies ("conditional universals") are not more than a "myth." Although the structuralist and macroparametric hope that languages have an overall coherence across the entire system has not been confirmed, there are many generalizations with narrower domains that hold true, and for which there is massive empirical support. Consider the examples (1) to (6) (see Haspelmath 2008b for more discussion).

1. If a language has an [ø] sound (front rounded mid vowel), it also has an [e] sound (front unrounded mid vowel).
2. If a language has a [θ] sound (interdental fricative), it also has an [s] sound.
3. If a language allows two consonants at the end of a syllable, it also allows one or zero consonants at the end of a syllable.
4. If a language requires accusative marking on a patient object of some type, it also requires marking on patients of all types to the left on the following scale: speaker/hearer > 3rd-person pronoun > definite NP > indefinite specific NP > nonspecific NP.
5. If a language can form a relative clause on some clause position, it allows relativization on all other positions to the left on the following scale (Hawkins 2004): subject > object > oblique > possessor.
6. If a language uses a special reflexive pronoun for an adnominal possessor (e.g., Latin *suum*), then it also uses a special reflexive pronoun for the patient object (e.g., Latin *se*) (see Haspelmath 2008a).

All these universals, and quite a few more of a similar kind, refer to scales (or "hierarchies"; cf. Croft's [2003, p. 122] notion of "implicational hierarchies"). These scales can generally be interpreted in terms of processing cost. For example, the [θ] sound requires more articulatory effort than the [s] sound; the [ø] sound is less easy to distinguish from other vowels than the [e] sound; accusative marking is less redundant on patient objects that are pronouns or definite than on indefinite and nonspecific NPs; and reflexive use is more remarkable (and hence better coded) in patient object position than in adnominal possessive position (cf. *He$_1$ criticized himself$_1$* versus *He$_1$ criticized his$_1$ colleague*).

E&L primarily attack the Chomskyan rhetoric of "Universal Grammar" (UG) for understandable reasons, given the widespread acceptance of his approach outside of linguistics. But the cross-linguistic research of the last few decades has put to rest quite a few earlier hunches and claims by linguists about deep dependencies. What remains, however, is no less interesting from a cognitive science perspective than Chomsky's innate UG: Language structures can be shown to follow scalar patterns that in most cases have fairly straightforward explanations in terms of processing cost.

Animal comparative studies should be part of linguistics

doi:10.1017/S0140525X09990690

Daniel Margoliash[a] and Howard C. Nusbaum[b]

[a]*Departments of Organismal Biology and Anatomy, and Psychology, The University of Chicago, Chicago, IL 60637;* [b]*Department of Psychology, The University of Chicago, Chicago, IL 60637.*

dan@bigbird.uchicago.edu
http://margoliashlab.uchicago.edu
h-nusbaum@uchicago.edu
http://psychology.uchicago.edu/people/faculty/hnusbaum.shtml

Abstract: Universal Grammar promotes the study of an idealization of language behavior and language learning. In examining the diversity of actual behavioral strategies used to achieve linguistic goals, Evans & Levinson (E&L) move towards studying language as a behavior. This approach can benefit from studying communicative and cognitive capacities more broadly – across species. We exhort like-minded

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

linguists to cast off the remaining intellectual shackles of linguistic speciesism.

The fundamental premise for generativist linguists is that the diversity of languages springs from a common cognitive structure, latent in all humans, called Universal Grammar (UG). The assumption of this core cognitive construct, from which language emerges and by which language structures are constrained, is closely tied to assumptions that the structural framework of language is innate and unlearnable, that the processing capacities necessary for language are unique to humans, and that language is inherently a cognitively isolated module and therefore a computationally unique aspect of human cognition. This black box approach to a fundamental question in cognitive science ignores the potential of an evolutionary approach to cognitive processing, including such an approach to language. It also trains linguists to avoid consideration of evolutionary biology, and animal research in general. Evans & Levinson (E&L) address the first limitation head on but fall prey to the second. The good news is that it is a modest and easy extension of their work to complete the revolution.

By shifting the terms of scientific discourse from the unsupported and untestable assumption of a common core set of linguistic universals, a sort of mystical construction, to the actual distinctive behaviors of language populations, E&L move the study of language in the direction of the research methods of the experimental sciences and away from enclosed personal belief systems. The idea of studying real behavior as opposed to a platonic idealization also brings with it the obligation to consider real mechanisms that can explain that behavior. This calls into question the longstanding barrier of the competence/performance distinction to empirical study of language use and processing. This also brings with it a different standard for language research and what should count as evidence in understanding language processes.

This is particularly important in considering the issue of the human uniqueness of language. Hauser et al. (2002) have claimed that the single property of recursion as applied to vocalizations is the uniquely human computational capacity of the faculty of language. To address this point, E&L discuss the definition of recursion at length, from recursively enumerable patterns to nested recursion of patterns, and they allude to the computational differences among these. However, it is important to have clarity regarding what capacity is being described. If recursive syntactic processing in humans is a concept tractable to experimental science, then recursion refers not to the concept of putatively infinite (or, in any case, exceedingly long) patterns that no human produces or ever produced, as best as we know, but to an aspect of the computational complexity of pattern processes in actual human vocalizations. The real functional limits on what humans do produce and can perceive matter, and these should not be disregarded under the rubric of performance limitations independent of linguistic competence. Research with animals is irrefutably essential to address any claims of human uniqueness. In psychology and biology a single counterexample may not destroy a theory as it does in physics, but a whole class or classes of counterexamples can do so.

E&L refer to a study by Fitch and Hauser (2004) that reports the "impressive learning powers over FSGs [finite state grammars]" displayed by cotton-top tamarins, in spite of their failure to "grasp the patterning in PSG-generated sequences" (i.e., phrase-structure grammar generated sequences) (see target article, sect. 6, para. 3). Here E&L fall prey to becoming the "outsiders . . . [heeding] . . . the articulate envoys from the universalising generativist" (sect. 1, para. 4). E&L accept a null (PSG) result reported in tamarin study with species-irrelevant signals (human speech syllables), no clear learning procedure (mere exposure to stimuli), and with possible biases by additional cues (male/female voicing distinctions available for the FSG stimuli). Indeed, it might not have even been clear to the tamarins that there was a task since they were not behaviorally shaped to make a particular response.

The experimental paradigm that Fitch and Hauser (2004) used to distinguish tamarin failure of perception at the PSG level became broadly controversial only after it was used to demonstrate that operantly trained European starlings are able to learn both the FSG and PSG sequences (Gentner et al. 2006). Claims and counterclaims surrounding these results have appeared elsewhere and need not be repeated here. Nevertheless, one clear distinction of the starling results stands out in comparison to the tamarin studies. Whether one accepts or rejects the claims regarding recursive processing in assessing the learning of the PSG sequences, the starlings displayed far more complex vocal processing behavior than was displayed by tamarins. Our point is that animals that engage in vocal sensorimotor learning (starlings are songbirds and are excellent mimics) also demonstrate heightened vocal perceptual skills. Understanding the functional relationship between perceptual and motor skills, and the functional utility of such abilities in the daily life and social needs of an organism – including humans – is critical to assessing biological and computational capacities. This fact alone importantly informs the program of E&L, but about it they are mute.

E&L appropriately reach towards cladistic methods developed by evolutionary biologists to help sort out relations among human languages. They do not reach far enough. There are also examples of rapid changes in learned vocal patterns within nonhuman species based on cultural transmission of vocal patterns – famously studied in the songbirds (e.g., Marler & Tamura 1964). These studies have stimulated assessment of different learning models and extensive analysis of learning mechanisms. Applying this knowledge must be part of E&L's program, but they failed to recognize it as such.

As long as linguistics will be tied up in generativist nativism, animal studies won't matter and will contribute little. If linguistics truly becomes an experimental science, then animal studies are informative, and a synthesis of animal and human work becomes a useful goal to try to conceive.

E&L have staked out an important position: that there is an ethology of language, that it is available for study, and that it is essential that models of language use and mechanisms of language be inexorably related to actual language usage across humans. We agree, and maintain that this program is fundamentally rooted in experimental science and evolutionary biology. As such, it can be deeply informed by biological work on animal communication systems. Linguistics, and psychology, are subsets of organismal biology.

Variability in languages, variability in learning?

doi:10.1017/S0140525X09990926

Bob McMurray and Edward Wasserman

Department of Psychology and Delta Center, University of Iowa, Iowa City, IA 52240.

bob-mcmurray@uiowa.edu

www.psychology.uiowa.edu/faculty/mcmurray

ed-wasserman@uiowa.edu

www.psychology.uiowa.edu/Faculty/Wasserman/

Abstract: In documenting the dizzying diversity of human languages, Evans & Levinson (E&L) highlight the lack of universals. This suggests the need for complex learning. Yet, just as there is no universal structure, there may be no universal learning mechanism responsible for language. Language is a behavior assembled by many processes, an assembly guided by the language being learned.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

A rough heuristic can be deployed in comparative approaches to learning: If a behavior is fixed, it derives from the organism (e.g., early genetic or epigenetic processes); if it is moderately variable, it is learned, although possibly with constraints; if it is highly flexible, it is acquired by highly context-sensitive learning processes. Language may be the quintessentially flexible behavior. So, why have classical approaches focused so heavily on what might lie within the organism?

Evans & Levinson (E&L) offer a remarkable gift to cognitive science. In their sweeping synthesis of typological linguistics and their daring criticism of prevailing orthodoxy, they push us to focus on the flexibility and prowess of the language capacity, if there even is such a thing anymore.

Chomsky proposed Universal Grammar (UG) to ease the learning of language; yet, thanks to E&L, we see that in many cases even basic assumptions like syllabification or word order might set children learning certain languages down the wrong path. Although it is possible to soften UG by positing *limits* on the form of languages (as opposed to features that languages must have), even this maneuver may fail. Swedish, for example, violates the hallowed "size principle" in phonology: using two features to represent a binary contrast for which one would suffice (Helgason & Ringen 2008). Given the linguistic data described by E&L, we suspect that a UG based on limits will fare no better than one based on necessary features.

E&L, by questioning whether anything about language is universal (and therefore could be properly deemed innate), compel us to consider a flexible learning process that is capable of giving rise to a great deal of emergent complexity. So, a language learner may rely on a rich, redundant set of cues, using seemingly irrelevant speaker information to acquire phonology (Rost & McMurray 2009) and seemingly irrelevant phonological detail to acquire syntax (Christiansen et al. 2009), all in the context of a reinforcement environment that surpasses that of all other species in complexity.

Therefore, why do only humans acquire language? Hold-outs might argue that disparities in basic learning processes among species (Newport et al. 2004) are responsible. And, indeed, there is longstanding and compelling evidence that animals differ in basic, although perhaps parametric ways (Bitterman 1975). Yet, other than Skinner (1957), no other learning theorist has attempted to apply basic learning processes to language acquisition as a whole (although connectionist approaches have begun to address this piecemeal; e.g., Christiansen & Chater 2001).

One might also observe that basic cognitive capacities like categorization (Lazareva & Wasserman 2008), same-different discrimination (Cook & Wasserman 2006), sequence learning (Gentner et al. 2006), and expectation generation and evaluation (Keller & Hahnloser 2009) can be seen in various forms among diverse species. Yet, exhibiting many of these complex forms of cognition, all necessary for language, may be insufficient to move animals to a level of communication performance comparable to humans.

So where does that leave us? When confronted with such diversity of possible language inputs, one might first look for even more powerful learning mechanisms. But, perhaps language is not the product of a single mechanism. There is tremendous diversity of learning and cognition among animals (Wasserman & Zentall 2006). Finding one simple mechanism across many different species is highly improbable (Bitterman 1975). Given the diversity of human language so richly detailed by E&L, perhaps the parallel holds across humans.

This possibility would seem at first blush to undermine the view of humans sharing fundamental cognitive capacities. Yet, perhaps that is not so. Language may not be a monolithic structure to be internalized. Rather, language may be an ensemble of skilled behaviors that are shaped by a variety of forces and are sensitive to any available source of information. In this light, humans (much like evolution) assemble language with a variety of learning mechanisms and sources of information, this assembly being guided by the particularities of the language they are learning.

Learning even something as restricted as the phonology of English may demand mechanisms as diverse as unsupervised clustering (McMurray et al. 2009), dimensional weighting or cue neutralization (Rost & McMurray 2009), anticipatory learning of transitions (Elman 1990; Saffran et al. 1996), top-down feedback from the lexicon (Thiessen 2007), and nonspecific social contingencies (Goldstein et al. 2003). Perhaps none of them is necessary for mastery of all languages; but most of them are necessary for any given language. Still, none of these learning mechanisms could operate without a rich information source consisting of overlapping categories, information buried in the relative variation in particular cues, sequential relationships between sounds, and vertical relationships between sound and meaning. These relationships are provided by the language and the culture. Although this may appear circular, as is so often the case in evolution, it is merely evidence of coevolving systems. Language is not a structure to be internalized; rather, it is an extremely complex assembly of behaviors to be developed.

Is language a thing? Or is it a skill? If it is a skill, we should not be studying "structure" at all, but rather, studying moment-by-moment behaviors. Evolution does not select for genes, it selects for phenotypes, or rather, for the developmental processes that give rise to them (Lickliter & Honeycutt 2003). Similarly, in language, the fundamental unit is not a piece of structure – it is an ongoing behavior, the learning processes that give rise to it, and the cultural factors that make it useful. So what is language? It is a behavior . . . that develops . . . in a culture . . . that develops . . .

The evidence and argument of E&L have freed developmental theorists from having to explain the abstract and arcane of narrow linguistic theories; language can be virtually anything. The authors have moved us to a more productive consideration of what is central to language: how language is used and acquired, and what flexible and emergent mechanisms are needed to accomplish this remarkable behavior.

Returning language to culture by way of biology

doi:10.1017/S0140525X09990707

Bjorn Merker
Gamla Kyrkv. 44, SE-14171 Segeltorp, Sweden.
gyr694c@tninet.se

Abstract: Conflation of our unique human endowment for language with innate, so-called universal, grammar has banished language from its biological home. The facts reviewed by Evans & Levinson (E&L) fit the biology of cultural transmission. My commentary highlights our dedicated learning capacity for vocal production learning as the form of our language endowment compatible with those facts.

The synoptic sketch of language diversity presented by Evans & Levinson (E&L) suggests that the scope and depth of this diversity itself is a pivotal fact of language. The sheer variety in patterns of usage by which different languages map between linguistic form and meaning points to cultural transmission as the biological model of choice for human language. Handed down the generations in a family setting, languages not only provide a good match to the details of human population structure worldwide (Cavalli-Sforza 1997), but each one of them, linguistic isolates included, is in a state of perpetual change. Grammars (as the rest of language) undergo constant transformations in historical time, exhibiting shared modes of change driven by language *use*, modes that may in fact harbor the true laws of language (Bybee 2006; Greenberg 1969; Lehmann 1995).

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Add the fact that any human being, placed in any one of these language communities at birth, will acquire full competence in the local language, and the cultural model equates to two null hypotheses: Every extant and past mapping between linguistic form and meaning is learnable under the circumstances in which humans typically encounter them, and we come into the world equipped with the capacity to acquire, and eventually to transmit, such mappings across generations by cultural learning. The impression that language requires constraints on its *forms* (such as its grammar) other than those imposed by the logic of cultural transmission itself stems from a half-century old strategic confound introduced into linguistics by Noam Chomsky (1959).

One argument Chomsky opposed to behaviorist claims regarding language acquisition was a formal one: The sophisticated Suppes-Estes formalization of behaviorist learning theory lacks the power to acquire a grammar of human caliber (Chomsky 1975, in Piattelli-Palmarini 1980, p. 111). Although the logical possibility remained that behaviorist learning theory is inadequate *as learning theory*, Chomsky took the unlearnability of human grammar by its means to support the proposition that grammar is innate. Through his persistent efforts, the fact that humans have a unique biological endowment for language has come to be identified with a presumptive *innate grammar* to which learning contributes but "minor modifications that give one language or another, depending on experience" (Chomsky, in Harnad 1976), as if this were the only way to have an endowment for language.

There is neither need nor justification for biologically oriented work on human language to continue to accommodate this strategic confound. Cultural transmission delivers the restricted search space needed to enable language learning, not by constraining the form language takes on an innate basis, but by ensuring that the form in which language is presented to the learner is learnable (Zuidema 2003). The target of the new generation's learning process is itself the outcome of a learning process (previous generation). Gold's 1967 theorem – cited as recently as 2002 by Hauser, Chomsky, and Fitch (see Hauser et al. 2002) – accordingly cannot decide between learned and innate grammar (Zuidema 2003; see also Clark 2001; Horning 1969; Johnson 2004; Lappin & Shieber 2007). Moreover, across generations of learning agents, cultural transmission has the power to transform a state of arbitrary strings randomly paired with meanings into a shared and efficient grammar without intervention of natural selection or differential reinforcement of outcomes (Kirby 2002). This apparent magic results from competition among strings for access to subsequent generations via the "learner bottleneck," a central aspect of the "poverty of the stimulus" argument. In this competition more efficient and general forms tend to outlast others over generations.

What remains is to complete the cultural model with a way to sustain the chain of transmission of initial nonsense (unsemanticized, ungrammaticized strings) over the many generations required for convergence on a shared grammar. The ideal prior state for this can be found among the cultural traditions of complex learned song of many songbirds and a few mammals. Some of these are not confined to cultural variations on an innate pattern, the "species-specific song" recently invoked by Fitch (2008), but are emancipated from innate constraint on the form song takes (calls and songs of 76 species of birds from two continents have been identified in the repertoire of the marsh warbler: Baylis 1982; Dowsett-Lemaire 1979). Under circumstances explored in Merker and Okanoya (2007), a prelinguistic human adaptation for emancipated song would provide the mechanism to sustain string transmission for however long convergence might take, because it would be driven by the need to impress by elaborate vocal display rather than to communicate meaning.

But "*Something* about the faculty of language must be unique in order to explain the differences between humans and other animals" (Fitch et al. 2005, p. 182, emphasis in original). Indeed, this is a condition met by the very adaptation needed for song learning to be possible, as first proposed by Darwin (1871). It is a dedicated learning capacity missing in all other primates – our ape relatives included – but essential for every word and sentence we know how to pronounce, namely the capacity to learn to reproduce, by voice, patterns of sound originally received by ear, technically known as vocal production learning (Janik & Slater 1997). We share this capacity with some of the singers and all mimics among the birds, and with a few mammals, yet we alone, of all species, have evolved vocal learning in the setting of a primate brain.

Supported by a conformal motive (Merker 2005) and de novo evolution of a direct projection from primary motor cortex to the respiratory and phonatory motor nuclei of the lower brainstem (Brown et al. 2008; Okanoya & Merker 2007), vocal learning turns the cerebral territories centered on Wernicke's and Broca's areas from their non-language uses in other primates to the service of human language by recruiting them to the generative production and intergenerational transmission of culturally learned vocal lore. To it we owe not only our developmental trajectory for language learning, infant babbling included, but our propensity for imitation and ritual culture more generally (Merker 2005), along with a robust selection pressure for encephalization (Merker, in press; Merker & Okanoya 2007). As repeatedly urged by students of birdsong (Doupé & Kuhl 1999; Marler 1970; Nottebohm 1975), this distinctive capacity of ours for vocal learning holds the biological key to the singularity of human language. Through it, and unconstrained by innate so-called universal grammar, the historical filter of cultural transmission – which passes only the possible – continually adapts the actual forms of languages to multiple interacting constraints such as use, utility, learnability, and neural resources (Christiansen & Chater 2008), as well as cultural norms (Everett 2005), with diversity in train.

ACKNOWLEDGMENT

The author is indebted to Shimon Edelman for helpful comments on an earlier version of this commentary.

On formal universals in phonology

doi:10.1017/S0140525X09990537

Andrew Nevins

Department of Linguistics, Harvard University, Cambridge, MA 02139.
nevins@fas.harvard.edu
www.fas.harvard.edu/~nevins

Abstract: Understanding the universal aspects of human language structure requires comparison at multiple levels of analysis. While Evans & Levinson (E&L) focus mostly on substantive variation in language, equally revealing insights can come from studying formal universals. I first discuss how Artificial Grammar Experiments can test universal preferences for certain types of abstract phonological generalizations over others. I then discuss moraic onsets in the language Arrernte, and how its apparent substantive variation ultimately rests on a formal universal regarding syllable-weight sensitivity.

The target article by Evans & Levinson (E&L), questioning not only the existence of universals but the methodology of pursuing them, raises many issues for discussion. In this commentary, I limit my remarks to the importance of formal (as opposed to substantive) universals – a distinction E&L do not draw with sufficient clarity. Formal universals in phonology are constituted by the analytic elements that human minds employ in constructing representations of sound structure. Put simply, formal universals refer to the set of available data structures (e.g., binary features, metrical grids, autosegmental tiers) and the possible operations on them that can be used in constructing a grammar of a language.

Phonotactic dependencies of the form *Given segments A, B in the same word, if A has feature F, then B must have feature G* are

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

constrained by analytic restrictions on what can be referred to. These analytic restrictions are, by hypothesis, formal universals that are independent of the historical contingencies or cultural practices of any given language community. One of the best ways of studying formal universals of this kind is by constructing Artificial Grammar Learning experiments, using the methodology of cognitive science.

In one such recent study, Moreton (2008) conducted an experiment in which participants were taught a miniature artificial language containing phonotactic dependencies of the form outlined above. There were three conditions: in one, *F* and *G* were both vowel height; in a second, *F* and *G* were both obstruent voicing; and in a third, *F* was vowel height and *G* was obstruent voicing. Importantly, the rules of English phonotactics do not contain any of these three dependencies. The results, however, showed that the height-voice dependency was not learned by participants. Moreton's conclusion was that an analytic bias favors learning certain phonotactic dependencies over others; the resulting formal phonological universal is in (1):

(1) Learning phonotactic dependencies of the form *Given segments A, B in the same word, if A has feature F, then B must have feature G* is universally easier when *F* and *G* are the same feature than when *F* and *G* are different features.

Formal universals like (1) lend themselves to eminently more possibilities for integration with the cognitive sciences than E&L's proposed research program based on "the dual role of biological and cultural-historical attractors" (target article, sect. 8, para. 6, E&L's thesis 5). Formal universals allow for experimental testing in laboratory conditions under which the historical-cultural factors are completely controlled for, and hence irrelevant to the outcomes.

It is worth considering how apparent exceptions to universals are analyzed in other fields. As an example, consider the case of the Jacana bird, one of nature's species exhibiting a "sex-role reversal," whereby it is the females that engage in polyandry and cuckolding of the males. At the right level of analysis, the sex-role reversal in these shorebirds is entirely unsurprising, because it is the males that perform the raising of the chicks. The correct asymmetry between sexes is not that males have multiple mates while females do not, but rather, that the sex that commits to what biologists call "parental investment" is the one who is stuck in the harem. When the universal is correctly formulated, the Jacana bird is actually an exception that proves the rule.

I argue that E&L err in concluding that there are no universals within human phonology based on apparent substantive exceptions:

> But in 1999, Breen and Pensalfini published a clear demonstration that Arrernte organizes its syllables around a VC(C) structure and does not permit consonantal onsets(Breen & Pensalfini 1999). 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. (target article, sect. 2.2.2, para. 2)

Arrernte is not, as E&L would have it, but one language that recently "ruined the entire sample," so to speak. The question of VC syllabification in Australian languages was raised by Sommer (1970; 1981) on the language Oykangand, later insightfully analyzed in terms of onset-maximization by McCarthy and Prince (1986). There was, historically, a widespread loss of initial consonants throughout Australian languages, which Hale (1964) and Blevins (2001) attributed to stress shift and lenition processes. Although Arrernte was apparently no exception to this sweeping change, nonetheless, "25% of Arrernte words are pronounced in isolation with an initial consonant" (Breen & Pensalfini 1999, p. 2). To account for words such as *mpʷar* and *taŋk*, Breen and Pensalfini (1999) have to propose that these words have an underlying hidden initial vowel, a red-flag for any "clear demonstration" that the language disallows consonantal onsets.

In general, the deduction of which syllabification pattern a word contains depends on particular phonological processes that refer to syllabic divisions. In this light, consider the following formal universal:

(2) Stress assignment, weight-sensitive allomorphy, compensatory lengthening and prosodic morphology, when sensitive to distinctions among syllable types, refer exclusively to the representational unit of weight called the *mora.*

The phonological universal in (2), developed by Hyman (1985), McCarthy and Prince (1986), and Hayes (1989), is formal, not substantive in nature: it restricts the data structures that can be referred to by morphophonological processes, and is not about the substantive question of which segments can bear moras. In fact, Topintzi (2009) has gathered evidence from a wide range of languages demonstrating the ability of onset consonants to be moraic. The existence of metrical processes referring to onsets has been a topic of research for many years; see Davis (1988), Downing (1998), Goedemans (1998), and Gordon (2005), who discuss onset-sensitivity of stress in languages ranging from English and Italian to Pirahã and Iowa-Oto.

If vowels and onset consonants, but not coda consonants, are moraic in Arrernte, the statement of stress assignment and weight-sensitive allomorphy become quite straightforward in the light of (2). In Arrernte, stress is assigned within a word to the first vowel preceded by a consonant: *mpʷá.rem*, "is making," versus *i.kʷént*, "policeman." Since onset consonants are moraic, the stress rule is simple: the left-most bimoraic syllable receives stress. Similarly, the statement of plural allomorphy in Arrernte is simple: bimoraic-or-greater forms like *i.ŋel* and *taŋk* take the suffix *-ewar*, while monomoraic forms like *ar* and *aŋk* take the suffix *-erir*. The reduplication patterns can receive a similar treatment in terms of moraic targets, within the prosodic morphology framework: for example, the copying of VC strings to a reduplicant is driven by the demand to fill a bimoraic template.

Like the Jacana bird's sex-role reversal, which has a mechanistic and principled explanation at a different level of primitives (partner with greater parental investment, instead of male and female), the patterning of weight-sensitive process in Arrernte exhibits a principled conformity to a formal universal at the level of which consonants are moraic, instead of in terms of syllabification. Rather than positing a silent initial vowel for 25% of Arrernte words, attention to the statement of formal universals enables a consistent representational property for syllables throughout the language. The universal in this case pertains to the set of data structures that learners use to encode sound patterns: moras, and only moras, are the formal unit that can be referred to by weight-sensitive properties.

E&L trumpet the slogan "A linguist who asks 'Why?' must be a historian" (sect. 7, epigram, quoting Haspelmath 1999, p. 205). Integration with the cognitive sciences, however, will come from mechanistic explanations, not from hand-waving at diachronic contingencies. Formal universals are restrictions on representational vocabulary, and they rear their heads even when history deals them an odd shuffle or, as in the case of artificial grammar experiments, no historical shuffle at all.

Universal grammar and mental continuity: Two modern myths

doi:10.1017/S0140525X09990719

Derek C. Penn,[a] Keith J. Holyoak,[b] and Daniel J. Povinelli[a]

[a]*Cognitive Evolution Group, University of Louisiana at Lafayette, New Iberia, LA, 70560;* [b]*Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095.*

dcpenn@ucla.edu **http://www.cognitiveevolutiongroup.org/**
holyoak@lifesci.ucla.edu **http://reasoninglab.psych.ucla.edu/**
ceg@louisiana.edu **http://www.cognitiveevolutiongroup.org/**

Abstract: In our opinion, the discontinuity between extant human and nonhuman minds is much broader and deeper than most researchers

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

admit. We are happy to report that Evans & Levinson's (E&L's) target article strongly corroborates our unpopular hypothesis, and that the comparative evidence, in turn, bolsters E&L's provocative argument. Both a Universal Grammar and the "mental continuity" between human and nonhuman minds turn out to be modern myths.

If Evans & Levinson (E&L) are right (and we believe they are), the dominant consensus among comparative psychologists is as specious as the Universal Grammar (UG) hypothesis in linguistics.

At present, most comparative psychologists believe that the difference between human and other animal minds is "one of degree and not of kind" (Darwin 1871; for recent examples of this pervasive consensus, see commentaries on Penn et al. [2008] in *BBS*, Vol. 31, No. 2). Among researchers willing to admit that the human mind might be qualitatively different, most argue that our species' cognitive uniqueness is limited to certain domain-specific faculties, such as language and/or social-communicative intelligence.

We believe this view of the human mind is profoundly mistaken. In our opinion, the discontinuity between extant human and nonhuman minds is much broader and deeper than most researchers admit. We are happy to report that E&L's target article strongly corroborates our unpopular hypothesis; conversely, the comparative evidence bolsters E&L's provocative argument.

The case for a domain-general discontinuity. Linguists can (and should) argue about which human languages employ constituency and which do not. Or about which employ syntactic recursion and which do not. But from a comparative psychologist's perspective, the spectacular fact of the matter is that any normal human child can learn any human language, and no human language is learnable by any other extant species. Why? Why are human languages so easy for us to learn and so unthinkable for everyone else?

The standard explanation is that only humans have a "language instinct" (Pinker 1994), which fits nicely with the presumption that the rest of the human mind is more or less like that of any other ape (e.g., Hauser et al. 2002). But as E&L point out, the diversity of human languages suggests that our faculty for language relies largely on domain-general cognitive systems that originally evolved for other purposes and still perform these non-linguistic functions to this day.

If E&L are right, there should be significant differences between human and nonhuman minds outside of language. Conversely, E&L's case would be in bad shape if the nonverbal cognitive abilities of nonhuman apes turned out to be highly similar to those of humans.

We recently reviewed the available comparative evidence across a number of domains – from "same-different" reasoning and Theory of Mind (ToM) to tool-use and spatial cognition (Penn et al. 2008). Across all these disparate domains, a consistent pattern emerges: Although there is a profound similarity between the ability of human and nonhuman animals to learn about perceptually grounded relations, only humans reason in terms of higher-order structural relations. Nonhuman animals, for example, are capable of generalizing abstract rules about "same" and "different" relations in terms of the perceptual variability between stimuli. But unlike human children (Holyoak et al. 1984), they appear incapable of making analogical inferences based on structural, logical, or functional relationships. Nonhuman animals are able to predict the actions of others on the basis of the past and occurrent behavior of these others. But they are incapable of forming higher-order representations of others' mental states or of understanding others' subjective experience by analogy to their own (Penn & Povinelli 2007b; in press; Povinelli & Vonk 2003). Nonhuman animals can infer the first-order causal relation between observable contingencies and reason in a means-end fashion. But they show no evidence of being able to cognize the analogy between perceptually disparate causal relationships or to reason in terms of unobservable causal mechanisms (Penn & Povinelli 2007a; Povinelli 2000). And although nonhuman animals can learn the first-order relationship between symbols and objects in the world, they are incapable of cognizing combinatorial hierarchical schemas in any domain – causal, social, spatial, or symbolic.

So here is the central explanadum that most comparative psychologists and linguists seem to be avoiding: Why is it that the discontinuity between human and nonhuman forms of communication appears at the same degree of relational complexity as does the discontinuity between human and nonhuman cognition in every other domain?

The hypothesis that human and nonhuman minds differ only in their social-communicative intelligence (cf. Tomasello 2008) does not explain why the discontinuities between human and nonhuman minds in non-communicative tasks (e.g., causal reasoning) are just as profound as those between human and nonhuman forms of communication; nor does it explain why the human learning system is the only one that can cope with the relational complexity of human language. Furthermore, the evidence from high-functioning autistic populations demonstrates that normal relational intelligence can be preserved in the absence of normal social-communicative intelligence, whereas the converse is not the case (see Morsanyi & Holyoak [in press] and references therein). This suggests that our unique social-communicative skills rely on our unique relational intelligence – not the other way around.

The hypothesis that the communicative and cognitive functions of language played an important role in rewiring the human brain makes good sense to us (Bermudez 2005; Bickerton 2009). And it is clear that language still enables, extends, and shapes human cognition in many profound ways (Clark 2006; Loewenstein & Gentner 2005; Majid et al. 2004). But the evidence from comparative psychology points to the same conclusion as does that from comparative linguistics: It is not language *alone* that made or makes humans so smart (see our discussion in Penn et al. 2008). Rather, the diversity of human languages seems to have been shaped by the capabilities and limitations of the human mind (Christiansen & Chater 2008). And our species' unique relational abilities seem to be a necessary precursor for the advent of even the most rudimentary human language. As Darwin himself put it: "The mental powers of some early progenitor of man must have been more highly developed than in any existing ape, before even the most imperfect form of speech could have come into use" (Darwin 1871, p. 57).

The model of a physical symbol system (PSS; Newell 1980) provides a useful heuristic framework for understanding what happened to the human mind. According to our hypothesis (Penn et al. 2008), animals of many taxa evolved the ability to represent first-order symbolic relations; but only humans evolved the ability to approximate the higher-order capabilities of a PSS. As a result, only humans can make the kind of analogical, modality-independent, role-governed inferences necessary to master the spectacular complexity of human language (Holyoak & Hummel 2000). This capability is not specific to language, nor did it originally evolve for language. Rather, our domain-general ability to reason about higher-order relations co-evolved with and continues to subserve all our uniquely human abilities.

An object lesson for comparative research. We cannot help but notice an unfortunate parallel between the current state of comparative psychology and generative linguistics: in both cases, researchers have gone to elaborate lengths to minimize any possible differences from English-speaking humans.

To our mind, this is a gross travesty of Darwin's legacy. Wasn't the entire point of Darwin's (1859) magnum opus to show that the natural processes of variation and selection combined with "Extinction of less-improved forms" could create enormous differences between extant organisms, some so great as to constitute different species altogether? Human minds and human

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

languages may both be bio-cultural hybrids. But, in our view, the principles of evolution that apply to bodies apply, *mutatis mutandis*, to minds and languages.

Against taking linguistic diversity at "face value"

doi:10.1017/S0140525X09990562

David Pesetsky

Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139.

pesetsk@mit.edu

http://web.mit.edu/linguistics/people/faculty/pesetsky/index.html

Abstract: Evans & Levinson (E&L) advocate taking linguistic diversity at "face value." Their argument consists of a list of diverse phenomena and the assertion that no non-vacuous theory could possibly uncover a meaningful unity underlying them. I argue, with evidence from Tlingit and Warlpiri, that E&L's list itself should not be taken at face value – and that the actual research record already demonstrates unity amidst diversity.

From a distance, the structures of the world's languages do look gloriously diverse and endlessly varied. But since when is it sound strategy to take diversity at "face value"? All other sciences have progressed precisely by taking *nothing* at face value – diversity included. Evans & Levinson (E&L) claim, in effect, that linguistics is different from all other fields. If they are right, the search for deeper laws behind linguistic structure is a fool's errand, and languages are just as inexplicably diverse as they seem at first glance. It is thus surprising that E&L's article contains no discussion of the actual research to which they supposedly object. Instead, their article offers only (1) a parade of capsule descriptions of phenomena from the world's languages, coupled with (2) blanket assertions that each phenomenon falls outside the scope of all (non-vacuous) general theories of linguistic structure. The argument must therefore rest on two premises: (1) that the capsule descriptions of phenomena are correct, and (2) that these phenomena fall so obviously beyond the pale of research on Universal Grammar (UG) that no argument is necessary. I believe there are reasons for caution about the first supposition and that the second is simply wrong.

Confidence in the accuracy of E&L's examples is undermined at the outset by a comparison of E&L's claims with their sources. For example, as a demonstration that "semantic systems may carve the world at quite different joints" (sect. 2, para. 2), E&L cite Mithun (1999, p. 81) for an "unexpected number" marker (sect. 2.2.5, para. 3) in Jemez (misidentified by E&L as Kiowa). It is supposedly because "two" is an expected number of legs but not stones that the marker means "any quantity but two" when added to "leg" but "exactly two" when added to "stone." According to Mithun, however, the actual determinant is the Noun-Class to which the noun belongs, a partly arbitrary classification (like Indo-European gender). In Jemez, "nose" belongs to the same class as "leg," and "arm" belongs to the same class as "stone." Consequently, if the number marker truly indicates how Jemez speakers "carve the world," "one" must be an unexpected quantity of noses, but an expected quantity for arms (cf. Harbour 2006). Similarly, although E&L use Mundari ideophones to exemplify a "major word class" whose very existence "dilutes the plausibility of the innatist UG position" (sect. 2.2.4, para. 10), their source, Osada (1992, pp. 140–44), actually does not discuss ideophone syntax at all. Mayan positionals are mentioned for the same reason – but although they do constitute a distinct category of *root* (used to derive words belonging to standard classes), they do not appear to constitute a syntactically distinguishable *word* class – at least not in Mam (England 1983, p. 78), Tzotzil (Haviland 1994, p. 700), Chol (Coon & Preminger 2009), or Tzeltal (Brown 1994, p. 754; cited by E&L).

But even if E&L's sources had been quoted perfectly, there is a deeper danger in arguments founded on capsule descriptions. Because linguistic puzzles are complex, even the best descriptions are incomplete. Fundamental generalizations often remain undiscovered for years, until a creative researcher suddenly asks a question whose answer unlocks one of the language's secrets. And how does one find such questions? As it happens, among the most productive question-generating devices is the very idea that E&L deride: that languages share structural properties, often masked by one of their differences. Again and again, investigations guided by this possibility have altered our basic picture of the world's languages.

Compare E&L's assertion that "syntactic constituency . . . is not a universal feature of languages" (sect. 5, para. 8), supposedly supported by the existence of languages with "free word order," with the alternative – that such languages do not fundamentally differ from others in basic constituent structure, but merely allow these constituents to be reordered more freely. Although E&L attribute the appeal of this alternative to mere anglophone prejudice, its actual appeal rests on a body of hard-won *results*, achieved by researchers pursuing lines of inquiry that no one previously thought to pursue.

One recent example is Cable's (2007; 2008) investigation of *wh*-questions in Tlingit, a language of Alaska with considerable word order freedom. In languages with more rigid word order, there is a well-known split in how they form *wh*-questions. In languages like English and Italian, interrogative *wh*-phrases move to a dedicated position near the left periphery of the clause, precedable only by discourse topics (cf. *This book, who would read it?*; Rizzi 1997). In languages like Chinese and Korean, by contrast, there is no dedicated position for *wh*-phrases. Instead, they appear wherever their noninterrogative counterparts would:

(1) (a) English
Who did John see?

(b) Chinese (Huang 1982)

Zhangsan	kajian-le	shei
Zhangsan	see-asp	who

At first glance, Tlingit appears to pattern with Chinese rather than English in lacking a dedicated *wh*-position, as expected if free word order entails absence of constituent structure:

(2) Tlingit (Cable 2007; 2008)

a.

Yá x'úx'	aadóoch	sá	kwgwatóow?
this book	who-erg	Q	he.will.read.it

"Who will read this book?"

b.

Aadóoch	sá	yá x'úx'	kwgwatóow?
who-erg	Q	this book	he.will.read.it

(= a)

c.

Kéet	daa	sá	axá?
killer.whale	what	Q	he.eats.it

"What do killer whales eat?"

On the basis of his fieldwork and analysis of Tlingit texts, Cable showed that, contrary to appearances, Tlingit *wh*-phrases do occupy a dedicated *wh*-position – in fact, a position syntactically and semantically identical to its counterpart in languages like English. Specifically, Cable discovered that phrases to the left of an interrogative *wh*-phrase – "this book" in (2a) and "killer whale" in (2c) – are always discourse topics, and if a *wh*-phrase appears to the *right* of the verb, the result is not interpretable as a *wh*-question. These findings are explained if the interrogative phrase in both Tlingit and English must occupy the same

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

position at the left periphery, precedable only by a discourse topic – that is, if this aspect of Tlingit clause structure resembles English, despite first impressions. Cable's findings constitute a *discovery*, not anglophone prejudice. Earlier work on Tlingit, though excellent, had failed to spot the evidence for a *wh*-position. Cable's discoveries were made precisely because his investigation was informed by what linguists had learned about *wh*-questions in other languages.

Remarkably, the same type of discovery has been reported by Legate (2001; 2002, building on Laughren 2002) for Warlpiri, the free word order language par excellence. Legate not only found the same dedicated wh-position in Warlpiri (precedable by topics) but also showed that the full array of standard tests for phrase structure in languages like English yield the same results for Warlpiri – important findings directly attributable to the very research program that E&L consider bankrupt.

Questions like those addressed by Legate and Cable must be asked for every item in E&L's linguistic *Wunderkammer*. Are they truly as "different" as they seem? Sometimes, of course, the answer might be "yes." But how will we ever distinguish real differences from those that merely reflect our ignorance if we grant – even for an instant – E&L's dictum that apparent structural differences must just "be accepted for what they are" (sect. 1, para. 1)?

ACKNOWLEDGMENTS

I wish to thank Tom Bever, Seth Cable, Guglielmo Cinque, Jessica Coon, Daniel Harbour, and Massimo Piattelli-Palmarini for useful comments on an earlier draft.

The reality of a universal language faculty

doi:10.1017/S0140525X09990720

Steven Pinker[a] and Ray Jackendoff[b,c]

[a]*Department of Psychology, Harvard University, Cambridge, MA 02138;* [b]*Center for Cognitive Studies, Tufts University, Medford, MA 02155; and* [c]*Santa Fe Institute, Santa Fe, NM 87501.*

pinker@wjh.harvard.edu
http://pinker.wjh.harvard.edu
Ray.jackendoff@tufts.edu
http://ase.tufts.edu/cogstud/incbios/RayJackendoff/index.htm

Abstract: While endorsing Evans & Levinson's (E&L's) call for rigorous documentation of variation, we defend the idea of Universal Grammar as a toolkit of language acquisition mechanisms. The authors exaggerate diversity by ignoring the space of conceivable but nonexistent languages, trivializing major design universals, conflating quantitative with qualitative variation, and assuming that the utility of a linguistic feature suffices to explain how children acquire it.

Though Evans & Levinson (E&L) cite us as foils, we agree with many of their points: that the documentation of linguistic diversity is important for cognitive science; that linguists have been too casual in assuming universals and formulating a defensible theory of universal grammar; that the language faculty is a product of gene-culture coevolution; that languages are historical compromises among competing desiderata; and that cross-linguistic generalizations are unlikely to be exceptionless. We do, however, endorse a version of the Universal Grammar (UG) hypothesis: that the human brain is equipped with circuitry, partly specific to language, that makes language acquisition possible and that constrains human languages to a characteristic design.

E&L do not make it easy to evaluate their position. Their claim that linguistic variation is "extraordinary," "fundamental," "amazing," and "remarkable" is not just unfalsifiably vague but also myopic, focusing on differences while taking commonalities for granted. Any survey that fails to consider the larger design space for conceivable languages is simply unequipped to specify how "remarkable" the actual diversity of languages is. Consider these hypothetical languages:

Abba specifies grammatical relations not in terms of agent, patient, theme, location, goal, and so on, but in terms of evolutionarily significant relationships: predator-prey, eater-food, enemy-ally, permissible-impermissible sexual partners, and so on. All other semantic relations are metaphorical extensions of these.

Bacca resembles the two-word stage in children's language development. Speakers productively combine words to express basic semantic relations like recurrence, absence, and agent-patient, but no utterance contains more than two morphemes. Listeners use the context to disambiguate them.

The speakers of Cadda have no productive capacity: they draw on a huge lexicon of one-word holophrases and memorized formulas and idioms. New combinations are occasionally coined by a shaman or borrowed from neighboring groups.

The grammar of Daffa corresponds to quantificational logic, distinguishing only predicates and arguments, and using morphological markers for quantifiers, parentheses, and variables.

Fagga is a "rational" language spoken in a utopian commune. It lacks arbitrary signs larger than a phoneme, allowing the meaning of any word to be deduced from its phonology.

Gahha is a musical language, which uses melodic motifs for words, major and minor keys for polarity, patterns of tension and relaxation for modalities, and so on.

These and countless other conceivable languages are not obviously incompatible with cognitive limitations or communicative limitations (unless stipulated post hoc), and fall well outside the envelope reviewed by E&L.

Indeed, E&L concede an enormous stratum of universals in endorsing an innate core of Hockett-syle "design features": discreteness, arbitrariness, productivity, duality of patterning, auditory-vocal specialization, multiple levels of structure between sound and meaning (including morphosyntax and semantics), and linearity combined with nonlinear structure (constituency and dependency). But if this set of universal mechanisms were laid out explicitly, as computational machinery consistent with the details of actual languages, we doubt that it would differ "starkly" from the kind of Universal Grammar that the two of us have posited (Jackendoff 1997; 2002; Pinker & Jackendoff 2005; Culicover & Jackendoff 2005).

E&L's focus on differences of detail, moreover, leads them to exaggerate differences in underlying design. Whether a language's inventory of phonemes or verbs is large or small, whether grammatical agreement plays a role that is indispensable or vestigial, whether a noun-verb difference is obvious to a linguist in every language or subtle in a few, whether grammatical combination is florid inside phrases but rudimentary inside words, or vice versa, whether embeddings are found in a high or a low percentage of phrases – none of these pertains to underlying computational systems, just the extent to which different parts of it are deployed (which we agree depends on the vagaries of history).

The force of many of E&L's examples depends more on having a splitter's (rather than a lumper's) temperament than on the existence of "extraordinary" variation. English, for example, has phenomena similar to Chinese classifiers (e.g., *a piece of paper, a stick of wood*), Athabaskan verb distinctions (among locative verbs; Levin 1993; Pinker 1989; 2007), ideophones (response cries such as *yum, splat, hubba-hubba, pow!*; Goffman 1978), and geocentric spatial terms (e.g., *north, upstream, crosstown*; Li & Gleitman 2002). There are differences, to be sure, but it is unparsimonious to insist that they lack any common psycholinguistic mechanisms.

E&L try to minimize their own commitment to Universal Grammar in two ways. First, they try to trivialize it as "true by definition," claiming that any system lacking these features

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

would not be called a "language." But that is both factually false (the word "language" is commonly applied to bee dance, ape signing, computer programming protocols, etc.) and theoretically misleading. The proposition that all human societies have "language" *according to that definition* is neither circular nor trivial.

Second, the authors note that these features are functional, as if that suffices to explain how children acquire them. But functionally useful structures do not materialize by magic. Even after they have developed historically as part of a tacit convention within a community of speakers, a child still has to acquire them, including the myriad grammatical details which do not perceptibly contribute to that child's ability to communicate but are simply part of the target language (semi-regular inflectional paradigms, arbitrary grammatical categories, complex co-occurrence restrictions, etc.). This requires sufficiently powerful learning abilities, which have evolved as the "gene" side of the "gene-culture coevolution" we all embrace.

The construct "Universal Grammar" (UG), as we have invoked it, refers to these learning abilities. UG is emphatically not a compendium of language universals (though it has often been taken to be such). UG, in our conception, is a toolkit. This implies that it does not, as E&L claim, contain all possibilities for all languages: one builds structures *with* tools, not *from* tools. We have suggested that UG may be a set of "attractors" for grammars (Culicover 1999; Culicover & Jackendoff 2005; Jackendoff 2002) – strong biases on the child's learning that produce statistical tendencies in grammars, but that can be overcome by exceptional input, a possibility that E&L endorse.

These learning mechanisms are a black box for E&L, who suggest that they can be borrowed off the shelf from other domains of cognition. But they do not cite any analysis showing that some independently motivated set of cognitive processes actually suffices to acquire the detailed and complex structures mastered by children. In practice, attempts to model specific phenomena of language development using generic information-processing machinery (notably, connectionist networks) habitually smuggle innate linguistic structure in the back door in the form of unheralded design features and kluges (Marcus 2001; Pinker & Ullman 2002a; 2002b; 2003).

For all these disagreements, we applaud E&L's call for a more rigorous study of linguistic variation and universals. When combined with rigorous specification of the psychological mechanisms that make language acquisition and use possible, we predict that the resulting psycholinguistic universals that give rise to the diversity of human language will be far from "unprofound."

For universals (but not finite-state learning) visit the zoo

doi:10.1017/S0140525X09990732

Geoffrey K. Pullum and Barbara C. Scholz

School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, EH8 9AD, Scotland, United Kingdom.

gpullum@ling.ed.ac.uk

http://ling.ed.ac.uk/~gpullum

bcscholz@gmail.com

http://www.philosophy.ed.ac.uk/people/honorary-fellows/barbara-scholz.html

Abstract: Evans & Levinson's (E&L's) major point is that human languages are intriguingly diverse rather than (like animal communication systems) uniform within the species. This does not establish a "myth" about language universals, or advance the ill-framed pseudo-debate over universal grammar. The target article does, however, repeat a troublesome myth about Fitch and Hauser's (2004) work on pattern learning in cotton-top tamarins.

The take-home message from the target article by Evans & Levinson (E&L) is not in their title but in their subtitle. They don't show that language universals are a myth; their point is that what makes human languages really interesting for cognitive science is their diversity, not their uniformity. Boas would have endorsed this view, but it seems fresh and novel in the current context. You want species-wide universal grammar? Visit the zoo. Study puttynose monkeys (Arnold & Zuberbühler 2006), or your cat. What human beings bring to animal communication is not rigid universals but a flexible ability to employ any of a gigantic range of strikingly varied systems. That seems to be what E&L are saying.

Regrettably, though, the authors repeat a wildly false claim about results on syntactic learning in nonhuman primates. E&L were apparently misled by a statement in the literature they critique: they cite Fitch and Hauser (2004) as having demonstrated that cotton-top tamarins have "impressive learning powers over FSGs [finite state grammars]" (sect. 6, para. 3).

This meme is spreading, alarmingly: E&L even cite a paper by a brain researcher (Friederici 2004) who is looking for distinct neural systems for processing FSGs on the one hand and PSGs (phrase-structure grammar) on the other. The truth is that no one has shown monkeys to have any general ability to learn finite-state (FS) languages. It is extremely unlikely that anyone ever will. FS parsing is powerful; it would suffice for pretty much all the linguistic processing that humans ever do.

Reflect for a moment on the likelihood that tamarins could be habituated to strings matching this regular expression:

$$a\,(c^*d\,c^*d\,c^*)^*\,a + b\,(c^*\,d\,c^*\,d\,c^*)^*\,b$$

All such strings begin with either *a* or *b*; the middle is an indefinitely long stretch of *c* and *d* in random order, but always containing an even number of *d*; and strings end with whatever they began with (notice, an unbounded dependency!).

This language has a very simple FSG, but passively learning it from being exposed to examples (*acca*, *accdcccddcccccda*; *bdccccdb*; *acdccddcccddda* . . .) is surely not plausible. Figuring out the grammar would surely be way beyond the abilities of any mammal other than a skilled human puzzle-solver with pencil and paper.

People have unfortunately been confusing FS languages with a vastly smaller proper subset known as the strictly local (SL) languages (Pullum & Rogers 2006; Pullum & Scholz 2007; Rogers & Pullum 2007). Fitch and Hauser unwittingly encouraged the error by remarking that FSGs "can be fully specified by transition probabilities between a finite number of 'states' (e.g., corresponding to words or calls)" (Fitch & Hauser 2004, p. 377). The equation of states with words here is an error. States in FSGs are much more abstract. Languages that can be described purely in terms of transitions between particular terminal symbols (words or calls or whatever) are SL languages. (It is not clear why Fitch and Hauser mentioned the orthogonal issue of transition probability.)

The SL class is the infinite union of a hierarchy of SL_k languages ($k > 1$), the k setting the maximal distance over which dependencies can be stated. The most basic SL languages are the SL_2 languages, describable by finite sets of *bigrams*. The SL_2 class is right at the bottom of several infinite hierarchies within the FS languages (see Pullum & Rogers [2006] or Rogers & Pullum [2007] for the mathematics).

Fitch and Hauser found that cotton-top tamarins could be habituated to an SL_2 pattern, namely the one denoted by (ab)*. They remark that perhaps tamarins fail to learn non-FS PSG languages "because their ability to differentiate successive items is limited to runs of two" (Fitch & Hauser 2004, p. 379), conceding the point

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

that a limitation to recognizing bigrams might well be involved. Their results do not in any way imply that monkeys can acquire arbitrary FS languages from exposure to primary data. They may not be much better at pattern learning than your cat. E&L have unfortunately contributed to the spread of a myth.

As for the supposed myth of E&L's title, that of language universals, we see little prospect of sensible debate at this stage. People trying to set one up usually depict a clash between Chomsky, who has purportedly "shown that there is really only one human language" (Smith 1999 p. 1), and Joos, who is alleged to have claimed that languages may "differ from each other without limit and in unpredictable ways" (e.g., Smith 1999, p. 105).

But Chomsky (in Kasher 1991, p. 26) says merely that if all parameters of syntactic variation are "reducible to lexical properties" and if we ignore all parameters that are so reducible (hence we ignore all parameters), there is no syntactic variation at all, so the number of distinct syntactic systems is 1. This is not an empirical claim about human languages; it is a tautology. And Joos (1966, p. 96), while setting a phonology paper in historical context, merely alluded to an "American (Boas) tradition" that valued cataloguing language features over explanatory speculation. The passage quoted does not endorse that tradition or extend it to syntax.

It should be obvious that we must assume languages may differ in unpredictable ways: we do not know the limits of variation, so fieldwork often brings surprises. That was Boas's point. But equally obviously, not all conceivable differences between languages will be attested. Logically there could be dekatransitive verbs (taking ten obligatory object NPs), but there are not, because using them would outstrip our cognitive resources. In that sense there will be all sorts of limits.

This does not look like the seeds of an interesting debate, so it is just as well that E&L do not really try to pursue one. Their conclusions are not about universally quantified linguistic generalizations being mythical, but about how "the diversity of language is, from a biological point of view, its most remarkable property" (sect. 8, para. 6, their thesis 1). That is an interesting thought, and it deserves extended consideration by linguistic and cognitive scientists.

ACKNOWLEDGMENT

We thank Rob Truswell for helpful critical comments on an earlier draft. He bears no responsibility for the opinions expressed in this one, or for any errors.

The discovery of language invariance and variation, and its relevance for the cognitive sciences

doi:10.1017/S0140525X09990574

Luigi Rizzi

CISCL (Interdepartmental Centre for Cognitive Studies on Language), University of Siena, 53100 Siena, Italy.

rizzil unisi.it

Abstract: Modern linguistics has highlighted the fundamental invariance of human language: A rich invariant structure has emerged from comparative studies nourished by sophisticated formal models; languages also differ along important dimensions, but variation is constrained in severe and systematic ways. I illustrate this research direction in the domains of island constraints, word order restrictions, and the expression of referential dependencies. Both language invariance and language variability within systematic limits are highly relevant for the cognitive sciences.

Evans & Levinson (E&L) claim that diversity, rather than uniformity, is the fundamental fact of language relevant for cognitive scientists. This is not the conclusion that modern linguistics arrived at over the last half century: careful testing of specific universalist hypotheses permitted in many cases the discovery of precise and stable patterns of invariance.

First, consider locality. In the late 1970s, the observation that *wh*-phrases can be extracted from indirect questions in Italian led to abandoning the hypothesis of a universal Wh Island Constraint (WhIC), and to the formulation of a parametrized version of the relevant Universal Grammar (UG) principle, Subjacency (Rizzi 1978). The attempt to systematically check the empirical validity of the model led to the discovery of argument/adjunct asymmetries. *Wh*-adjuncts (*why, how, ...*) are strictly non-extractable even in languages which are permissive with argument extraction:

(1) a. Quale problema non sai come risolvere?
"Which problem don't you know how to solve?"

b. *Come non sai quale problema risolvere?
"How don't you know which problem to solve?"

The critical cross-linguistic observation came from the study of *wh-in-situ* languages (with *wh*-elements remaining in clause-internal position). Huang(1982) discovered that in Chinese, the argument-adjunct asymmetry (1) holds at the interpretive level. Sentence (2) can mean (3)a, with interpretive extraction of the argument, but not (3)b, with interpretive extraction of the adjunct:

(2) Akiu xiang.zhidao [women weishenme jiegu-le shei] ne?
Akiu want.know [we why fire-Prf who] Qwh

(3) a. "Who is the person x such that Akiu wonders [why we fired person x]?"

b. "*What is the reason x such that Akiu wonders [whom we fired for reason x]?"

Huang's conclusion was that in *wh-in-situ* languages the *wh*-elements move abstractly to the front, and the abstract movement in (2) is sensitive to the same fundamental locality constraint as the overt movement in (1) (see also, Tsai 1994). A huge literature was then devoted to understanding the nature of the asymmetry in (1) to (3) (see Cinque 1990; Pesetsky 2000; Starke 2001, among others), the generality of adjunct non-extractability (Lasnik & Saito 1992), and the parametrization involved in argument extraction. So, the refutation of an unqualified WhIC gave rise to a deeper cross-linguistic generalization, which was soon extended to a much larger class of Weak Islands (Szabolcsi 2006), and deeply influenced the development of the theory of locality (Chomsky 1995, 2000; Rizzi 1990).

Here and elsewhere, variation is not haphazard but gives rise to precise cross-linguistic patterns, with some logical possibilities systematically unattested: we don't find the "mirror image" of Italian or Chinese, that is, of a language allowing (overt or covert) extraction of adjuncts and barring extraction of arguments. If the non-universality of the WhIC had been interpreted as showing that anything goes in extraction across languages, linguists would have continued to collect catalogues of extraction facts for individual languages, with little hope of ever uncovering systematic generalizations in the patterns of variation.

In this case, the discovery of invariance required considerable abstractness and theoretical sophistication. Much of the typological tradition discussed by E&L has chosen to stick to an extremely impoverished, non-abstract descriptive apparatus, a legitimate methodological decision, but one which severely restricts the possibility of capturing forms of invariance. In spite of these limitations, it is far from true that the typological literature globally supports an "anything goes" position. In fact, the typological findings can be an important source of inspiration for more structured and abstract approaches. Consider Cinque's (2005) discussion of Greenberg's (1963a) Universal 20, a generalization essentially confirmed by 40 years of typological

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

investigations: in languages in which N appears at the end of the NP, the order of certain modifiers is rigid (Demonstrative – Numeral – Adjective – Noun); in languages in which N is initial, significant ordering variation is observed. Cinque offers an explanation of this pattern through the assumption that N movement is the fundamental engine determining reordering. If N moves, it can determine different orders on the basis of familiar parameters on movement. If N remains in the final position, the order is invariably the basic one, plausibly determined by compositional semantics. Again, only a small subset of the logical possibilities is attested, and precise patterns of variation hold, which are amenable to plausible computational principles.

Finally, consider the theory of binding, constraining referential dependencies (E&L, sect. 3, para. 8). There is some limited variation in the locality conditions for the principles holding for anaphors and pronominals, expressible through restrictive and highly falsifiable parametric models (Manzini & Wexler 1986). Clearly, we are very far from a picture of indefinite variability. Moreover, there is no variation at all in non-coreference effects: no language seems to allow coreference between a pronoun and a NP when the pronoun c-commands the NP (*He_i *said that* $John_i$ *was sick*: Lasnik 1989), a property which children are sensitive to as early as testing is feasible (Crain 1991). Furthermore, referential dependencies are systematically regulated by the hierarchical relation c-command (Reinhart 1976), not by other imaginable relations. In sum, we find highly restricted variation, or strict uniformity. The same kinds of considerations would apply to many other domains discussed by E&L (constituent structure, the status of subjects, etc.).

I certainly agree that language variation is of great interest for the cognitive sciences. Much promising research already focuses on that. Brain-imaging studies try to determine how the brain learns formal rules falling within or outside the range of attested cross-linguistic variation (Moro 2008; Tettamanti et al. 2002); acquisition studies uncover systematic "errors" in language development which correspond to structures possible in other adult languages, thus suggesting that children explore UG-constrained grammatical options which are not supported by experience (Rizzi 2006; Thornton 2008); and so on. The cognitive sciences should not overlook the results of half a century of linguistic research which has seriously attempted to identify the limits of variation: it is simply not true that languages can vary indefinitely. This fundamental fact is of critical importance for addressing central questions of acquisition and processing, questions on the expression of the linguistic capacities in the brain, and on the domain specificity or generality of the relevant cognitive constraints. The astounding observational diversity of living organisms has not deflected biologists from the search for the fundamental common features of life. There is no reason why linguists and cognitive scientists should give up the theory-guided search for deep invariants and the ultimate factors of linguistic variation.

Universals in cognitive theories of language

doi:10.1017/S0140525X09990586

Paul Smolensky[a] and Emmanuel Dupoux[b]

[a]*Department of Cognitive Science, Johns Hopkins University, Baltimore, MD 21218-2685;* [b]*Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, Département d'Etudes Cognitives, Ecole Normale Supérieure, Centre National de la Recherche Scientifique, 75005 Paris, France.*

smolensky@jhu.edu

http://www.cogsci.jhu.edu/faculty/smolensky.html

emmanuel.dupoux@gmail.com

http://www.lscp.net/persons/dupoux.html

Abstract: Generative linguistics' search for linguistic universals (1) is not comparable to the vague explanatory suggestions of the article; (2) clearly merits a more central place than linguistic typology in cognitive science; (3) is fundamentally untouched by the article's empirical arguments; (4) best explains the important facts of linguistic diversity; and (5) illuminates the dominant component of language's "biocultural" nature: biology.

1. A science of cognition needs falsifiable theories. Although the target article's final seven theses (sect. 8) include suggestions we find promising, they are presented as vague speculation, rather than as a formal theory that makes falsifiable predictions. It is thus nonsensical to construe them as superior to a falsifiable theory on the grounds that that theory has been falsified. Every theory is certain to make some predictions that are empirically inadequate, but the appropriate response within a science of cognition is to improve the theory and not to take refuge in the safety of unfalsifiable speculation. Insightful speculation *is* vital – not because speculation can *replace* formal theorizing, but because speculation can be sharpened to *become* formal theory. Theory and speculation are simply not empirically comparable.

2. In a theory of cognition, a universal principle is a property true of all human minds – a cog-universal – not a superficial descriptive property true of the expressions of all languages – a des-universal. This is why generative grammar, with its explicit goal of seeking cog-universals, has always been more central to cognitive science than linguistic typology, which only speaks to des-universals. Unlike descriptive linguistic typology, generative grammar merits a central place in cognitive science because its topic is cognition and its method is science – falsifiable theory formulation.

3a. Counterexamples to des-universals are not counterexamples to cog-universals. The des-universals of Evans & Levinson's (E&L's) Box 1 must not be confused with the cog-universals sought by generative grammar. This general point applies to all cases addressed in their article, but we only illustrate with one example. That Chinese questions do not locate *wh*-expressions in a different superficial position than the corresponding declarative sentence (Box 1) is a counterexample to a *wh*-movement *des*-universal but, famously, generative syntax has revealed that *Chinese behaves like English* with respect to syntactically determined restrictions on possible interpretations of questions; this follows if questions in both languages involve the same dependency between the same two syntactic positions, one of them "fronted." In English, the fronted position is occupied by the *wh*-phrase and the other is empty, whereas in Chinese the reverse holds (Huang 1982/1998; Legendre et al. 1998). It is the syntactic relation between these positions, not the superficial location of the *wh*-phrase, that restricts possible interpretations. Such a hypothesized *cog*-universal can only be falsified by engaging the full apparatus of the formal theory – it establishes nothing to point to the superficial fact that *wh*-expressions in Chinese are not fronted.

3b. There are two types of cog-universals: Architectural and specific universals. The former specify the computational architecture of language: levels of representation (phonological, syntactic, semantic, etc.) data structures (features, hierarchical trees, indexes, etc.), operations (rule application, constraint satisfaction, etc.). The authors correctly recognize these as "design features" of human languages, but erroneously exclude them from the set of relevant universals. These architectural universals do not yield falsifiable predictions regarding typology, but they yield falsifiable predictions regarding language learnability. For instance, Peperkamp et al (2006) showed that without architectural universals regarding phonological rules, general-purpose unsupervised learning algorithms simply fail to acquire the phonemes of a language. The latter, specific universals, are tied to particular formal theories specifying in detail the architecture's levels, structures, and operations, thus yielding falsifiable predictions regarding language typology.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

4a. Optimality Theory (OT), mentioned in the target article as a promising direction, contains the strongest architectural and specific universals currently available within generative grammar. According to OT's architectural universals (Prince & Smolensky 1993/2004; 1997), grammatical computation is optimization over a set of ranked constraints. This strong hypothesis (more than the hypothesis of "parameters") has contributed insight into all levels of grammatical structure from phonology to pragmatics, and has addressed acquisition, processing, and probabilistic variation (the website http://roa.rutgers.edu hosts more than 1,000 OT papers). In a particular OT *theory*, specific universals take the form of a set of constraints (e.g., C1 = "a sentence requires a subject"; C2 = "each word must have an interpretation," etc.) A *grammar* for a particular language is then a priority ranking of these constraints. For instance, C1 is ranked higher than C2 in the English grammar, so we say "it is raining," although expletive "it" contributes nothing to the meaning; in Italian, the reverse priority relation holds, making the subjectless sentence "piove" *optimal* – grammatical (Grimshaw & Samek-Lodovici 1998).

4b. OT's cog-universals yield theories of cross-linguistic typology that generally predict the absence of des-universals. Each ranking of a constraint set mechanically predicts the possible existence of a human language. OT therefore provides theories of linguistic typology that aim, as rightly urged by the target article, to grapple with the full spectrum of cross-linguistic variation. OT makes use of a large set of specific universals (i.e., constraints), but because of the resolution of constraint conflict through optimization, these do not translate into des-universals: In the preceding example, C1 is violated in Italian, and C2 in English. Some des-universals can, however, emerge as general properties of the entire typology, and can be falsified by the data (as, perhaps, the existence of onsetless languages). This does not entail abandoning the Generative Linguistics program, nor the OT framework, but rather, revising the theory with an improved set of specific universals.

5. Language is more a biological trait than a cultural construct. The authors do not provide criteria to determine where language is located on the continuum of bio-cultural hybrids. Lenneberg, quoted in the target article, presented four criteria for distinguishing biological traits from cultural phenomena (universality across the species, across time, absence of learning of the trait, rigid developmental schedule) and concluded that oral (but not written) language is a biological trait (Lenneberg 1964). The validity of this argument is ignored by the authors. Ironically, OT is more readily connected to biology than to culture: the archictural-universals of OT are emergent symbolic-level effects of subsymbolic optimization over "soft" constraints in neural networks (Smolensky & Legendre 2006); and Soderstrom et al. (2006) have derived an explicit abstract genome that encodes the growth of neural networks containing connections implementing universal constraints.

If language is a jungle, why are we all cultivating the same plot?

doi:10.1017/S0140525X09990598

Maggie Tallerman

Centre for Research in Linguistics and Language Sciences (CRiLLS), Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom.

maggie.tallerman@ncl.ac.uk

http://www.ncl.ac.uk/elll/staff/profile/maggie.tallerman

Abstract: Evans & Levinson (E&L) focus on differences between languages at a superficial level, rather than examining common processes. Their emphasis on trivial details conceals uniform design features and universally shared strategies. Lexical category distinctions between nouns and verbs are probably universal. Non-local dependencies are a general property of languages, not merely non-configurational languages. Even the latter class exhibits constituency.

Languages exhibit hugely more diverse phenomena than are displayed in well-studied European families. However, citing a collection of exotica does not prove Evans & Levinson's (E&L's) claim that "it's a jungle out there" (sect. 3, para. 17). Examining languages more closely, or at a higher level of abstraction, often reveals critical similarities which superficial descriptions can obscure. Moreover, languages frequently employ distinct grammatical strategies to achieve parallel outcomes; thus, the universal is the end result, not the means of achieving it. Finally, unrelated languages often "choose" the same strategy, despite the lack of a single universal solution, suggesting that homogeneity is widespread.

Lexical category distinctions (sect. 2.2.4). Certainly, there is no invariant set of lexical or functional categories. But it remains to be demonstrated that a language may lack any distinctions between lexical categories, or, more specifically, may lack a noun/verb distinction. E&L note that languages of the Pacific Northwest Coast are frequently claimed to have no noun/verb distinction, illustrating with Straits Salish. Similar claims have been made for a nearby, unrelated family, Southern Wakashan (e.g., Makah, Nuuchahnulth). Here, nouns can function as predicates (i.e., not only arguments) and bear predicative inflections, including tense, aspectual, and person/number marking, and verbs can function as arguments (i.e., not only predicates) and bear nominal inflections, including determiners; (1) and (2) give Nuuchahnulth examples from Swadesh (1939):

1. mamuuk-maa quu?as-?i
   work-3s:INDIC man-the
   "The man is working."

2. quu?as-maa mamuuk-?i
   man-3s:INDIC work-the
   "The working one is a man."

Thus, nominal and verbal roots cannot be identified either by distribution or morphology. Additionally, essentially any lexical root in Nuuchahnulth, including (the equivalents of) nouns, adjectives, and quantifiers, can take verbal inflectional morphology, superficially suggesting that all words are predicative, and thus that there is no noun/verb distinction. Immediate evidence against this (Braithwaite 2008) is that verbs only function as arguments when a determiner is present, whereas nouns function as arguments even without a determiner.

Close inspection reveals further behavioral differences between noun and verb roots (Braithwaite 2008). For instance, proper names can take nominal inflections, such as the definite *-?i*, shown on noun and verb stems in (1) and (2), but cannot take the third singular indicative verbal inflection *-maa*:

3. *Jack-maa
   Jack-3s:INDIC
   ("He is Jack.")

Names, a subclass of nouns, therefore cannot be predicates, clearly distinguishing them from verb roots.

Moreover, although both nominal and verbal predicates can bear possessive markers, nominal predicates with possessive morphemes display a systematic ambiguity in terms of which argument an accompanying person marker is understood to refer to, whereas verbal predicates display no such ambiguity. A similar ambiguity arises in tense marking. Verbal predicates in Nuuchahnulth display a past tense suffix: *-(m)it*:

4. mamuuk-(m)it -(m)aɬ
   work-PAST-1s.INDIC
   "I was working."

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

This suffix also appears on nouns. Even nonpredicative nouns, including names, can bear tense morphology, apparently supporting the lack of a noun/verb distinction:

5. ʔaħʔaaʔaɬ qaħšiɬ–'a ɬ mista-(m)it
and.then die-EVENTIVE Mista-PAST
"Then (the late) Mista died."

The past-tense marker *-(m)it* on the name conveys the specific meaning "former"; since names cannot be predicative in Nuuchahnulth, as (3) shows, this is evidently not a nominal predicate. However, past-tense markers also attach to nominal predicates, which are then interpreted in one of two ways: (6) shows a past-tense nominal predicate, exactly parallel to (4), except with a noun root; (7) displays a predicate nominal in which *-(m)it* bears the alternative "former" meaning:

6. quuʔas-(m)it-(m)aħ
person-PAST-1s.INDIC
"I was a man."

7. ʔuunuuɬ ʔani ʔuumiik-(m)it-qa
because that whaler-PAST-SUBORDINATE
"because he was a former whaler"

Critically, *-(m)it* on a verbal predicate never exhibits the "former" meaning but is always interpreted simply as past tense. In sum, careful investigation such as that of Braithwaite provides ample evidence for a noun/verb distinction in Wakashan languages, despite superficial appearances.

Constituent structure (sect. 5). As E&L note, "non-configurational" languages display free word order and discontinuous constituents: in (8), from the Australian language Kalkatungu, the underscore shows the components of the ergative subject, and italics show the (nominative) object:

8. Tjipa-yi *tjaa* kunka-(ng)ku *pukutjurrka* lhayi nguyi-nyin-tu.
this-ERG this branch-ERG mouse kill fall-PARTICIPLE-ERG
"The falling branch hit the mouse." (Blake 2001, p. 419)

E&L state that "the parsing system for English cannot be remotely like the one for such a language" (sect. 2, para. 3), because case-tagging indicates relationships between words, rather than constituency and fixed word order. But, in fact, the parsing system for English is well used to non-local dependencies – that is, to relating items not contiguous in the string. Note the discontinuous constituents in the following examples, and that the dependency even occurs across a clause boundary in the second instance: *A student sauntered in wearing a large fedora*; *Which girl did you say he gave the books to* __?. Parsing in Kalkatungu (or Latin) therefore utilizes a strategy also found in languages which do have clear constituents. Moreover, completely unrelated non-configurational languages like Kalkatungu and Latin share the same method of signaling relationships between words (case-marking). All this is hardly indicative of the jungle E&L assume; rather, it is evidence that very few solutions are available, and that languages make differential use of options from a small pool of possibilities.

Furthermore, certain non-configurational Australian languages (e.g., Wambaya; Nordlinger 2006) actually have one strict word order requirement, namely that the auxiliary is in second position, thus either second word, (9), or second constituent, (10) (Hale 1973 outlines the parallel requirement in Warlpiri):

9. Nganki ngiy-a lurrgbanyi wardangarringa-ni alaji
this.ERG 3SF-PAST grab moon-ERG boy
"The moon grabbed (her) child."

10. Naniyawulu nagawulu baraj-bulu wurlu-n duwa.
that.DUAL.NOM female.DUAL.NOM old.person-DUAL.(NOM) 3.DUAL-PROG get.up
"Those two old women are getting up."

Crucially, the auxiliary cannot appear as, say, third word within a four-word noun phrase. Contra E&L, this demonstrates the psychological reality of word order and of constituent structure in such languages. Moreover, while by no means universal, second-position phenomena occur widely (e.g., Sanskrit, Celtic, Germanic), demonstrating remarkable formal homogeneity cross-linguistically.

Finally, E&L claim linguistic diversity is not characterized by "selection from a finite set of types" (sect. 8, para 9, their thesis 3). Case-encoding systems are few indeed, and familiar strategies (such as ergativity) even occur in language isolates such as Basque.

Universal grammar is dead

doi:10.1017/S0140525X09990744

Michael Tomasello
Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.
tomas@eva.mpg.de

Abstract: The idea of a biologically evolved, universal grammar with linguistic content is a myth, perpetuated by three spurious explanatory strategies of generative linguists. To make progress in understanding human linguistic competence, cognitive scientists must abandon the idea of an innate universal grammar and instead try to build theories that explain both linguistic universals and diversity and how they emerge.

Universal grammar is, and has been for some time, a completely empty concept. Ask yourself: what exactly is in universal grammar? Oh, you don't know – but you are sure that the experts (generative linguists) do. Wrong; they don't. And not only that, they have no method for finding out. If there is a method, it would be looking carefully at all the world's thousands of languages to discern universals. But that is what linguistic typologists have been doing for the past several decades, and, as Evans & Levinson (E&L) report, they find no universal grammar.

I am told that a number of supporters of universal grammar will be writing commentaries on this article. Though I have not seen them, here is what is certain. You will not be seeing arguments of the following type: I have systematically looked at a well-chosen sample of the world's languages, and I have discerned the following universals … And you will not even be seeing specific hypotheses about what we might find in universal grammar if we followed such a procedure. What you will be seeing are in-principle arguments about why there have to be constraints, how there is a poverty of the stimulus, and other arguments that are basically continuations of Chomsky's original attack on behaviorism; to wit, that the mind is not a blank slate and language learning is not rat-like conditioning. Granted, behaviorism cannot account for language. But modern cognitive scientists do not assume that the mind is a blank slate, and they work with much more powerful, cognitively based forms of learning such as categorization, analogy, statistical learning, and intention-reading. The in-principle arguments against the sufficiency of "learning" to account for language acquisition (without a universal grammar) assume a long-gone theoretical adversary.

Given all of the data that E&L cite, how could anyone maintain the notion of a universal grammar with linguistic content? Traditionally, there have been three basic strategies. First, just as we may force English grammar into the Procrustean bed of Latin grammar – that is how I was taught the structure of English in grade school – the grammars of the world's so-called exotic

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

languages may be forced into an abstract scheme based mainly on European languages. For example, one can say that all the world's languages have "subject." But actually there are about 30 different grammatical features that have been used with this concept, and any one language has only a subset – often with almost non-overlapping subsets between languages. Or take noun phrase. Yes, all languages may be used to make reference to things in the world. But some languages have a large repertoire of specially dedicated words (nouns) that play the central role in this function, whereas others do not: they mostly have a stock of all-purpose words which can be used for this, as well as other, functions. So are subjects and noun phrases universal? As you please.

Second, from the beginning a central role in universal grammar has been played by the notion of transformations, or "movement." A paradigm phenomenon in English and many European languages is so-called wh- movement, in which the wh- word in questions always comes at the beginning no matter which element is being questioned. Thus, we ask, "What did John eat?", which "moves" the thing eaten to the beginning of the sentence (from the end of the sentence in the statement "John ate X"). But in many of the world's languages, questions are formed by substituting the wh- word for the element being questioned in situ, with no "movement" at all, as in "John ate what?". In classic generative grammar analyses, it is posited that all languages have wh- movement, it is just that one cannot always see it on the surface – there is *underlying* movement. But the evidence for this is, to say the least, indirect.

The third, more recent, strategy has been to say that not all languages must have all features of universal grammar. Thus, E&L note that some languages do not seem to have any recursive structures, and recursion has also been posited as a central aspect of universal grammar (in a very different way than such notions as noun phrase). The response has been that, first of all, these languages do have recursive structures, it is just that one cannot see them on the surface. But even if they do not have such structures, that is fine because the components of universal grammar do not all apply universally. This strategy is the most effective because it basically immunizes the Universal Grammar (UG) hypothesis from falsification.

For sure, all of the world's languages have things in common, and E&L document a number of them. But these commonalities come not from any universal grammar, but rather from universal aspects of human cognition, social interaction, and information processing – most of which were in existence in humans before anything like modern languages arose. Thus, in one account (Tomasello 2003a; 2008), human linguistic universals derive from the fact that all humans everywhere: (1) conceive nonlinguistically of agents of actions, patients of actions, possessors, locations, and so forth; (2) read the intentions of others, including communicative intentions; (3) follow into, direct, and share attention with others; (4) imitatively learn things from others, using categorization, analogy, and statistical learning to extract hierarchically structured patterns of language use; and (5) process vocal-auditory information in specific ways. The evolution of human capacities for linguistic communication draw on what was already there cognitively and socially ahead of time, and this is what provides the many and varied "constraints" on human languages; that is, this is what constrains the way speech communities grammaticalize linguistic constructions historically (what E&L call "stable engineering solutions satisfying multiple design constraints"; target article, Abstract, para. 2).

Why don't we just call this universal grammar? The reason is because historically, universal grammar referred to specific linguistic content, not general cognitive principles, and so it would be a misuse of the term. It is not the idea of universals of language that is dead, but rather, it is the idea that there is a biological adaptation with specific linguistic content that is dead.

The neglected universals: Learnability constraints and discourse cues

doi:10.1017/S0140525X09990756

Heidi Waterfall[a] and Shimon Edelman[b]

[a]*Department of Psychology, Cornell University, Ithaca, NY 14853, and Department of Psychology, University of Chicago, Chicago, IL 60637;* [b]*Department of Psychology, Cornell University, Ithaca, NY 14853, and Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, South Korea.*

heidi.waterfall@gmail.com
http://kybele.psych.cornell.edu/~heidi
se37@cornell.edu
http://kybele.psych.cornell.edu/~edelman

Abstract: Converging findings from English, Mandarin, and other languages suggest that observed "universals" may be algorithmic. First, computational principles behind recently developed algorithms that acquire productive constructions from raw texts or transcribed child-directed speech impose family resemblance on learnable languages. Second, child-directed speech is particularly rich in statistical (and social) cues that facilitate learning of certain types of structures.

Having surveyed a wide range of posited universals and found them wanting, Evans & Levinson (E&L) propose instead that the "common patterns" observed in the organization of human languages are due to cognitive constraints and cultural factors. We offer empirical evidence in support of both these ideas. (See Fig. 1.)

One kind of common pattern is readily apparent in the six examples of child-directed speech in Figure 1, in each of which partial matches between successive utterances serve to highlight the structural regularities of the underlying language. Two universal principles facilitating the identification of such regularities can be traced to the work of Zellig Harris (1946; 1991). First, the discovery of language structure, from morphemes to phrases, can proceed by cross-utterance alignment and comparison (Edelman & Waterfall 2007; Harris 1946). Second, the fundamental task in describing a language is to state the departures from equiprobability in its sound- and word-sequences (Harris 1991, p. 32; cf. Goldsmith 2007).

These principles are precisely those used by the only two unsupervised algorithms currently capable of learning productive construction grammars from large-scale raw corpus data, ADIOS (Solan et al. 2005) and ConText (Waterfall et al., under review). Both algorithms bootstrap from completely unsegmented text to words and to phrase structure by recursively identifying candidate constructions in patterns of partial alignment between utterances in the training corpus. Furthermore, in both algorithms, candidate structures must pass a statistical significance test before they join the growing grammar and the learning resumes (the algorithms differ in the way they represent corpus data and in the kinds of significance tests they impose).

These algorithms exhibited hitherto unrivaled – albeit still very far from perfect – capacity for language learning, as measured by (1) precision, or acceptability of novel generated utterances, (2) recall, or coverage of withheld test corpus, (3) perplexity, or average uncertainty about the next lexical element in test utterances, and (4) performance in certain comprehension-related tasks (Edelman & Solan, under review; Edelman et al. 2005; 2004; Solan et al. 2005). They have been tested, to varying extents, in English, French, Hebrew, Mandarin, Spanish, and a few other languages. The learning algorithms proved particularly effective when applied to raw, transcribed, child-directed speech (MacWhinney 2000), achieving precision of 54% and 63% in Mandarin and English, respectively, and recall of about 30% in both languages (Brodsky et al. 2007; Solan et al. 2003).

To the extent that human learners rely on the same principles of aligning and comparing potentially relatable utterances, one may put these principles forward as the source of part of

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

English:
those are checkers
two checkers yes
play checkers

Italian:
dove sono
dove sono i coniglietti

Hebrew:
מה לא
מה את לא רוצה
את רוצה לספר

Korean:
제일 이뻐
누가 제일 이뻐
지원이 제일 이뻐 맞어

Mandarin:
这 是 什么 呀
哎呀 是 什么 呀

Russian:
вот твой папа не хочет с тобой остаться
как не хочет хочет
хочет папа хочет

Figure 1 (Waterfall & Edelman). **Examples of child-directed speech in six languages.** It is not necessary to be able to read, let alone understand, any of these languages to identify the most prominent structural characteristics common to these examples (see text for a hint). These characteristics should, therefore, be readily apparent to a prelinguistic infant, which is indeed the case, as the evidence we mention suggests. All the examples are from CHILDES corpora (MacWhinney 2000).

speech, phrase structure, and other structural "universals." In other words, certain forms may be common across languages because they are easier to learn, given the algorithmic constraints on the learner.[1]

Language acquisition becomes easier not only when linguistic forms match the algorithmic capabilities of the learner, but also when the learner's social environment is structured in various helpful ways. One possibility here is for mature speakers to embed structural cues in child-directed speech (CDS). Indeed, a growing body of evidence suggests that language acquisition is made easier than it would have been otherwise because of the way CDS is shaped by caregivers during their interaction with children.[2] One seemingly universal property of CDS is the prevalence of variation sets (Hoff-Ginsberg 1990; Küntay & Slobin 1996; Waterfall 2006; under review) – partial alignment among phrases uttered in temporal proximity, of the kind illustrated in Figure 1. The proportion of CDS utterances contained in variation sets is surprisingly constant across languages: 22% in Mandarin, 20% in Turkish, and 25% in English (when variation sets are defined by requiring consecutive caregiver utterances to have in common at least two lexical items in the same order; cf. Küntay & Slobin 1996; this proportion grows to about 50% if a gap of two utterances is allowed between the partially matching ones). Furthermore, the lexical items (types) on which CDS utterances are aligned constitute a significant proportion of the corpus vocabulary, ranging from 9% in Mandarin to 32% in English.

Crucially, the nouns and verbs in variation sets in CDS were shown to be related to children's verb and noun use at the same observation, as well as to their production of verbs, pronouns, and subcategorization frames four months later (Hoff-Ginsberg, 1990; Waterfall 2006; under review). Moreover, experiments involving artificial language learning highlighted the causal role of variation sets: adults exposed to input which contained variation sets performed better in word segmentation and phrase boundary judgment tasks than a control group that heard the same utterances in a scrambled order, which had no variation sets (Onnis et al. 2008).

The convergence of the three lines of evidence mentioned – the ubiquity of variation sets in child-directed speech in widely different languages, their proven effectiveness in facilitating acquisition, and the algorithmic revival of the principles of acquisition intuited by Harris – supports E&L's proposal of the origin of observed universals. More research is needed to integrate the computational framework outlined here with models of social interaction during acquisition and with neurobiological constraints on learning that undoubtedly contribute to the emergence of cognitive/cultural language universals.

NOTES

1. Language may also be expected to evolve in the direction of a better fit between its structure and the learners' abilities (Christiansen & Chater 2008).

2. Social cues complement and reinforce structural ones in this context (Goldstein & Schwade 2008).

Authors' Response

With diversity in mind: Freeing the language sciences from Universal Grammar

doi:10.1017/S0140525X09990525

Nicholas Evans[a] and Stephen C. Levinson[b]

[a]*Department of Linguistics, Research School of Asian and Pacific Studies, Australian National University, ACT 0200, Australia;* [b]*Max Planck Institute for Psycholinguistics, Wundtlaan 1, NL-6525 XD Nijmegen, The Netherlands; and Radboud University, The Netherlands.*

nicholas.evans@anu.edu.au
http://rspas.anu.edu.au/people/personal/evann_ling.php
stephen.levinson@mpi.nl
http://www.mpi.nl/Members/StephenLevinson

Abstract: Our response takes advantage of the wide-ranging commentary to clarify some aspects of our original proposal and augment others. We argue against the generative critics of our coevolutionary program for the language sciences, defend the use of close-to-surface models as minimizing cross-linguistic data distortion, and stress the growing role of stochastic simulations in making generalized historical accounts testable. These methods lead the search for general principles away from idealized representations and towards selective processes. Putting cultural evolution central in understanding language diversity makes learning fundamental in the cognition of language: increasingly powerful models of general learning, paired with channelled caregiver input, seem set to manage language acquisition without recourse to any innate "universal grammar." Understanding why human language has no clear parallels in the animal world requires a cross-species perspective: crucial ingredients are vocal learning (for which there are clear non-primate parallels) and an intention-attributing cognitive infrastructure that provides a universal base for language evolution. We conclude by situating linguistic diversity within a broader trend towards understanding human

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

cognition through the study of variation in, for example, human genetics, neurocognition, and psycholinguistic processing.

R1. Introduction

The purpose of our target article was to draw attention to linguistic diversity and its implications for theories of human cognition: Structural diversity at every level is not consonant with a theory of fixed innate language structure, but instead suggests remarkable cognitive plasticity and powerful learning mechanisms. We pointed out that human communication is the *only* animal communication system that varies in myriad ways in both form and meaning across the species, and this must be a central fact that should never be lost sight of.

The responses in the commentaries show that opinion in the language sciences, and especially in linguistics, is still sharply divided on "the myth of language universals," or at least our telling of it. The comments of the typological and functional linguists (**Croft**, **Goldberg**, **Haspelmath**) show that much of our argument is already widely accepted there: "evolutionary linguistics is already here" (Croft). Positive responses from many commentators in experimental and cross-species comparative psychology suggest that researchers in experimental psychology and cross-species studies of communication are ready for the kind of coevolutionary, variability-centred approach we outlined (**Bavin**, **Catania**, **McMurray & Wasserman**, **Merker**, **Tomasello**, and **Waterfall & Edelman**). Generative linguists, by contrast, disagreed sharply with our presentation, laying bare some fundamental differences in how linguistics is conceived as a science.[1]

We have organized the response as follows.

Section R2 responds to the critical comments from the generative camp, suggesting that the assumptions behind many of these responses are misplaced.

Section R3 looks at the question of whether we have overstated the range of diversity by ignoring unpopulated regions of the design space.

Section R4 takes the commentaries from the non-generative linguists and the psychological, animal behavior, and computational learning research communities, which were overwhelmingly positive, and indicates how these might be used to round out, or in places correct, our position.

Section R5 sketches where we think these new developments are heading, and their relationship to what else is happening in the cognitive sciences.

We set aside the specific data questions till an appendix at the end, where we concede two factual mistakes, clarify disputed facts and generalizations, and examine more specific linguistic points that would bog down the main argument – on nearly all of them, we think the criticisms from our generativist colleagues do not stand up to scrutiny.

R2. Incompatible evaluation metrics reflect different paradigms

It was never our intention to engage in mud-slinging with our generative colleagues, but as **Tomasello** has predicted there was a certain inevitability that familiar sibling quarrels would be rerun. Most of the criticisms from the generative camp reflect deep differences between generative and typological/functionalist approaches in their overall assumptions about many issues. Where do we locate causal explanations? Where do we seek the general unifying laws behind surface diversity – in structure or in process? Do we use only discrete mathematical models (favoring regularized representations), or do we bring in continuous and stochastic models as well (favoring representations sticking closer to surface variety)? Should generalizations purport to directly represent mental reality, or are they modelling the summed information of thousands of different coevolutionary products shaped by multiple selective factors? Should we adopt essentializing categorizations (as "formal universals"), or abandon these as misleading and adopt a strategy that measures surface diversity directly so as not to lose data that is useful for evaluating the fit of models?

Generative and typological/functionalist accounts will give different answers to each of these questions, and it is this difference in overall scientific paradigm that accounts for the seemingly irreconcilable conflict between generativist commentators like **Freidin** and **Pesetsky**, who see our proposals as so imprecise as to be unfalsifiable, and psychologists like **Tomasello** and **Margoliash & Nusbaum**, for whom it is the generative approach that has moved away from falsifiability.

To clarify these differences, we try here to give a brief and constructive account of where the real differences lie (as **Pullum & Scholz** opine, more could be fruitless). The generativist critique includes the following interlinked charges:

1. Lack of theory, precise representation, or falsifiability (**Smolensky & Dupoux**, **Freidin**)
2. Mistaken ontology, mistaking behavior for cognition and (a point we hold off till sect. R4.1) history for science (**Smolensky & Dupoux**)
3. Lack of abstractness – that we are misled by surface variation into ignoring underlying structural regularities (**Baker**, **Harbour**)
4. That taking surface diversity at face value leads away from the quest for general principles (**Smolensky & Dupoux**, **Nevins**)
5. That we have neglected the presence of "formal universals" (**Nevins**)
6. That the typologists' preference for using a non-abstract descriptive apparatus is the wrong methodological choice (**Rizzi**)
7. That we have merely presented an under-analyzed *Wunderkammer* of variation that can be shown to reduce to well-known phenomena (**Pesetsky**).

We now take up these issues one at a time. A further criticism, that we may have overstated the range of diversity by ignoring the fact that languages all lie within a bounded corner of the possibility space (**Pinker & Jackendoff**, **Tallerman**) is dealt with separately in section R3.

R.2.1. What kind of theory?

Smolensky & Dupoux and **Freidin** complain that we did not offer a fully articulated theory with precise predictions about sentential structure. But that was not what we set out to do. Our goal was to survey our current understanding of language variation, explain its import for the cognitive sciences, and outline a fertile area for future research. We sketched the kind of biological models into which this

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

variation neatly seems to fit and the ones that invite future development in a number of directions. A lot of these materials and ideas have not been sufficiently taken into account, we felt, by researchers in the cognitive sciences. We were gently suggesting that the time has come for a paradigm change, and at the end of this response we will say a little more.

Nevertheless, at the end of the target article we did sketch directions for future research (see also sect. R5 in this response). Commentators outside the generative camp (e.g., **Waterfall & Edelman**, **Christiansen & Chater**) in many cases saw no difficulty in deriving a range of predictions or consequences, and indeed saw the target article as "mov[ing] the study of language in the direction of the research methods of the experimental sciences and away from enclosed personal belief systems" (**Margoliash & Nusbaum**).

The radically different assessments of empirical grounding here reflect (we think) a narrow view of theory and evidence on the part of some of our critics. Within the language sciences there is a wide variety of theory – theories about language change (historical linguistics), language usage (pragmatics), microvariation within individual languages (sociolinguistics), language production, acquisition and comprehension (psycholinguistics), language variation (typology), and language structure (the traditional heart of linguistics), to name just a few. Generative theory is just one version of a theory of linguistic structure and representation, and it is marked by a lack of external explanatory variables, making no reference to function, use, or psychological or neural implementation. It has delivered important insights into linguistic complexity, but has now run into severely diminishing returns. It is time to look at the larger context and develop theories that are more responsive to "external" constraints, be they anatomical and neural, cognitive, functional, cultural, or historical. Here we think an evolutionary framework has a great deal to offer in the multiple directions we sketched.

We pointed out the central fact that the human communication system is characterized by a diversity in form and meaning that has no parallel in the animal kingdom. Generative theory has never come fully to terms with this, and a theory of universal grammar that isn't answerable to linguistic variation consequently has distinctly limited appeal.

R.2.2. Cognition, behavior, and representation

Various Chomskyan dichotomies (competence vs. performance, i-language vs. e-language, **Smolensky & Dupoux**'s cog-universals vs. des-universals) have been used to drive a wedge between cognition and behavior. There are distinct dangers in this.

First, most cognitive scientists will agree that cognition exists to service perception and behavior. Second, the evidence for cognition remains behavioral and perceptual (even when we can look at the functioning brain in vivo, we look at its response to an event), and most cognitive scientists will want all theories measured ultimately in terms of predictions over brain events and behavior or response (as the very title of this journal suggests; cf. **Margoliash & Nusbaum**). Third, many cognitive scientists view favorably the new "embodiment" perspectives which blur the line between representation and process.

Chomsky, in his initial work on formal grammars, suggested that the descriptive apparatus chosen to model language should be just sufficient and not more powerful than is required – in that way, some match to cognition may be achieved. From then on, in the generative tradition there has been a systematic conflation between the language of description and what is attributed to the language learner and user: "the brains of all speakers represent a shared set of grammatical categories" (**Berent**), and "formal universals in phonology are constituted by the analytic elements that human minds employ in constructing representations of sound structure" (**Nevins**).

Many generativist approaches – particularly parametric formulations – consequently attribute cognitive reality to conditionals of the form "if structural decision X, then also structural decision Y" or "learning X is universally easier than learning Y" (essentially **Nevins**' Example [1]). No language typologist would maintain that conditional regularities of this type would be found in speakers' heads. Yet this is precisely what is advocated in the OT (Optimality Theory) framework advocated by **Smolensky & Dupoux**:

> OT . . . is inherently typological: the grammar of one language inevitably incorporates claims about the grammars of all languages. This joining of the individual and the universal, which OT accomplishes through ranking permutation, is probably the most important insight of the theory. (McCarthy 2002, p. 1)

To make this work, an infinite set of possible sentences are first generated then filtered by (among other things) comparisons of this type. Instead of putting the filtering where it belongs, in cultural system evolution across generations, OT effectively burdens each individual mind with a précis of the functional history of all known human languages, and loads the entire optimization process onto on-line grammatical computation. This is not just cognitively unrealistic – it is computationally intractable (Idsardi 2006).

This conflation of the metalanguage with the object of description is a peculiar trick of the generative tradition. By going down this path, it has opened up a huge gap between theory and the behavioral data that would verify it. The complex representational structures look undermotivated, and covert processes proliferate where alternative models deftly avoid them (see the discussion of Subjacency and covert movement in sect. R6.8).

A biologist does not assume that a snail maintains an internalized representation of the mathematical equations that describe the helical growth of its shell. Even for the internal characterization of a mental faculty, the strategy is odd: computer scientists interested in characterizing the properties of programming languages use a more general auxiliary language to describe them, as in Scott-Strachey denotational semantics. Once explanatory theories hook external factors (e.g., psycholinguistic or evolutionary factors) into the account, this conflation of cognition and metalanguage must be dropped.

Smolensky & Dupoux's aphorism "Generative grammar merits a central place in cognitive science because its topic is cognition and its method is science," then, will not find universal approval: other branches of linguistics are much more in tune with psychological reality as reflected in language acquisition, production, and comprehension. Nor has generative grammar of the

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Chomskyan variety been particularly successful as an explicit theory of linguistic representation. Many other representational formats, such as HPSG and LFG, have had greater uptake in computational circles (see, e.g., Butt et al. 2006, Reuer 2004). LFG, for example, adopts a parallel constraint-based architecture that includes dependency as well as constituency relations. This allows for the direct representation of crucial types of variability discussed in sect. 5 of our target article, while avoiding the need for movement rules or large numbers of empty nodes (see sect. R6.8 for further discussion of how this works for subjacency). These formats, which represent different outgrowths from the same generative roots, show that precise, testable, computationally tractable models of language can be developed that reflect cross-linguistic diversity much more directly in their architecture.

R2.3. Abstractness and universal generalizations

A number of commentators (**Baker**, **Harbour**, **Nevins**, **Pesetsky**) felt that we were unwilling to entertain the sorts of abstract analyses which allow us to find unity in diversity. But we are simply pointing out that the proposals on the table haven't worked. Abstractness has a cost: the more unverifiable unobservables, the greater the explanatory payoff we expect. Judging the point where explanatory superstructure becomes epicyclic and unproductive may be tough, and generative and non-generative camps clearly have different thresholds here. But the increasingly abstruse theoretical apparatus is like a spiralling loan that risks never being paid by the theory's meagre empirical income (cf. Edelman & Christiansen 2003). Even attempts to deal with the growing evidence of variability through the theory of parameters – projecting out diversity by a limited number of "switches" pre-provided in Universal Grammar (UG) – has empirically collapsed (Newmeyer 2004, p. 545), a point largely undisputed by our commentators (although **Rizzi** continues to use the notion – see the discussion of Subjacency in sect. R6.8).

All sciences search for underlying regularities – that's the game, and there is no branch of linguistics (least of all historical linguistics, with its laws of sound change) that is not a player. For this reason **Harbour**'s commentary misses the target – of course some middle level generalizations about the semantics of grammatical number are valid in any framework (although his account of the plural seems to not generalize beyond three participants, and there are additional problems that we discuss in sect. R6.4). The art is to find the highest level generalization that still has empirical "bite."

R2.4. Recognizing structural diversity is not incompatible with seeking general laws

The criticisms by **Nevins, Pesetsky**, and **Smolensky & Dupoux** – that we are not interested in seeking deeper laws behind the surface variation in linguistic structure – reveal a failure to understand the typological/functional approach. In a coevolutionary model the underlying regularities in the cross-linguistic landscape are sought in the vocal-auditory, cognitive, sociolinguistic, functional, and acquisitional selectors which favor the development of some structures over others. The goal is to seek a constrained set of motivated selectors (each testable) that filter what structures can be learned, processed, and transmitted. The stochastic nature of the selection process, and the interaction and competition between multiple selectors, accounts for the characteristic balance we find, of recurrent but not universal patterns with marked diversity in the outliers.

Phonological structures, for example, will be favored to the extent that they can be easily said, easily heard, and easily learned.[2] But these targets regularly conflict, as when streamlined articulation weakens perceptual contrasts or creates formal alternations that are harder to learn. In fact it has been a key insight of optimality theory (OT) that many competing factors need to be juggled, but that not all are equally potent and most can be "non-fatally violated." The different weightings of these "constraints" generate a kaleidoscope of language-specific configurations, and modelling their interaction has been a strong appeal of the OT program. But the constraints identified by OT are more fruitfully treated as the sorts of scalar processing effects sketched in **Haspelmath**'s commentary. The typological sweep, like OT phonology, aims at a comprehensive documentation of all such constraints and their interactions, finding languages in which individual effects can best be isolated or recombined, with laboratory phonology studying why each effect occurs.

The line of attack that "languages share structural similarities often masked by one of their differences" (**Pesetsky**) thus misses the point of why it is useful to confront diversity head on. Like generative theory, the program we have outlined seeks to discover the general behind the particular. But it differs in where we seek the general laws. For our generativist critics, generality is to be found at the level of structural representation; for us, at the level of process. Our claim, in Darwinian mode, is that the unity of evolutionary mechanisms can best be discerned by reckoning with the full diversity of evolutionary products and processes.

R2.5. Non-abstract representations preserve information

Rizzi suggests that the typologist's strategy of using an "extremely impoverished, non-abstract descriptive apparatus" that takes diversity at face value in representing phenomena will have less success than the generative program in establishing universal patterns. Yet, as the burden of explanation for cross-linguistic patterning is moved out of the prewired mind and into the evolution of individual language systems under selection from the sorts of factors outlined earlier, the most appropriate mathematical models employ stochastical and continuous methods rather than the discrete methods that have characterized the generative tradition (Pierrehumbert et al. 2000). And once we employ these methods, there are positive benefits in "directly measuring the variation, instead of reducing it" (Bickel 2009): any other strategy risks degrading the information on which the methods are based.

Take the question of how perceptual discriminability and articulatory ease interact in the domain of vowel dispersion over the formant space to favor the emergence of some vowel systems over others. The classic study by Liljencrants and Lindblom (1972) simulated the evolution

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

of vowel systems over time under these twin selectional pressures and compared the results to the distribution of attested vowel inventories. The insights yielded by their model would not have been possible if the descriptions of vowel systems had been in terms of discrete binary features such as [front] and [round] rather than in terms of position in a continuous three-dimensional space based on formant frequencies.

Staying close to the surface thus avoids the essentializing fallacy critiqued by **Goldberg** and **Croft**, while retaining the maximum information for matching against stochastic models of how general evolutionary processes interact to produce a scatter of different structural outcomes across the language sample.

R2.6. Neglect of "formal universals"

We are criticized by **Nevins** for neglecting "formal universals" – "the analytic elements that human minds employ in constructing representations of sound structure … the available data structures (e.g., binary features, metrical grids, autosegmental tiers) and the possible operations on them that can be used in constructing a grammar of a language." (See also our discussion in sect. R6.8 of Subjacency, as raised by **Smolensky & Dupoux**, **Freidin**, and **Rizzi**.)

Data structures like these have undoubted value in constructing formal representations of phonological phenomena. But, first, it does not follow that they are the actual representations that humans learn and use. As **Tomasello** and **Bavin** argue, increasingly powerful general pattern learning mechanisms suggest that many of the relevant phenomena can be managed without needing the representations that **Nevins** advocates. Second, even if such structures prove to have psychological reality, it does not follow that we are natively endowed with them. Take the general issue of discrete combinatoriality – the fact that languages recombine discrete units like consonants and vowels – which is relevant both to binary features (like ± consonantal) and, in many models, the separation of consonantal and vocalic elements onto distinct autosegmental tiers.[3] Zuidema and De Boer (2009) have used evolutionary game theory simulations to investigate the hypothesis that combinatorial phonology results from optimizing signal systems for perceptual distinctiveness. Selection for acoustic distinctiveness, defined in terms of the probability of confusion, leads along a path of increasing fitness from unstructured, holistic signals to structured signals that can be analyzed as combinatorial. Some very general assumptions – temporal structuring of signals and selection for acoustic distinctiveness – lead over time to the emergence of combinatorial signals from holistic origins.

Should linguists use binary features and autosegmental tiers in the grammars and phonological descriptions they write? Sure, whenever they are useful and elegant. Do we need them to draw on a single, universal feature set to account for the mental representations that speakers have? Probably not, judging by the direction in which the psycholinguistic learning literature is headed. Do we need them to account for why languages all exhibit discrete combinatoriality? No – this can emerge through the sorts of processes that Zuidema and De Boer have modelled. Intriguingly, an empirical parallel has been identified in one new sign language: Meir et al. (in press) and Sandler et al. (2009) show that duality of patterning has only been gradually emerging over three generations of one Bedouin sign language variety.

R2.7. An underanalyzed Wunderkammer of variation

A number of commentators charge us with producing a *Wunderkammer* of exotica (**Pesetsky**), intended more to dazzle rather than illuminate. Pesetsky and **Tallerman** suggest that if properly analyzed these exotica will turn out just to be ordinary, universal-conforming languages. Both take up the issue of constituency, and argue that recent research finds it subtly present in languages previously claimed to lack it. A clarification is in order. There are two potential issues:

a. Is constituency universal in the sense that all languages exhibit it somewhere in their systems, if even marginally?

b. Is constituency universal in the sense that all languages use it as the main organizational principle of sentence structure and the main way of signalling grammatical relations?

Our target was (b) – different languages use different mixes, as has been well-modelled by approaches like LFG; but our commentators tend to target (a).

Pesetsky points out that Tlingit may after all have an initial slot into which constituents can be systematically shifted (we would need to know actually what can go there, and if that is actually predicted by a constituency analysis). But he is wrong in presenting Warlpiri as the "free word order language par excellence." It is well known that Warlpiri places its auxiliary after the first constituent, and that when words are grouped together into a contiguous NP only the last word needs to carry case, instead of the usual patterning of inflecting every word. Neither of these properties, however, are found in Jiwarli (Austin & Bresnan 1996), which is why we chose it as our example.

The point about free word order languages, whether or not they have small islands of constituency, is that they *cannot be parsed by a constituency-based algorithm* as in most NLP (natural language programming) today, because they do not use constituency as the systematic organizing principle of sentence structure. If constituency is not the universal architecture for sentence structure, then the entire generative apparatus of c-command, bounding nodes, subjacency, and so forth collapses, since all are defined in terms of constituency. In this way **Tallerman** is wrong in thinking that parsing free word order is just like parsing English discontinuous constructions – the latter are allowed by rule, which sets up precise expectations of what comes next in what order.

Incidentally, the reader should note the argumentation of these rejoinders: that we, Evans & Levinson (E&L), have cherry-picked exotic facts about language A, but look, language B yields to the normal universal analysis, so there's no reason to take A seriously. Since absolute universals can be falsified by a single counterexample, it is a logical fallacy to defend a universal by adducing facts from some other language which happens not to violate it.

The seven general charges we have discussed capture, we think, most of the sources of disagreement. **Freidin**'s commentary in particular indicates the deep rift in

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

contemporary linguistics between Chomskyans and the rest, which ultimately rests on different judgements about the interlocking of theory and evidence. This is regrettable, as generative grammar has served to open up the "deep unconscious" of language as it were, showing how languages are organized with far greater complexity than had hitherto been imagined. While Chomskyans have presumed that these complexities must be innate, we have argued that there are two blind watchmakers: cultural evolution acting over deep time, and genetic infrastructure, which for the most part, of course, will not be specific to language.

Finally, let us note that Chomsky's own position makes it clear that the generative enterprise simply has a different target than the program we are trying to promote, namely, (in our case) working out the implications of language diversity for theories of cognition and human evolution. The following recent quote makes this clear:

> Complexity, variety, effects of historical accident, and so on, are overwhelmingly restricted to morphology and phonology, the mapping to the sensorimotor interface. That's why these are virtually the only topics investigated in traditional linguistics, or that enter into language teaching. They are idiosyncrasies, so are noticed, and have to be learned. If so, then it appears that language evolved, and is designed, primarily as an instrument of thought. Emergence of unbounded Merge in human evolutionary history provides what has been called a "language of thought," *an internal generative system that constructs thoughts of arbitrary richness and complexity, exploiting conceptual resources that are already available or may develop with the availability of structured expressions*. (Chomsky 2007, p. 22; our emphasis)

On this view, UG primarily constrains the "language of thought," not the details of its external expression. The same conclusion was stoically reached by Newmeyer (2004, p. 545): "Typological generalizations are therefore phenomena whose explanation is not the task of grammatical theory. If such a conclusion is correct, then the explanatory domain of Universal Grammar is considerably smaller than has been assumed in much work in the Principles-and-Parameters approach" and Chomsky (2007, p. 18) seems in part to concur: "Diversity of language provides an upper bound on what may be attributed to UG."

These then are simply different enterprises – Chomsky is concerned with the nature of recursive thought capacities, whereas linguistic typology and the non-generative linguists are concerned with what external language behavior indicates about the nature of cognition and its evolution. We have argued that the latter program has more to offer cognitive science at this juncture in intellectual history. Perhaps a mark of this is that our cross-linguistic enterprise is actually close to Chomsky's new position in some respects, locating recursion not as a universal property of (linguistic) syntax, but as a universal property of language use (pragmatics, or mind) – a fact, though, that emerges from empirical cross-linguistic work.

R3. How much of the design space is populated?

Pinker & Jackendoff point out no doubt correctly that the possible design space for human languages is much greater than the space actually explored by existing languages. Two basic questions arise: (1) What exactly are the dimensions of the possible design space, of which just one corner is actually occupied? (2) What exactly does this sequestration in a small corner imply?

Before we get too excited by (1), we should consider (2). **Pinker & Jackendoff** imply that languages are locked into the corner by intrinsic, innate constraints, and that's why we don't find languages with really outlandish properties. But there is a fundamental fact they have overlooked. The earliest modern human remains date back to about 200,000 BP, and outside Africa date from only 100,000 years or so ago. If that is the date of the great diaspora, there has been relatively little time for diversification. Let us explain.

We have to presume that most likely all the languages we have now are descended by cultural evolution from a single ancestral tongue (it would take an event of total spoken language loss to be otherwise – not impossible, but requiring a highly unlikely scenario, such as an isolated lineage descended from a deaf couple). Now consider the following surprising fact. The structural properties of language change on a near-glacial time scale. In an ongoing study using Bayesian phylogenetics, Dunn et al. (in preparation) have found that taken individually, a structural feature within a single large language-family like Austronesian changes on average just once about every 50,000 years.[4] What that implies is that *all the languages we now sample from are within structural spitting distance of the ancestral tongue!* It is quite surprising in this light that typologists have been able to catalogue so much linguistic variation. Once again, a coevolutionary perspective is an essential corrective to the enterprise.

So whether we need a lot of further explanation for the fact that languages seem to be cultivating the same garden (**Tallerman**), to the degree that this can be shown, depends crucially on the extent to which you think the languages of the world are independent experiments. Francis Galton, who stressed the need for genealogical independence in statistical sampling, would urge caution!

Let us turn now to the properties of the design space. **Pinker & Jackendoff** point out that we set aside a rich set of functional universals on the grounds that they are definitional of language (a move we borrowed directly from Greenberg). Of course it is not trivial that these seem shared by all human groups (although very little empirical work has actually been done to establish this – but see, e.g., Stivers et al. 2009). We think that there is a clear biological infrastructure for language, which is distinct from the structural properties of language. This consists of two key elements: the vocal apparatus and the capacity for vocal learning, on the one hand (both biological properties unique in our immediate biological family, the Hominidae), and a rich set of pragmatic universals (communicative intention recognition prime among them), on the other. This is the platform from which languages arise by cultural evolution, and yes, it limits the design space, like our general cognitive and learning capacities (**Christiansen & Chater**). We emphasized that those interested in the evolution of the biological preconditions for language have been looking in the wrong place: Instead of looking at the input-output system (as Philip Lieberman has been urging for years; see, e.g., Lieberman 2006), or the pragmatics of communicative exchange, they've been focussed on the syntax and combinatorics, the least determined part of the system, as demonstrated by linguistic typology.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

A functional perspective has been a long running undercurrent in typological and descriptive linguistics, as **Croft** and **Goldberg** remind us. Goldberg suggests that the design space is highly constrained by systems motivations; for example, pressures to keep forms distinct while staying within the normal sound patterns of a language. These pressures provide explanations for the internal coherence of language structure, a perspective that is indeed necessary to explain how languages are not for the most part a heap of flotsam and jetsam accumulated during cultural evolution, but rather, beautifully machined systems, with innovations constantly being adjusted to their functions.

Returning to the question of how saturated or otherwise the design space is, **Pinker & Jackendoff** maintain it is easy to think of highly improbable but possible language types, and they suggest a few. Quite a few of these simply fail on the functional front – they are unproductive like their Cadda, or limited in expressiveness like their Bacca, and groups confined to speaking such languages would rapidly lose out to groups with more expressive systems. Daffa, the quantificational-logic language, lacks any form of deictics like "I," "'you," "this," "now," or "here": The presence of some deictics is certainly a functional universal of human language and follows from the emergence of human language from interactional, socially situated transactions.

Interestingly, though, some natural languages do have properties that partake of **Pinker & Jackendoff**'s thought experiments. For example, their imaginary Cadda, a language of one word holophrases, lacks double articulation. The three-generation sign language of Al Sayyid is also said to lack double articulation (Meir et al., in press; Sandler et al. 2009), showing that this has to arise by cultural evolution: it is not given by instinct.

The musical language Gahha, likewise, isn't too far off attested reality. The West Papuan language Iau (Bateman 1986a; 1986b; 1990a; 1990b) has eight phonemic tones (including melodic contours), close to the number of phonemic segments, and uses them both for lexical distinctions and for grammatical distinctions including aspect, mood, and speech-act distinctions; other tone languages use pitch to indicate modality or case (e.g., Maasai).

Nor is the "rational" Fagga too far "outside the envelope." Sure, it would require a level of semantic factorization down to a number of combinable semantic components not larger than the number of phonemes, but some semantic theories posit a few score "semantic primitives" in terms of which all meanings can be stated (e.g., Goddard & Wierzbicka 2002), and Demiin, the Lardil initiation language, maps the entire lexicon down to around 150 elements (Hale 1982). Combine Demiin semantics with !Xóõ phonology (159 consonant phonemes on some analyses), pair one semantic element to each phoneme, and Fagga might just squeak in.[5] Whether or not it then actually existed would depend on whether a possibly evolutionary route past the "historical filters" could be found – in other words whether an evolutionary pathway could exist to reach this highly economical mapping of meaning elements onto phonological segments.

Finally, it is salutary to recollect that it is only relatively recently that we have come to recognize sign languages as fully developed languages with equal expressive power to spoken languages. These languages with their easy access to iconicity and analog spatial coding break out of the design space restricted by the strictly linear coding of the vocal-auditory channel. The typology of these languages is still in development, and there are plenty of surprises yet to come (see Meir et al., in press; Zeshan 2006a; 2006b).

R4. Language variation and the future directions of cognitive science

R4.1. Is history bunk? Linguistics as a science of change

> *History is more or less bunk. It's tradition. We don't want tradition. We want to live in the present and the only history that is worth a tinker's dam is the history we make today.*
>
> — Henry Ford (Interview in *Chicago Tribune*, May 25, 1916)

Nevins' dismissal of the coevolutionary approach we are advocating as "hand-waving at diachronic contingencies" hints at another kind of dissatisfaction with the general program we outlined in the target article: the suspicion that we advocate an old-fashioned historical and cultural approach, which will return linguistics wholly to the humanities. The antipathy to history is based on the view that (a) it is the study of particularities, whereas we should be in the business of generalizing, (b) it cannot be predictive, while any empirical science should make falsifiable predictions.

But the study of evolution is centrally about history, the study of the match between organisms and environment over time, and few would doubt its scientific credentials. And modern linguistics began as a historical discipline, that rapidly felt able to announce laws of historical change, while recent sociolinguistics has been able to catch language change in the making.

A fundamental shift is that modern computational methods have revolutionized the possibility of studying change in systems using more and more complex and realistic simulations. Within the study of evolution, computational cladistics exploits this to the full, using, for example, Bayesian inference to run millions of simulations and Monte Carlo Markov chains to search for the optimum model that predicts back the data with the greatest likelihood. We can make history today, as Henry Ford thought we should.

In the coda of the target article (sect. 8) we sketched a set of future directions for the language sciences based on evolutionary ideas, and these new methods put those directions within our grasp right now. Take the idea stated in thesis (3), that recurrent clustering of solutions will occur in grammars of non-closely related languages – such a claim can be tested by simulations. Equally tractable is the idea that changes cascade (thesis [4]), so that a few crucial ones may land a language in a gully of future developments. Thesis (5) about coevolution between brain, vocal organs, and language has already begun being intensively explored by simulation (Christiansen & Chater 2008; Christian et al. 2009). Thesis (7) suggests that we should investigate how the full range of attested language systems could have arisen – pie in the sky without computational simulation, but now thinkable. For example, we could follow Bickerton (1981) and start with a simple Creole-like language, described by a set of formal features or characters, and use the rates and parameters of character change derived from recent work on the Bayesian

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

phylogenetics of language families to simulate cultural evolution over more than 100,000 years. Do we derive patterns of diversity like we now see, or would we need to model special historical circumstances such as massive hybridization?

Smolensky & Dupoux ignore the recent synthesis of biological and cultural evolution. Thus they assert "language is more a biological than a cultural construct." We would go further: "language is one hundred percent a biological phenomenon." It is absurd to imagine that humans by means of culture have escaped the biosphere – we are just a species with a very highly developed "extended phenotype" or "niche construction" (Laland et al. 1999), using culture to create conditions favorable to our survival. The twin-track approach to human evolution that we sketched (derivatively from, e.g., Boyd & Richerson 1985; Durham 1991) tries to explicate this, unifying perspectives on history and phylogeny as the science of likely state changes in a population. There is immense room for future theoretical and modelling work here: without it we are not going to understand how we as a species evolved with the particular cognitive capacities we have.

R4.2. Learning and development

A number of commentators stress how two further avenues of research will help to situate our understanding of human cognition in a biological context: human development, and comparative psychology across species. For reasons of space, and reflecting the limits of our own expertise, we underplayed the crucial role of learning and cognitive development that is presupposed by the linguistic variation we outlined. These commentators offer a valuable corrective, summarizing the human and cross-species literature. They show how much more powerful are the learning mechanisms we can now draw on than the basic associationist models available in the 1950s when Chomsky argued that their lack of power forced us to postulate rich innate endowments for language learning. Indeed, the combined arguments put forth by the commentators go some way towards providing a solution to a problem we left unanswered at the end of section 7 of the target article: accounting for how language learning is possible in the face of the levels of diversity we describe.

Bavin does a good job of reminding readers what the basic issues are here, and especially the central debate over the domain-specificity of language learning. **Tomasello** observes that the Chomskyan argument about the unlearnability of language structure crucially relies on the assumption of a simple association learning: once we take into account the rich context of communication, with shared attention and intention interpretation, not to mention capacities for analogy and statistical learning, the argument fails. **Catania** also refers to work on other species showing that category discrimination can be triggered right across the board by a single new stimulus. Catania, **Christiansen & Chater**, and **Merker** all stress the funnelling effects of the learner bottleneck via "C-induction": In Merker's words. "cultural transmission delivers the restricted search space needed to enable language learning, not by constraining the form language takes on an innate basis, but by ensuring that the form in which language is presented to the learner is learnable." Catania suggests we explicitly incorporate a "third track" – acquisition – into our coevolutionary model, but we would prefer to maintain it as one (albeit powerful) set of selectors on linguistic structure alongside the others we outline in our article.

A number of commentators dwelt on Chomsky's "poverty of the stimulus" argument for rich innate language capacities. **Bavin** points out that the complex sentential syntax that motivates the argument is learnt so late that the child has wide experience of language on which to build. Perhaps the neatest refutation is provided by **Waterfall & Edelman**, who note a crucial property of the linguistic input to children: namely, repetition with minor variation, which draws attention to the structural properties of strings, exhibiting for the infant the "transformations" of Zellig Harris. They show how learning algorithms can effectively use this information to bootstrap from unsegmented text to grammatical analysis.

McMurray & Wasserman correctly point out that our position radically moves the goal posts for language learning, suggesting that not only are a slew of specialized learning strategies employed to learn a language (and these commentators provide very useful bibliographic leads here), but *which* of these strategies is deployed may depend on the language being learnt. We don't necessarily learn Yélî Dnye with its 90 phonemes, flexible phrase order, and widespread verb suppletion using the same strategies we use for English: As McMurray & Wasserman write, "humans ... assemble language with a variety of learning mechanisms and sources of information, this assembly being guided by the particularities of the language they are learning." Instead of talking about the passive acquisition of language, we should be talking about the active construction of many different skills. This perspective buries the idea of a single language acquisition device (LAD).

Christiansen & Chater, as well as **Catania**, emphasize that learning in development is the crucial filter through which languages have to pass. Languages have to be good to think with (to modify an adage of Levi-Strauss), otherwise they won't make it. Christiansen & Chater have described (both in their 2008 *BBS* article [see *BBS* 31(5)] and in Christiansen et al. 2009) interesting modelling that shows that the learning filter must be largely language-independent, and thus that properties of learning are unlikely to have evolved specifically for language. This is a new kind of evidence against the position taken by **Pinker & Jackendoff** that language-specific learning principles must be involved in the acquisition of language.

Finally, we would like to draw attention to one other crucial aspect of development, namely, the way in which the environment is known to modulate developmental timing in the underlying biology of organisms, so that phenotypic variation can be achieved from the same genotype ("phenotypic plasticity"), and conversely, phenotypic identity can be obtained from variant genotypes ("developmental buffering"). In the conclusion to our target article we drew attention to the extraordinary achievement that is culture – generating phenotypic difference where there is no genetic difference, and phenotypic identity where there is genetic difference. These issues have been much explored in the biological literature on epigenesis and

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

development (see West-Eberhard [2003] for a fascinating overview).

R4.3. The comparative perspective across species

Our other major omission, as some commentators noticed, is the lack of reference to the comparative psychology of other species. **Margoliash & Nusbaum** appeal to linguists and others interested in the evolution of language to "cast off the remaining intellectual shackles of linguistic speciesism" and take the findings of animal research more into account. They usefully remind us of the importance of the relationship between perceptual and motor skills. **Merker** notes how findings about complex learned birdsong can explain how a prelinguistic human adaptation for emancipated song could provide a mechanism for sustaining and elaborating string transmission, even if this was timed before the full emergence of social cognition: it can be driven by the need to impress by elaborate vocal display even when not yet used to communicate meaning. Darwin (1871) had, of course, imagined that language evolved from song (see Fisher & Scharff 2009; Fitch 2006, for an update).

Penn, Holyoak, & Povinelli (**Penn et al.**) point out that our demonstration of the variability in language, and the implication that there is no simple innate basis for it, has interesting implications for a central issue in comparative psychology: what exactly is the Rubicon which divides us from apes? If the crucial ingredient was a chance language gene or the genetic substrate for UG, it might be possible to argue that language alone is responsible for the sea-change in our cognition. But if there is no such magic bullet, then languages must be learnt by principles of general cognition, and the Rubicon must be constituted by more fundamental and more general differences in cognition.

Penn et al. err, though, when they try to extend the argument to downplay Tomasello's (2008) thesis that the crucial divide is the special assemblage of abilities that make up the pragmatic infrastructure for human language. Tomasello's assemblage of specialized social cognition is precisely what we need to explain the genesis of language diversity – it provides a general platform both for language learning and for the elaboration of distinct systems. Still, bringing their point together with those by **Margoliash & Nusbaum** and **Merker** is a useful reminder that we need to account both for the emergence of *patterned form* (where cross-species studies of sophisticated vocalizers must take on greater importance) and of *productive meaning* (where social cognition is likely to remain the main driver).

Penn et al. see in our display of language variation more evidence for their identification of a major discontinuity between apes and humans in the capacity for relational thought. If this capacity is not introduced by a single new evolved trait, human language, then the gulf is a feature of general cognition. But we note two caveats here: First, in our very nearest cousins (chimps and bonobos), there are pale shadows of relational thinking (Haun & Call 2009). Second, no one doubts the importance of language in delivering ready-made relational concepts (Gentner 2003). Beyond that, we probably agree about the facts, but might value them differently: Is 10% continuity with chimps a telling bit of continuity, or is 90% discontinuity a hopeless Rubicon?

R5. Situating language and cognition in the biology of variation

Science moves in new directions blown by winds of different kinds – Kuhnian collapses, new technologies, new integrative insights, newly developing fields, funding biases, even boredom with old paradigms. We think it is pretty clear that for a mix of these reasons, the cognitive sciences are about to undergo a major upheaval. Classical cognitive science was based on a mechanistic analogy with a serial computational device, where serial algebraic algorithms could represent models of the mind. A simplifying assumption was made at the outset: we need only characterize one invariant system. That is, the human mind is essentially an invariant processing device, processing different content to be sure, but running the same basic algorithms regardless of its instantiations in different individuals with different experiences, different environments, and different languages (cf. **Smolensky & Dupoux**'s "a universal principle is a property true of all minds").

This view has taken a number of knocks in the last twenty years; for example, from the success of parallel computational models and the rise of the brain sciences. The brain sciences were at first harnessed to the classical enterprise, with invariance sought beneath individual variation in brain structure and function through selecting only right-handed or male subjects, pooling data, and normalizing brains. But cognitive neuroscience has increasingly broken free, and now the range of individual biological variation is a subject of interest in its own right.

Pushing this development is genetics. It is now feasible to correlate brain structure and function with scans across half a million single nucleotide polymorphisms (SNPs) or genetic markers. We already know detailed facts about, for example, the alleles that favor better long-term memory (Papassotiropoulos et al. 2006), and we are well on the way to knowing something about the genetic bases of language (Fisher & Marcus 2006, Vernes et al. 2008). On the processing side, we know that about 8% of individuals have right-lateralized language, that individuals differ markedly in the degree of language lateralization, and that on specific tasks about 10% of individuals may not show activation of the classic language areas at all (Müller 2009). (True, most individuals will have circuitry special to language, as **Pinker & Jackendoff** remark, but that may be only because using language bundles specific mental tasks, and because adults have built the circuitry in extended development.) We even have preliminary evidence that gene pools with certain biases in allele distribution are more likely to harbour languages of specific sorts (Dediu & Ladd 2007). We are not dealing, then, with an invariant machine at all, but with a biological system whose evolution has relied on keeping variance in the gene pool.

This research is going to revolutionize what we know about the mind and brain and how it works. By putting variation central, as the fuel of evolution, it will recast the language sciences. Some aspects of the language sciences are pre-adapted to the sea-change –sociolinguistics, dialectology, historical linguistics, and typology – provided they can take the new mathematical methods on board. But we can look forward to the new psycholinguistics, centrally concerned with variation in human performance in the language domain both within and across language groups, and the new neurocognition of language which will

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

explore both the varying demands that different languages put on the neural circuitry and the way in which superficial phenotypic standardization is achieved by distinct underlying processing strategies in different individuals.

In this context, renewed interest in the variation in human experience and expertise, in the cultural contexts of learning, and the diversity in our highest learned skill – language – is inevitable. For the cognitive and language sciences to engage with these developments, a first step is to take on board the lessons of those linguistic approaches that place variation and process at centre stage. Then the very diversity of languages becomes no longer an embarrassment but a serious scientific resource. That is the message we have been trying to convey.

R6. Appendix: Disputed data and generalizations

R6.1. Kayardild nominal tense

The occurrence of tense on Kayardild nominals was cited by us as a counterexample to Pinker and Bloom's (1990) claim that languages will not code tense on nominals. **Baker**'s commentary does not dispute this, but then tries to use it to establish an orthogonal issue, namely, his verb-object constraint (see sect. R.6.10). While it is true that in Kayardild, tense appears on objects rather than subjects, it is not hard to find other languages, such as Pitta-Pitta (Blake 1979), where it is the subject rather than the object that codes for tense – so the general phenomenon gives no succor to Baker's hoped-for universal. Needless to say, all this only reinforces the fact that tense can occur on nominals.

R6.2. Positionals and ideophones

We noted in the target article that not only are the "big four" word classes (noun, verb, adjective, adverb) not wholly universal, but there were plenty of other word classes out there, including positionals and ideophones. We used the example of Mayan positionals. **Pesetsky** is right that Mayan positionals are classically defined as a root class, not a stem class, but the facts are actually more complex (see, e.g., Haviland 1994). Positionals have their own unique distribution at the stem level too, occurring, for example, in a special class of mensural classifiers (de Léon 1988), body-part constructions (Haviland 1988, p. 92) and color-plus-position constructions (Haviland, submitted). In any case, many languages from around the world (such as Yélî Dnye; Levinson 2000) have positionals as a special word class with their own distinctive distributions. (See Ameka and Levinson [2007] for detailed examples and a typology.)

Pesetsky similarly tries to undermine the status of ideophones/expressives as a word class (the terms are more or less synonymous, but come from different linguistic descriptive traditions). He correctly notes that Osada (1992) does not discuss their syntax in Mundari, and this reflects a general neglect of their syntactic characteristics in linguistic descriptions, apart from simplistic characterizations of them as "syntactically unintegrated." However, a careful discussion of the syntax of the functionally similar class of *expressives* in another Austroasiatic language, Semelai, can be found in Kruspe (2004): their syntactic distribution closely parallels that of direct speech complements. Likewise in Southern Sotho (Molotsi 1993), ideophones pattern like complements of "say," with the further property that they can be passivized, so that "John snatched the woman's purse" is literally "John said snatch woman's purse," which can be passivized as "snatch was said woman's purse." In short, ideophones and expressives have a syntax, if sometimes limited.

R6.3. Straits Salish noun versus verb distinction

We pointed out that it was still unclear whether in fact there is a universal noun/verb distinction. We mentioned the Wakashan language Straits Salish as an example of a language plausibly claimed to lack a noun/verb distinction. Instead of presenting counteranalyses of the Straits Salish data, **Tallerman** cites data from Nuuchahnulth (Nootka), from another language family, with no demonstration that the arguments can be transferred to Straits Salish. A crucial difference between the languages is that names can be predicative in Straits Salish but not in Nootka. Tallerman's major arguments for the existence of a noun/verb distinction in Nuuchahnulth were already given in Jacobsen (1979) and Schachter (1985), and Jelinek (1995) takes care to show that they don't apply to Straits Salish, which is why we used Salish rather than Nootka as an example. We agree with her, though, that further investigation of the Salish case is needed (a point also articulated in Evans & Osada 2005); hence our statement that no definitive consensus has been reached.

R6.4. Jemez/Kiowa number

Harbour reproaches us for attributing the "unexpected number" facts to Jemez rather than Kiowa; in fact, the languages are related and both exhibit similar phenomena (Mithun 1999, p. 81, and personal communication). We thank Harbour for picking up the factual errors he points out, but for our part would like to correct his mischaracterization of this case as our "prime example" of "something we would never think of" – it was one of many, and the rest still stand. More importantly, further cross-linguistic data disputes his claim that "singular 'we' arises because Winnebago uses only [± augmented]." The use of "because" here illustrates the fallacy of inferring cause from single cases. Harbour's formulation predicts that if a language uses a more elaborated grammatical number system than just [± augmented] it should not treat "1 + 2" as singular. Yet there are many languages which have a three-way number system and which nonetheless treat 1+2 in the same series as the singulars, like Bininj Gun-wok (Evans 2003a).

R6.5. Arrernte syllable structure

Nevins, and (briefly) **Berent**, take issue with our citing Arrernte as an example of a language that defies the "Universal CV preference" by taking VC as the underlying syllable type. To contextualize their riposte, it is worth quoting Hyman (2008, p. 13):

> In each of the above cases, there is no "knock-out argument." Anyone determined to maintain [these] universals can continue to do so, the worst consequence being an indeterminate or more awkward analysis. . . . Architectural universals have

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

this property: it all depends on your model and on what complications you are willing to live with.

Nevins' purported counter-analysis is of this type. To make it work is not just a matter of allowing onset-sensitive morae, not a problem in itself, but also of leaving the coda out of weight considerations, which is more problematic. Moreover, he only considers some of the phenomena that Breen and Pensalfini (1999) cite – such as the fact that the language game known as "Rabbit Talk" picks out exactly the VC syllable to move to the end of the word – and ignores the arguments they give for postulating an initial vowel in words which start with a C when pronounced in isolation; namely, that this vowel appears when the word is not pronounced breath-group initially, and that postulating it simplifies other morphonological processes. A further argument in favor of the VC analysis (see Evans 1995b) is that although there is considerable variation in how words are pronounced in isolation (e.g., "sits" can be pronounced [anə́mə], [anə́m], [nə́mə], or [nə́m]), the number of syllables remains constant under the VC syllable analysis (at 2 in this instance), whereas the number of syllables under other analyses remains inconstant, even with the moraic adjustments that Nevins proposes. In short, proposing VC syllables lines up beautifully with a whole range of rules, whereas adopting the alternative, while workable, is crabbed by inelegancies.

A deeper problem than mere inelegance in forcing a language like Arrernte into a procrustean CV bed is that it draws attention away from explaining what forces have shaped the unusual Arrernte structure. There is growing evidence from phonetic work by Butcher (2006) that the Arrernte VC syllable represents the phonologization of a whole syndrome of phonetic and phonological effects at work in Australian languages, linking a number of phenomena like: (a) the unusual proliferation of distinctive heterorganic clusters intervocalically (e.g., nk vs. ŋk vs. ɲk vs. ɳk); (b) the large set of place contrasts for oral and nasal stops, including contrasts like alveolar versus postalveolar, that are most effectively cued by the leading rather than following vowel; (c) the neutralization of the apico-alveolar versus apico-postalveolar contrast word-initially; and (d) the widespread pre-stopping of intervocalic nasals and laterals.

The joint effect of all these features is to concentrate the maximum amount of contrasting information in intervocalic position, and make the leading vowel crucial for signalling the place of following consonants through F2 and F3 formant transitions. In other words, it is VC rather than CV units (or, more accurately, the continuous phonetic signals that correspond to them) which are the most informative, in terms of cueing the greater number of contrasts. This now allows us to give an insightful account of why VC syllables emerge as phonological units in some Australian languages. We would not be led to this explanation if we use too much abstract representational machinery to conjure away the existence of an aberrant pattern.

R6.6. Finite state grammars and cotton-top tamarins

Pullum & Scholz pull us up for propagating a misinterpretation of the findings in Fitch and Hauser (2004), by stating that cotton-top tamarins have a general ability to learn finite state languages. We stand corrected, and urge the reader to heed Pullum & Scholz's clarification that Fitch and Hauser's findings are restricted to the much smaller subset known as SL (strictly local) languages.

The investigation of recursive and augmentative structures in animal cognition is a current minor industry in cognitive science. If this is meant to shed light on the human language capacity, it is arguably quite misguided. Indefinite recursion, or discrete infinity as Chomsky prefers, is not an actual property of human language – no human is capable of indefinite centre-embedding, for example. Only in the light of a radical distinction between competence and performance does this minor industry make any sense at all, and that little sense is undermined by the impossibility of testing animals directly for indefinite recursion.

R6.7. Cinque's generalization about Greenberg's Universal #20

Specifying strict ordering in noun phrases where the noun comes last, is raised by **Rizzi** as an example of how implicational universals can be made to follow from parameterized rules. However, Dryer (2009), drawing on a larger cross-linguistic sample, shows that you get better fit with the data if Cinque's formal categories (like Adjective) are replaced by semantic categories (like "modifier denoting size, color, etc."). Cinque's parameterization just gives a discrete and approximate characterization of statistical trends reflecting the interaction of many functional selectors.

R6.8. Subjacency and "invisible Wh-movement"

A number of commentators (**Smolensky & Dupoux, Freidin, Rizzi**) appealed to the Chomskyan notion of "Subjacency" as a convincing example of a highly abstract principle or rule-constraint which is manifested directly in languages like English. The idea in a nutshell is that movement of constituents is constrained so that they may not cross more than one "bounding node" in the syntactic tree (in English, bounding nodes are a NP, i.e., noun phrase, or a complementizer phrase headed by *that*). Hence you can say "What does John believe that Mary saw __?" but not "*What does John believe the rumor that Mary saw _?".

Now consider **Rizzi**'s point that many languages, including Chinese, do not move their Wh-words (so called in situ *Wh*) – they would stay in the corresponding slots indicated in the just provided sentences – but appear to exhibit semantic interpretations that might constitute a parallel phenomenon. The apparent lack of Wh-movement in Chinese, which at first seems an embarrassment to the theory, is claimed however to mask *covert movement* at an underlying level, close to semantic interpretation: consequently the range of construals of a Chinese Wh-question is argued to be limited by the very same abstract constraint postulated for languages with overt movement (see examples in Rizzi's commentary). For generativists, this may seem like a double scoop: Not only is the constraint of an abstract enough kind that children would find it hard to learn in English, but it even holds in Chinese where it is, in effect, invisible, so could not possibly be learnt! Moreover, it is a completely arbitrary and

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

unmotivated constraint, so there is no apparent way for the child to infer its existence. Therefore, it must be part of UG, a quirk of our innate language organ.

But this in fact is not at all a convincing example to the other camp. First, to make it work in languages with and without overt "movement," it has to be so particularized ("parameterized") for each language so that, as we noted in the target article, the child might as well learn the whole thing (Newmeyer 2005). Second, there are perfectly good alternative models that do not use movement: Wh-words are simply generated in the right place from the start, using other methods than movement to get the correct logical interpretations. Within LFG, a combination of the FOCUS discourse function and prosodic structure can get in situ Wh interpretation with no covert movement required (Mycock 2006). Through methods like these, LFG, HPSG, and Role and Reference Grammar have all developed ways of modelling both the English syntactic constraints and the Chinese interpretation constraints without any covert operations or unlearnable constraints.

Van Valin (1998) offers one of these rival explanations.[6] He notes that for entirely general purposes one needs to have a notion of "focus domain" – roughly the unit that can be focussed on as new information in a sentence. A chunk like *Mary did X* is such a unit, but *the rumor that Mary did X* is not, because it marks the information as already presumed. So it makes no sense to question part of it. Focus domains have a precise structural characterization, and the informational structure of this kind explains both the English and the Chinese facts without positing covert entities or unmotivated rule constraints. Van Valin shows that the focus domains are easily learned by children from the range of possible yes/no question elliptical answers. This explanation needs the minimum equipment (a definition of focus domain) and no magic or UG.

Take your pick between the two explanations – an unmotivated, unlearnable, hidden constraint implying innate complex architecture, or a general design for communication requiring nothing you wouldn't need for other explanatory purposes. As C.-R. Huang (1993) notes after discussing the Mandarin data, "there is no concrete evidence for an abstract movement account … invoking Ockham's razor would exclude movements at an abstract level."

R6.9. C-command

Rizzi claims that "no language allows coreference between a pronoun and a NP when the pronoun c-commands the NP" (**He said that John was sick*; **each other saw the men*). We pointed out that in languages (like Jiwarli) which lack constituency as the main organizing principle of sentence structure, notions like c-command cannot be defined (c-command is defined in terms of a particular kind of position higher in a syntactic constituency tree). But let us interpret this relation loosely and charitably, in terms of some general notion of domination or control. Then the observation would have very wide validity, but it would still be only a strong tendency. Counterexamples include Abaza reciprocals (Hewitt 1979) where the verbal affix corresponding to "each other" occupies the subject rather than the object slot, and Guugu Yimidhirr pronominalization, where it is possible to have a pronoun in the higher clause coreferential with a full NP in the lower clause (Levinson 1987).

Once again, then, we are dealing with a widespread but not universal pattern. The typological/functional paradigm explains it as emerging from a more general tendency in discourse (not just syntax): reference to entities proceeds with increasing generality, which is why "She came in. Barbara sat down" is not a way of expressing "Barbara came in. She sat down." (see Levinson [2000] for a detailed Gricean account). Many languages have grammaticalized the results of this more general tendency, producing grammatical rules which can then be described by c-command (if you want to use that formalism) but also by other formalisms. Seeking the most general explanation for cross-linguistic patterning here directs us to more general pragmatic principles ("use the least informative form compatible with ensuring successful reference given the current state of common ground"), rather than in terms of a specific syntactic constraint which only applies in a subset (even if a majority) of the world's languages. Many strong tendencies across languages appear to have a pragmatic or functional base, undermining a presumption of innate syntax.

R6.10. The "Verb-Object Constraint"

Baker offers his "Verb-Object Constraint (VOC)" as a proposal for a "true linguistic universal" of this high level kind – the generalization that the verb "combines" with the theme/patient before a nominal that expresses the agent/cause ("combines" is not defined, so we take it loosely). But this, too, rapidly runs afoul of the cross-linguistic facts. Note that his formulation equivocates between whether the constraint is formulated in terms of semantic roles such as agent and patient, or grammatical relations such as subject and object; some of the problems below pertain to one of these, some to the other, some to both:

1. Many languages don't have a clear notion of subject and object (see remarks in our target article). If we avoid this problem by stating the universal in terms of thematic roles (theme, patient, agent, experiencer), then we'll find such anomalies as languages which effectively idiomatize the subject-verb combination, only combining secondarily with the patient, employing idioms like "headache strikes me/the girl" or "fever burns him" (Evans 2004; Pawley et al. 2000).

2. Although polysynthetic languages like Mohawk usually incorporate objects rather than subjects into the verb, there are some that do incorporate transitive subjects/agents (not just objects as **Baker**'s generalization would predict), most famously the Munda language Sora (Ramamurti 1931; cf. Anderson 2007).

3. There are twice as many VSO languages as VOS languages, 14% versus 7%, respectively, in a worldwide sample by Dryer (2009), but only VOS languages seem likely to facilitate a "combination" of verb and object.

4. Languages with ergative syntax group the object of transitives and the subject of intransitives as one type of entity, around which the syntax is organized (**Baker** notes this as a potential problem, but doesn't offer the solution).

Taken together, these problems make the VOC just one more observation that is certainly a statistical tendency, but which it is misleading to elevate to "universal" status.

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

ACKNOWLEDGMENTS

We thank the following people for discussion, comments, and ideas in preparing this response: Mary Beckman, Balthasar Bickel, Penelope Brown, Andy Butcher, Grev Corbett, Bill Croft, Nick Enfield, Adele Goldberg, Martin Haspelmath, Yan Huang, Larry Hyman, Rachel Nordlinger, Masha Polinsky, Arie Verhagen, and Robert Van Valin.

NOTES

1. We use the term *generative linguists* to refer to linguists working specifically within frameworks deriving from the various theories of Chomsky. The term also has a wider sense, referring to a larger body of researchers working in fully explicit formal models of language such as LFG, HPSG, and their derivatives. These alternative theoretical developments have been much less wedded to the Chomskyan notion of Universal Grammar. LFG, in particular, has explicitly developed a much more flexible multidimensional architecture allowing for both constituency and dependency relations as well as the direct representation of prosodic units.

2. Of course these need to be relativized to modality: facts about the position of the larynx or the stability of some vowel formants across varying vocal tract configurations are only relevant to sound, whereas constraints on the production of hand or arm gestures are only relevant to manual sign. There will be some parallels, but the degree to which "sonority" is the same phenomenon in both, as **Berent** suggests, is still controversial (Sandler 2009; Sandler & Lillo-Martin 2006, p. 245).

3. Hockett (1960) correctly identified this as part of the "duality of patterning" (together with combinatorial semantics) necessary if language is to be unlimited in its productivity.

4. Lest this finding invite incredulity, given that the language family is assumed to be less than 6,000 years old, this figure is worked out by summing independent path-lengths in many branches of the family tree and looking for the total numbers of changes from an ancestral language. The number should be taken with a pinch of salt but is probably in the right general ballpark.

5. Abui, on František Kratochvil's (2007) analysis, comes rather close.

6. For other kinds of explanation in terms of processing costs, see Kluender (1992; 1998), Hawkins (1999), and Sag et al. (2007).

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively.]

Ameka, F. & Levinson, S. C., eds. (2007) The typology and semantics of locative predicates: Posturals, positionals and other beasts. *Linguistics* 45(5/6):847–72. [Special issue] [arNE]

Anderson, G. D. S. (2007) *The Munda verb. Typological perspectives.* Mouton de Gruyter. [rNE]

Anderson, P. W. (1972) More is different. *Science* 177(4047):393–96. [AEG]

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. [aNE]

Aoki, K. & Feldman, M. W. (1989) Pleiotropy and preadaptation in the evolution of human language capacity. *Theoretical Population Biology* 35:181–94. [aNE]

(1994) Cultural transmission of a sign language when deafness is caused by recessive alleles at two independent loci. *Theoretical Population Biology* 45:253–61. [aNE]

Arbib, M. A. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2):105–24. [aNE]

Archibald, L. M. & Gathercole, S. E. (2007) The complexities of complex memory span: Storage and processing deficits in specific language impairment. *Journal of Memory and Language* 57:177–94. [ELB]

Arnold, K. & Zuberbühler, K. (2006) Language evolution: Semantic combinations in primate calls. *Nature* 441(7091):303. [GKP]

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. 133–49. John Benjamins. [aNE]

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. [aNE]

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

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. [aNE]

Austin, P. & Bresnan, J. (1996) Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory* 14(2):215–68. [arNE]

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

(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. [aNE]

(1996) *The polysynthesis parameter*. Oxford University Press. [MCB, aNE]

(2001) *Atoms of language: The mind's hidden rules of grammar*. Basic Books. [MCB, aNE, MH]

(2003) Linguistic differences and language design. *Trends in Cognitive Science* 7:349–53. [aNE]

(in press) Formal generative typology. In: *Oxford handbook of linguistic analysis*, ed. B. Heine & H. Narrog. Oxford University Press. [MCB]

Barwise, J. & Cooper, R. (1981) Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219. [aNE]

Bateman, J. (1986a) Tone morphemes and aspect in Iau. *Nusa* 26:1–50. [rNE]

(1986b) Tone morphemes and status in Iau. *Nusa* 26:51–76. [rNE]

(1990a) Iau segmental and tone phonology. *Nusa* 32:29–42. [rNE]

(1990b) Pragmatic functions of the tone morphemes and illocutionary force particles in Iau. *Nusa* 32:1–28. [rNE]

Bates, E., Devescovi, A. & Wulfeck, B. (2001) Psycholinguistics: A cross-language perspective. *Annual Review of Psychology* 52:369–96. [aNE]

Bates, E. & Goodman, J. (1999) On the emergence of grammar from the lexicon. In: *Mechanisms of language acquisition*, ed. B. MacWhinney, pp. 29–79. Erlbaum. [ELB]

Bates, E. & MacWhinney, B. (1987) Competition, variation and language learning. In: *Mechanisms of language acquisition*, ed. B. MacWhinney, pp. 157–94. Erlbaum. [ELB]

Bavin, E. L., Wilson, P., Maruff, P. & Sleeman, F. (2005) Spatio-visual memory of children with specific language impairment: Evidence for generalized processing problems. *International Journal of Language and Communication Disorders* 40:319–32. [ELB]

Baylis, J. R. (1982) Avian vocal mimicry: Its function and evolution. In: *Acoustic communication in birds*, ed. D. E. Kroodsma & E. H. Miller, pp. 51–83. Academic Press. [BMe]

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

Becker, M., Ketrez, N. & Nevins, A. (submitted) The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. [IB]

Berent, I. (2008) Are phonological representations of printed and spoken language isomorphic? Evidence from the restrictions on unattested onsets. *Journal of Experimental Psychology: Human Perception and Performance* 34:1288–1304. [IB]

Berent, I., Lennertz, T., Jun, J., Moreno, M. A. & Smolensky, P. (2008) Language universals in human brains. *Proceedings of the National Academy of Sciences USA* 105:5321–25. [IB]

Berent, I., Lennertz, T., Smolensky, P. & Vaknin-Nusbaum, V. (2009) Listeners' knowledge of phonological universals: Evidence from nasal clusters. *Phonology* 26:1–34. [IB]

Berent, I., Steriade, D., Lennertz, T. & Vaknin, V. (2007) What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104:591–630. [IB]

Bermudez, J. L. (2005) *Philosophy of psychology: A contemporary introduction.* Routledge. [DCP]

Berry, L. B. (1998) Alignment and adjacency in optimality theory: Evidence from Walpiri and Arrernte. Doctoral dissertation, University of Sydney, Sydney, Australia. [IB]

Bickel, B. (2009) Typological patterns and hidden diversity. Plenary Talk, 8th Association for Linguistic Typology Conference, Berkeley, CA, July 24, 2009. [rNE]

Bickerton, D. (1981) *Roots of language.* Karoma. [arNE]

(2009) *Adam's tongue: How humans made language, how language made humans.* Hill and Wang. [DCP]

Bitterman, M. E. (1975) The comparative analysis of learning. *Science* 188(4189):699–709. [BMc]

Blake, B. J. (1979) Pitta-Pitta. In: *Handbook of Australian languages, vol. 1*, ed. R. M. W. Dixon & B. J. Blake, pp. 182–242. Australian National University (ANU) Press. [rNE]

(2001) The noun phrase in Australian languages. In: *Forty years on: Ken Hale and Australian languages*, ed. D. Nash, M. Laughren, P. Austin & B. Alpher, pp. 415–25. Pacific Linguistics. [MTa]

Blevins, J. (1995) The syllable in phonological theory. In: *Handbook of phonological theory*, ed. J. Goldsmith, pp. 206–44. Blackwell. [aNE]

(2001) Where have all the onsets gone? Initial consonant loss in Australian aboriginal languages. In: *Forty years on: Ken Hale and Australian languages*,

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

ed. J. Simpson, D. Nash, M. Laughren, P. Austin & B. Alpher, pp. 481–92. *Pacific Linguistics 512*. The Australian National University. [AN]
Boeckx, C. & Hornstein, N. (2008) Superiority, reconstruction, and islands. In: *Foundational issues in linguistic theory*, ed. R. Freidin, C. Otero & M.-L. Zubizaretta, pp. 197–225. MIT Press. [RF]
Bohnemeyer, J. & Brown P. (2007) Standing divided: Dispositional verbs and locative predications in two Mayan languages. *Linguistics* 45(5/6):1105–51. [aNE]
Bok-Bennema, R. (1991) *Case and agreement in Inuit*. Foris. [MCB]
Boroditsky, L. (2001) Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology* 43(1):1–22. [aNE]
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. [aNE]
Bowerman, M. & Levinson, S., eds. (2001) *Language acquisition and conceptual development*. Cambridge University Press. [aNE]
Boyd, R. & Richerson, P. J. (1985) *Culture and the evolutionary process*. University of Chicago Press. [arNE]
(2005) Solving the puzzle of human cooperation in evolution and culture. In: *Evolution and culture. A Fyssen foundation symposium*, ed. S. C. Levinson & P. Jaisson, pp. 105–32. MIT Press. [aNE]
Braithwaite, B. (2008) Word and sentence structure in Nuuchahnulth. Unpublished doctoral dissertation, Linguistics Section, Newcastle University, United Kingdom. [MTa]
Breen, G. & Pensalfini, R. (1999) Arrernte: A language with no syllable onsets. *Linguistic Inquiry* 30(1):1–26. [arNE, AN]
Bresnan, J. (2001) *Lexical functional syntax*. Blackwell. [aNE]
Brodsky, P., Waterfall, H. R. & Edelman, S. (2007) Characterizing motherese: On the computational structure of child-directed language. In: *Proceedings of the 29th Cognitive Science Society Conference*, ed. D. S. McNamara & J. G. Trafton, pp. 833–38. Cognitive Science Society. [HW]
Brown, P. (1994) The INs and ONs of Tzeltal locative expressions: The semantics of static descriptions of location. In: *Space in Mayan languages*, ed. J. B. Haviland & S. C. Levinson, pp. 743–90. [*Linguistics* 32(4/5): Special issue] [aNE, NP]
Brown, S., Ngan, E. & Liotti, M. (2008) A larynx area in the human motor cortex. *Cerebral Cortex* 18:837–45. [BMe]
Burzio, L. (2002) Missing players: Phonology and the past-tense debate. *Lingua* 112(3):157–99. [AEG]
Butcher, A. (2006) Australian Aboriginal languages: Consonant-salient phonologies and the "place-of-articulation imperative". In: *Speech production: Models, phonetics processes and techniques*, ed. J. M. Harrington & M. Tabain, pp. 187–210. Psychology Press. [rNE]
Butt, M., Dalrymple, M. & Holloway King, T., eds. (2006) *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan.* CSLI. [rNE]
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. 250–68. Cambridge University Press. [aNE]
(2006) Language change and universals. In: *Linguistic universals*, ed. R. Mairal & J. Gil, pp. 179–94. Cambridge University Press. [BMe]
Cable, S. (2007) The grammar of Q: Q-particles and the nature of Wh-fronting, as revealed by the Wh-questions of Tlingit. Doctoral dissertation, Massachusetts Institute of Technology. Available at: http://ling.auf.net/lingbuzz/000553. [DP]
(2008) Q-particles and the nature of *wh*-fronting. In: *Quantification: A cross-linguistic perspective*, ed. L. Matthewson, pp. 105–78. Elsevier. Available at: http://people.umass.edu/scable/papers/Q-Particles-and-Wh-Fronting.pdf. [DP]
Catania, A. C. (1973) The psychologies of structure, function, and development. *American Psychologist* 28:434–43. [ACC]
(1987) Some Darwinian lessons for behavior analysis. A review of Peter J. Bowler's "The eclipse of Darwinism." *Journal of the Experimental Analysis of Behavior* 47:249–57. [ACC]
(1990) What good is five percent of a language competence? *Behavioral and Brain Sciences* 13:729–31. [ACC]
(1995) Single words, multiple words, and the functions of language. *Behavioral and Brain Sciences* 18:184–85. [ACC]
(2001) Three varieties of selection and their implications for the origins of language. In: *Language evolution: Biological, linguistic and philosophical perspectives*, ed. G. Gyori, pp. 55–71. Peter Lang. [ACC]
(2003) Why behavior should matter to linguists. *Behavioral and Brain Sciences* 26:670–72. [ACC]
(2004) Antecedents and consequences of words. *European Journal of Behavior Analysis* 5:55–64. [ACC]
(2008) Brain and behavior: Which way does the shaping go? *Behavioral and Brain Sciences* 31(5):516–17. [ACC]
Catania, A. C. & Cerutti, D. (1986) Some nonverbal properties of verbal behavior. In: *Analysis and integration of behavioral units*, ed. T. Thompson & M. D. Zeiler, pp. 185–211. Erlbaum. [ACC]
Catania, A. C. & Shimoff, E. (1998) The experimental analysis of verbal behavior. *Analysis of Verbal Behavior* 15:97–100. [ACC]
Cavalli-Sforza, L. L. (1997) Genes, peoples, and languages. *Proceedings of the National Academy of Sciences USA* 94:7719–24. [BMe]
Chater, N. & Christiansen, M. H. (in press) Language acquisition meets language evolution. *Cognitive Science*. [MHC]
Chater, N., Reali, F. & Christiansen, M. H. (2009) Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences USA* 106(4):1015–20. [MHC]
Chater, N. & Vitányi, P. (2007) 'Ideal learning' of natural language: Positive results about learning from positive evidence. *Journal of Mathematical Psychology* 51:135–63. [MHC]
Chomsky, N. (1955) Logical syntax and semantics: Their linguistic relevance. *Language* 31(1):36–45. [aNE]
(1957) *Syntactic structures*. De Gruyter. [MCB, aNE, RF]
(1959) A review of B. F. Skinner's *Verbal Behavior*. *Language* 35:26–58. [BMe]
(1964) The logical basis of linguistic theory. *Proceeding of the Ninth International Congress of Linguists*, pp. 914–1008. Mouton de Gruyter. [RF]
(1965) *Aspects of the theory of syntax*. MIT Press. [MCB, ACC, aNE]
(1973) Conditions on transformations. In: *A Festschrift for Morris Halle*, ed. S. Anderson & P. Kiparsky, pp. 232–86. Holt, Rinehart, & Winston. [RF]
(1977) On wh-movement. In: *Formal syntax*, ed. P. Culicover, T. Wasow & A. Akmajian, pp. 71–132. Academic Press. [RF]
(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. [aNE]
(1981) *Lectures on government and binding*. Foris. [aNE]
(1995) *The minimalist program*. MIT Press. [LR]
(2000) Minimalist inquiries: The framework. In: *Step by step – essays in minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka. MIT Press. [LR]
(2007) Of minds and language. *Biolinguistics* 1:1009–27. [rNE]
(2008) On phases. In: *Foundational issues in linguistic theory*, ed. R. Freidin, C. Otero & M-L. Zubizaretta, pp. 133–66. MIT Press. [RF]
Christiansen, M. H. & Chater, N. (1999) Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23:157–205. [MHC]
(2003) Constituency and recursion in language. In: *The handbook of brain theory and neural networks, 2nd edition*, ed. M. A. Arbib, pp. 267–71. MIT Press. [MHC]
(2008) Language as shaped by the brain. *Behavioral and Brain Sciences* 31(5):489–509; discussion 509–58. [BMe, MHC, arNE, AEG, DCP, HW]
eds. (2001) *Connectionist psycholinguistics*. Ablex. [BMc]
Christiansen, M. H., Chater, N. & Reali, F. (in press) The biological and cultural foundations of language. *Communicative and Integrative Biology*. [MHC]
Christiansen, M. H., Collins, C. & Edelman, S., eds. (2009) *Language universals*. Oxford University Press. [rNE]
Christiansen, M. H. & Kirby, S. (2003) *Language evolution*. Oxford University Press. [aNE]
Christiansen, M. H. & MacDonald, M.C. (in press) A usage-based approach to recursion in sentence processing. *Language Learning*. [MHC]
Christiansen, M. H., Onnis, L. & Hockema, S. A. (2009) The secret is in the sound: From unsegmented speech to lexical categories. *Developmental Science* 12:388–95. [BMc]
Christianson, K. & Ferreira, F. (2005) Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition* 98:105–35. [aNE]
Chung, S. (1989) On the notion "null anaphor" in Chamorro. In: *The null subject parameter*, ed. O. Jaeggli & K. Safir, pp. 143–84. Kluwer. [aNE]
(1998) *The design of agreement*. University of Chicago Press. [MCB]
Cinque, G. (1990) *Types of A'-Dependencies*. MIT Press. [LR]
(2005) Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36:315–32. [LR]
Clark, A. (2001) Unsupervised language acquisition: Theory and practice. Doctoral dissertation, University of Sussex, England. [BMe]
(2006) Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences* 10(8):370–74. [DCP]
Clark, H. H. (1975) Bridging. In: *Theoretical issues in natural language processing*, ed. R. C. Schank & B. L. Nash-Webber, pp. 169–74. Association for Computing Machinery. [MHC]
Clements, G. N. (1990) The role of the sonority cycle in core syllabification. In: *Papers in laboratory phonology. I: Between the grammar and physics of speech*, ed. J. Kingston & M. Beckman, pp. 282–333. Cambridge University Press. [IB]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Clements, N. & Keyser, S. J. (1983) *CV Phonology: A generative theory of the syllable.* MIT Press. [aNE]
Colarusso, J. (1982) Western Circassian vocalism. *Folia Slavica* 5:89–114. [aNE]
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. [aNE]
Comrie, B. (1976) *Aspect.* Cambridge University Press. [aNE]
(1985) *Tense.* Cambridge University Press. [aNE]
(1989) *Language universals and linguistic typology, 2nd edition*. Blackwell. [aNE]
Cook, R. G. & Wasserman, E. A. (2006) Relational discrimination learning in pigeons. In: *Comparative cognition: Experimental explorations of animal intelligence*, ed. E. A. Wasserman & T. R. Zentall, pp. 307–24. Oxford University Press. [BMc]
Coon, J. & Preminger, O. (2009) Positional roots and case absorption. In: *New perspectives in Mayan linguistics* (*MIT Working Papers in Linguistics 59*), ed. H. Avelino, J. Coon & E. Norcliffe, pp. 35–58. Massachusetts Institute of Technology. Available at: http://web.mit.edu/jcoon/www/Jessica_Coon/Publications_files/coonpremingerssila.pdf. [DP]
Coppock, E. (2008) The logical and empirical foundations of Baker's paradox. Doctoral dissertation, Stanford University. [AEG]
Corbett, G. (2000) *Number*. Cambridge University Press. [DH]
Corina, D. P. (1990) Reassessing the role of sonority in syllable structure: Evidence from visual gestural language. In: *Proceedings of the 26th regional meeting of the Chicago Linguistic Society, vol. 2*, ed. M. Ziolkowski, M. Noske & K. Deaton, pp. 33–43. University of Chicago Press. [IB]
Crain, S. (1991) Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14:597–650. [LR]
Croft, W. (1991) *Syntactic categories and grammatical relations: The cognitive organization of information.* University of Chicago Press. [WC]
(2000a) *Explaining language change: An evolutionary approach.* Longman. [WC]
(2000b) Parts of speech as language universals and as language-particular categories. In: *Approaches to the typology of word classes*, ed. P. Vogel & B. Comrie, pp. 65–102. Mouton de Gruyter. [MH]
(2001) *Radical construction grammar. Syntactic theory in typological perspective.* Oxford University Press. [WC, aNE, AEG]
(2003) *Typology and universals, 2nd edition.* Cambridge University Press. [WC, aNE, MH]
(2007) Beyond Aristotle and gradience: A reply to Aarts. *Studies in Language* 31:409–30. [WC]
(2009) Methods for finding language universals in syntax. In: *Universals of language today*, ed. S. Scalise, E. Magni & A. Bisetto, pp. 145–64. Springer. [WC]
Croft, W. & Cruse, D. A. (2004) *Cognitive linguistics.* Cambridge University Press. [WC]
Culicover, P. W. (1999) *Syntactic nuts: Hard cases in syntax*. Oxford University Press. [SP]
Culicover, P. W. & Jackendoff, R. (2005) *Simpler syntax*. Oxford University Press. [SP]
Cutler, A., Mehler, J., Norris, D. & Segui, J. (1983) A language-specific comprehension strategy. *Nature* 304:159–60. [aNE]
(1989) Limits on bilingualism. *Nature* 340:229–30. [aNE]
Cysouw, M. (2001) The paradigmatic structure of person marking. Unpublished doctoral dissertation, Radboud University, Nijmegen, Netherlands. [aNE]
Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray. [DCP]
(1871) *The descent of man and selection in relation to sex*. John Murray. [rNE, BMe, DCP]
Davidoff, J., Davies, I. & Roberson, D. (1999) Color categories of a stone-age tribe. *Nature* 398:203–204. [aNE]
Davidson, L. (2006) Phonotactics and articulatory coordination interact in phonology: Evidence from nonnative production. *Cognitive Science* 30:837–62. [IB]
Davis, S. (1988) Syllable onsets as a factor in stress rules. *Phonology* 5:1–19. [AN]
Dawkins, R. (1982) *The extended phenotype*. Freeman. [ACC]
de Lacy, P. (2006) *Markedness: Reduction and preservation in phonology*. Cambridge University Press. [IB]
(2008) Phonological evidence. In: *Phonological argumentation: Essays on evidence and motivation*, ed. S. Parker, pp. 43–77. Equinox. [IB]
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* 104:10944–49. [arNE]
de León, L. (1988) *Noun and numeral classifiers in Mixtec and Tzotzil: A referential view*. Unpublished Ph.D. dissertation, University of Sussex.
Dench, A. & Evans, N. (1988) Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8(1):1–47. [aNE]
Dixon, R. M. W. (1972) *The Dyirbal language of North Queensland*. Cambridge University Press. [aNE]
(1977) The syntactic development of Australian languages. In: *Mechanisms of syntactic change*, ed. C. Li, pp. 365–418. University of Texas Press. [aNE]
Doupé, A. J. & Kuhl, P. K. (1999) Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 22:567–631. [BMe]
Downing, L. (1998) On the prosodic misalignment of onsetless syllables. *Natural Language and Linguistic Theory* 16:1–52. [AN]
Dowsett-Lemaire, F. (1979) The imitation range of the song of the Marsh Warbler, *Acrocephalus palustris*, with special reference to imitations of African birds. *Ibis* 121:453–68. [BMe]
Dryer, M. (1998) Why statistical universals are better than absolute universals. *Chicago Linguistic Society: The Panels* 33:123–45. [aNE]
(2005) Order of subject, object, and verb. In: *The world atlas of language structures*, ed. M. Haspelmath, M. Dryer, D. Gil & B. Comrie, pp. 386–89. Oxford University Press. [MCB]
(2009) On the order of demonstrative, numeral, adjective and noun: An alternative to Cinque. Public Lecture, Max Planck Institute for Evolutionary Anthropology, Department of Linguistics, and University of Leipzig, Institute of Linguistics, May 19, 2009. [rNE]
Dubois, J. W. (1987) The discourse basis of ergativity. *Language* 63(4):805–55. [aNE]
Dunn, M., Greenhill, S. J., Levinson, S. C. & Gray, R. D. (in preparation) Phylogenetic trees reveal lineage specific trends in the evolved structure of language. [rNE]
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):710–59. [aNE]
Dunn, M., Terrill, A., Reesink, G., Foley, R. & Levinson, S. C. (2005) Structural phylogenetics and the reconstruction of ancient language history. *Science* 309:2072–75. [aNE]
Durham, W. H. (1991) *Coevolution: Genes, culture and human diversity*. Stanford University Press. [arNE]
Durie, M. (1985) *A grammar of Acehnese, on the basis of a dialect of North Aceh*. Foris. [aNE]
Edelman, S. & Christiansen, M. H. (2003) How seriously should we take minimalist syntax? *Trends in Cognitive Sciences* 7.2:60–61. [rNE]
Edelman, S. & Solan, Z. (in press) Translation using an automatically inferred structured language model. [HW]
Edelman, S., Solan, Z., Horn, D. & Ruppin, E. (2004) Bridging computational, formal and psycholinguistic approaches to language. In: *Proceedings of the 26th Conference of the Cognitive Science Society*, ed. K. Forbus, D. Gentner & T. Regier, pp. 345–50. Erlbaum. [HW]
(2005) Learning syntactic constructions from raw corpora. In: *Proceedings of the 29th annual Boston University Conference on language development*, ed. A. Brugos, M. R. Clark-Cotton, & S. Ha, pp. 180–91. Cascadilla Press. [HW]
Edelman, S. & Waterfall, H. R. (2007) Behavioral and computational aspects of language and its acquisition. *Physics of Life Reviews* 4:253–77. [HW]
Elman, J. L. (1990) Finding structure in time. *Cognitive Science* 14:179–211. [BMc]
Elman, J. L., Bates, E., Johnson, M. H. & Karmiloff-Smith, A. (1996) *Rethinking innateness: A connectionist perspective on development*. MIT Press. [aNE]
Emmorey, K. (2002) *Language, cognition, and the brain: Insights from sign language research*. Erlbaum. [aNE]
Enfield, N. J. (2004) Adjectives in Lao. In: *Adjective classes: A cross-linguistic typology*, ed. R. M. W. Dixon & A. Aikhenvald, pp. 323–47. Oxford University Press. [aNE]
Enfield, N. J. & Levinson, S. C. (2006) *Roots of human sociality: Cognition, culture and interaction.* Berg. [aNE]
England, N. C. (1983) *A grammar of Mam, a Mayan language.* University of Texas Press. [DP]
(2001) *Introducción a la gramática de los idiomas Mayas.* Cholsomaj. [aNE]
(2004) Adjectives in Mam. In: *Adjective classes: A cross-linguistic typology*, ed. R. M. W. Dixon, pp. 125–46. Oxford University Press. [aNE]
Evans, N. (1995a) *A grammar of Kayardild*. Mouton de Gruyter. [MCB, aNE]
(1995b) Current Issues in Australian phonology. In: *Handbook of phonological theory*, ed. J. Goldsmith, pp. 723–61. Blackwell. [rNE]
(1995c) 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. [aNE]
(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. [aNE]
(2003a) *Bininj Gun-wok: A pan-dialectal grammar of Mayali, Kunwinjku and Kune. 2 vols.* Pacific Linguistics. [arNE]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

(2003b) Context, culture and structuration in the languages of Australia. *Annual Review of Anthropology* 32:13–40. [aNE]
(2004) Experiencer objects in Iwaidjan languages. In: *Non-nominative subjects, vol. 1*, ed. B. Peri & S. Karumuri Venkata, pp. 169–92. John Benjamins. [rNE]
Evans, N. & Osada, T. (2005) Mundari: The myth of a language without word classes. *Linguistic Typology* 9(3):351–90. [arNE]
Evans, N. & Sasse, H.-J., eds. (2002) *Problems of polysynthesis*. Akademie. [aNE]
Everett, D. L. (2005) Cultural constraints on grammar and cognition in Pirahã. Another look at the design features of human language. *Current Anthropology* 46:621–46. [aNE, BMe]
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. 254–71. Oxford University Press. [aNE]
Fisher S. E. & Marcus, G. F. (2006) The eloquent ape: Genes, brains and the evolution of language. *Nature Reviews Genetics* 7:9–20. [arNE]
Fisher, S. & Scharff, C. (2009) FOXP2 as a molecular window into speech and language. *Trends in Genetics* 25(4):166–77. [rNE]
Fitch, W. T. (2006) The biology and evolution of music: A comparative perspective. *Cognition* 100(1):173–215. [rNE]
(2008) Co-evolution of phylogeny and glossogeny: There is no "logical problem of language evolution." *Behavioral and Brain Sciences* 31(5):521–22. [BMe]
Fitch, W. T. & Hauser, M. D. (2004) Computational constraints on syntactic processing in a nonhuman primate. *Science* 303(5656):377–80. [arNE, DM, GKP]
Fitch, W. T., Hauser, M. D. & Chomsky, N. (2005) The evolution of the language faculty: Clarifications and implications. *Cognition* 97:179–210. [BMe, aNE]
Fodor, J. A. (1975) *The language of thought*. Thomas Y. Crowell. [aNE]
(1983) *The modularity of mind*. MIT Press. [aNE]
François, A. (2005) A typological overview of Mwotlap, an Oceanic language of Vanuatu. *Linguistic Typology* 9(1):115–46. [aNE]
Freidin, R. (1992) *Foundations of generative syntax*. MIT Press. [RF]
Freidin, R. & Quicoli, C. (1989) Zero-stimulation for parameter setting. *Behavioral and Brain Sciences* 12:338–39. [RF]
Friederici, A. (2004) Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Science* 8:245–47. [aNE, GKP]
Gathercole, S. E. & Baddeley, A. D. (1989) Development of vocabulary in children and short-term phonological memory. *Journal of Memory and Language* 28:200–13. [ELB]
Gazdar, G., Klein, E., Pullum, G. & Sag, I. (1985) *Generalized phrase structure grammar*. Blackwell. [aNE]
Gazdar, G. & Pullum, G. K. (1982) *Generalized phrase structure grammar: A theoretical synopsis*. Indiana Linguistics Club. [aNE]
Gelman, S. (2003) *The essentialist child: Origins of essentialism in everyday thought*. Oxford University Press. [AEG]
Gentner, D. (2003) Why we are so smart. In: *Language in mind*, ed. D. Gentner & S. Goldin-Meadow, pp. 195–236. MIT Press. [rNE]
Gentner, D. & Boroditsky, L. (2001) Individuation, relativity, and early word learning. In: *Language acquisition and conceptual development*, ed. M. Bowerman & S. C. Levinson, pp. 215–56. Cambridge University Press. [aNE]
Gentner, T. Q., Fenn, K. M., Margoliash, D. & Nusbaum, H. C. (2006) Recursive syntactic pattern learning by songbirds. *Nature* 440(7088):1204–1207. [BMc, DM]
Gil, D. (2001) Escaping eurocentrism. In: *Linguistic fieldwork*, ed. P. Newman & M. Ratliff, pp. 102–32. Cambridge University Press. [aNE]
Givón, T. (2008) *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-Cognition, evolution*. John Benjamins. [aNE]
Gleitman, L. (1990) The structural sources of verb meanings. *Language Acquisition* 1:3–55. [aNE]
Goddard, C. & Wierzbicka, A., eds. (2002) *Meaning and universal grammar – theory and empirical findings. 2 volumes*. John Benjamins. [rNE]
Goedemans, R. (1998) *Weightless segments*. Holland Academic Graphics. [AN]
Goffman, E. (1978) Response cries. *Language* 54(4):787–815. [SP]
Goldberg, A. E. (2006) *Constructions at work: The nature of generalization in language*. Oxford University Press. [AEG]
Goldberg, A. E. & Boyd, J. K. (2009) Learning what not to say: Categorization, preemption and discounting in a-adjectives. Unpublished manuscript, Princeton University. [AEG]
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. [aNE]
Goldsmith, J. A. (2007) Towards a new empiricism. In: *Recherches linguistiques à Vincennes, vol. 36*, ed. J. B. de Carvalho. Presses universitaires de Vincenne. [HW]
Goldstein, M. H., King, A. P. & West, M. J. (2003) Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences USA* 100(13):8030–35. [BMc]
Goldstein, M. H. & Schwade, J. A. (2008) Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science* 19:515–23. [HW]
Gordon, M. (2005) A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23:595–653. [AN]
Gordon, P. (2004) Numerical cognition without words: Evidence from Amazonia. *Science* 306(5695):496–99. [aNE]
Gray, R. D. & Atkinson, Q. D. (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426:435–39. [aNE]
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. [aNE, LR]
ed. (1963b) *Universals of language*. MIT Press. [aNE]
ed. (1966) *Language universals*. Mouton de Gruyter. [aNE]
(1969) Some methods of dynamic comparison in linguistics. In: *Substance and structure of language*, ed. J. Puhvel, pp. 147–203. University of California Press. [WC, BMe]
(1978a) Diachrony, synchrony and language universals. In: *Universals of human language, vol. 1: Method and theory*, ed. J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik, pp. 61–92. Stanford University Press. [WC]
ed. (1978b) Generalizations about numeral systems. In: *Universals of human language*, ed. J. H. Greenberg, pp. 249–96. Stanford University Press. [aNE]
(1979) Rethinking linguistics diachronically. *Language* 55:275–90. [WC]
ed. (1986) On being a linguistic anthropologist. *Annual Review of Anthropology* 15:1–24. [aNE]
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. [aNE]
Greenfield, P. M. (1991) Language, tools, and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14(4):531–51. [aNE]
Grice, H. P. (1967) Logic and conversation. William James Lectures. Unpublished manuscript, Harvard University, Cambridge, MA. [MHC]
Grimshaw, J. & Samek-Lodovici, V. (1998) Optimal subjects and subject universals. In: *Is the best good enough? Optimality and competition in syntax*, ed. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, pp. 193–219. MIT Press. [PS]
Gross, M. (1979) On the failure of generative grammar. *Language* 55:859–85. [WC]
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. [aNE]
Hagoort, P. (2005) On Broca, brain and binding: A new framework. *Trends in Cognitive Science* 9:416–23. [aNE]
Haiman, J. (1985) *Iconicity in syntax*. Cambridge University Press. [AEG]
Hale, K. L. (1964) Classification of northern Paman languages, Cape York Peninsula, Australia: A research report. *Oceanic Linguistics* 3(2):248–64. [AN]
(1973) Person marking in Warlpiri. In: *A Festschrift for Morris Halle*, ed. S. A. Anderson & P. Kiparsky, pp. 308–44. Holt, Rinehart, & Winston. [MTa]
(1982) The logic of Damin kinship terminology. In: *Languages of kinship in Aboriginal Australia*, ed. J. Heath, F. Merlan & A. Rumsey, pp. 31–37. Oceania Linguistic Monographs. [rNE]
(1983) Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1:5–47. [aNE]
(1997) Some observations on the contribution of local languages to linguistic science. *Lingua* 100:71–89. [DH]
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. 1430–51. Walter de Gruyter. [aNE]
Halle, M. (1970) Is Kabardian a vowel-less language? *Foundations of Language* 6:95–103. [aNE]
Hankamer, J. (1989) Morphological parsing and the lexicon. In: *Lexical representation and process*, ed. W. Marslen-Wilson, pp. 392–408. MIT Press. [aNE]
Harbour, D. (2006) Valence and atomic number. Unpublished manuscript, Queen Mary, London. Available at: http://webspace.qmul.ac.uk/dharbour/Harbour%20Valence%20Revised.pdf. [DP, DH]
(2007) *Morphosemantic number: From Kiowa noun classes to UG number features*. Springer. [DH]
(in press) Descriptive and explanatory markedness. *Morphology*. [DH]
Harman, G. & Kulkarni, S. (2007) *Reliable reasoning: Induction and statistical Learning theory*. MIT Press. [MHC]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Harnad, S. (1976) Induction, evolution and accountability. Discussion paper. *Annals of the New York Academy of Sciences* 280:58–60. [BMe]
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. [aNE]
Harris, Z. S. (1946) From morpheme to utterance. *Language* 22:161–83. [HW]
(1991) *A theory of language and information*. Clarendon. [HW]
Hart, B. & Risley, T. R. (1995) *Meaningful differences in the everyday experience of young American children*. Brookes. [ACC]
Haspelmath, M. (1993) *A grammar of Lezgian* [Mouton Grammar Library, 9]. Mouton de Gruyter. [aNE]
(1999) Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18(2):180–205. [aNE, AN]
(2007) Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11(1):119–32. [aNE]
(2008a) A frequentist explanation of some universals of reflexive marking. *Linguistic Discovery* 6.1:40–63. [MH]
(2008b) Parametric versus functional explanations of syntactic universals. In: *The limits of syntactic variation*, ed. T. Biberauer, pp. 75–107, Benjamins. [MH]
(2009) The indeterminacy of word segmentation and the nature of morphology and syntax. Plenary talk presented at the Morphology of the World's Languages conference, University of Leipzig, June 2009, Benjamins. [MH]
Haun, D. & Call, J. (2009) Great apes' capacities to recognize relational similarity. *Cognition* 110:147–59. [rNE]
Hauser, M. D. (1997) *The evolution of communication*. MIT Press/Bradford Books. [aNE]
Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) The faculty of language: What is it, who has it and how did it evolve? *Science* 298(5598):1569–79. [aNE, BMe, DM, DCP]
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. [aNE]
(1988) "It's my own invention: A comparative grammatical sketch of colonial Tzotzil" and grammatical annotations. In: *The great Tzotzil dictionary of Santo Domingo Zinacantán, with grammatical analysis and historical commentary*, ed. R. M. Laughlin & J. B. Haviland, pp. 79–121. (*Smithsonian Contributions to Anthropology*, No. 31). Smithsonian Institution Press. [rNE]
(1994) "Te xa setel xulem" [The buzzards were circling]: Categories of verbal roots in (Zinacantec) Tzotzil. *Linguistics* 32:691–741. [rNE, DP]
(submitted) "White-blossomed on bended knee": Linguistic mediations of nature and culture. Book chapter for *Festschrift for Terry Kaufman*, ed. R. M. Zavala & T. Smith-Stark. Available at: http://anthro.ucsd.edu/~jhaviland/Publications/BLOSSOMEdit.pfd [rNE]
Hawkins, J. A. (1999) Processing complexity and filler-gap dependencies across grammars. *Language* 75:244–85. [rNE]
(2004) *Efficiency and complexity in grammars*. Oxford University Press. [MH]
Hayes, B. (1989) Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253–306. [AN]
Hayes, B. & Steriade, D. (2004) A review of perceptual cues and cue robustness. In: *Phonetically based phonology*, ed. B. Hayes, R. M. Kirchner & D. Steriade, pp. 1–33. Cambridge University Press. [IB]
Hayes, B., Zuraw, K., Siptar, P. & Londe, Z. (submitted) Natural and unnatural constraints in Hungarian vowel harmony. [IB]
Helgason, P. & Ringen, C. (2008) Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36(4):607–28. [BMc]
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. John Benjamins. [aNE]
Hewitt, B. G. (1979) Aspects of verbal affixation in Abkhaz (Abžui dialect). *Transactions of the Philological Society* 77(1): 211–38. [rNE]
Himmelmann, N. (1997) *Deiktikon, Artikel, Nominalphrase, Zur Emergenz Syntaktischer Struktur*. Niemeyer. [aNE]
Hockett, C. F. (1960) The origin of speech. *Scientific American* 203:89–96. [rNE]
(1963) The problem of universals in language. In: *Universals of language*, ed. J. H. Greenberg, pp. 1–29. MIT Press. [aNE]
Hoff-Ginsberg, E. (1990) Maternal speech and the child's development of syntax: A further look. *Journal of Child Language* 17:85–99. [HW]
Höhle, B., Weissenborn, J., Kiefer, D. Schulz, A. & Schmitz, M. (2004) Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy* 5:341–53. [ELB]
Hollich, G. J., Hirsh-Pasek, K. & Golinkoff, R. M. (2000) Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development* 65(3), Serial No. 262. [ELB]
Holyoak, K. J. & Hummel, J. E. (2000) The proper treatment of symbols in a connectionist architecture. In: *Cognitive dynamics: Conceptual change in humans and machines*, ed. E. Dietrich & A. B. Markman, pp. 229–63. Erlbaum. [DCP]
Holyoak, K. J., Junn, E. N. & Billman, D. O. (1984) Development of analogical problem-solving skill. *Child Development* 55(6):2042–55. [DCP]
Horning, J. J. (1969) A study of grammatical inference. Doctoral dissertation, Computer Science Department, Stanford University. [BMe]
Hornstein, N., Nunes, J. & Grohmann, K. (2005) *Understanding minimalism*. Cambridge University Press. [aNE]
Huang, C.-R. (1993) Reverse long-distance dependency and functional uncertainty: The interpretation of Mandarin questions. In: *Language, information, and computing*, ed. C. Lee & B. M. Kang, pp. 111–20. Thaehaksa. [rNE]
Huang, C.-T. J. (1982) Logical relations in Chinese and the theory of grammar. Doctoral dissertation, Massachusetts Institute of Technology. Available at: http://dspace.mit.edu/handle/1721.1/15215. [DP, LR]
(1982/1998) *Logical relations in Chinese and the theory of grammar*. Garland. (Doctoral dissertation of 1982; published by Garland in 1988). [PS]
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. 329–38. Walter de Gruyter. [aNE]
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. 141–54. Oxford University Press. [aNE]
Hyman, L. (1985) *A theory of phonological weight*. (*Publications in Language Sciences* 19). Foris. [AN]
(2008) Universals in phonology? *The Linguistic Review* 25:83–187. [rNE]
Idsardi, W. J. (2006) A simple proof that optimality theory is computationally intractable. *Linguistic Inquiry* 37:271–75. [rNE]
Jackendoff, R. (1997) *The architecture of the language faculty*. MIT Press. [SP]
(2002) *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press. [aNE, SP]
(2003a) Reintegrating generative grammar (Précis of *Foundations of language*). *Behavioral and Brain Sciences* 26:651–65. [aNE]
(2003b) Toward better mutual understanding [Response to peer commentaries on 2003a]. *Behavorial and Brain Sciences* 26:695–702. [aNE]
Jacobsen, W. H. (1979) Noun and verb in Nootkan. In: *The Victoria Conference on northwestern languages, Victoria, British Columbia, November 4/5, 1976*, ed. B. S. Efrat, pp. 83–155. British Columbia Provincial Museum. [rNE]
Jakobson, R. (1962) *Selected writings I: Phonological studies*. Mouton de Gruyter. [aNE]
Jakobson, R. & Halle, M. (1956) *Fundamentals of language*. Mouton de Gruyter. [aNE]
Janik, V. M. & Slater, P. J. B. (1997) Vocal learning in mammals. *Advances in the Study of Behavior* 26:59–99. [BMe]
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. [arNE]
Johnson, K. (2004) Gold's theorem and cognitive science. *Philosophy of Science* 71:571–92. [BMe]
Joos, M., ed. (1966) *Readings in linguistics I*, 2nd edition. University of Chicago Press. [GKP]
Jusczyk, P. (1997) *The discovery of spoken language*. MIT Press. [ELB]
Karmiloff-Smith, A. (2006) The tortuous route from genes to behaviour: A neuroconstructivist approach. *Cognitive, Affective and Behavioural Neuroscience* 6(1):9–17. [AEG]
Kasher, A. (1991) *The Chomskyan turn*. Blackwell. [GKP]
Kay, P. & Kempton, W. (1984) What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65–79. [aNE]
Kazenin, K. I. (1994) Split syntactic ergativity: Toward an implicational hierarchy. *Sprachtypololgie und Universalienforschung* 47:78–98. [WC]
Keller, G. & Hahnloser, R. (2009) Neural processing of auditory feedback during vocal practice in a songbird. *Nature* 457:187–90. [BMc]
Kirby, S. (2002) Learning, bottlenecks and the evolution of recursive syntax. In: *Linguistic evolution through language acquisition: Formal and computational models*, ed. T. Briscoe, pp. 173–204. Cambridge University Press. [BMe]
(2007) The evolution of meaning-space structure through iterated learning. In: *Emergence of communication and language*, ed. C. Lyon, C. Nehaniv & A. Cangelosi, pp. 253–68. Springer. [MHC]
Klima, E. S. & Bellugi, U. (1979) *The signs of language*. Harvard University Press. [aNE]
Kluender, R. (1992) Deriving island constraints from principles of predication. In: *Island constraints*, ed. H. Goodluck & M. Rochmont, pp. 223–58. Kluwer. [rNE]
(1998) On the distinction between strong and weak islands: A processing perspective. In: *Syntax and semantics*, ed. P. Culicover & L. McNally, pp. 241–79. Academic Press. [rNE]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge 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. [aNE]

Koster, J. & May, R. (1982) On the constituency of infinitives. *Language* 58:116–43. [aNE]

Kratochvíl, F. (2007) A grammar of Abui. Doctoral dissertation, Leiden University. [rNE]

Kruspe, N. (2004) *A grammar of Semelai.* Cambridge University Press. [rNE]

Kuhl, P. K. (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. [aNE]

(2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5:831–43. [ELB, aNE]

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

Küntay, A. & Slobin, D. (1996) Listening to a Turkish mother: Some puzzles for acquisition. In: *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp*, ed. D. Slobin & J. Gerhardt, pp. 265–86. Erlbaum. [HW]

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

(1994) *Principles of linguistic change, vol. 1: Internal factors.* Blackwell. [WC]

(2001) *Principles of linguistic change, vol. 2: Social factors.* Blackwell. [WC]

Ladefoged, P. & Maddieson, I. (1996) *The sounds of the world's languages.* Blackwell. [aNE]

Lakoff, G. (1987) *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press. [AEG]

Laland, K. N., Odling-Smee, J. & Feldman, M. W. (1999) Evolutionary consequences of niche construction and their implications for ecology. *Proceedings of the National Academy of Sciences USA* 96:10242–47. [rNE]

(2000) Niche construction, biological evolution and cultural change. *Behavioral and Brain Science* 23:131–75. [aNE]

Landau, B. & Jackendoff, R. (1993) "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16:217–65. [aNE]

Lander, E. S. (1994) Genetic dissection of complex traits. *Science* 265:2037–2048. [AEG]

Langacker, W. (2008) *Cognitive grammar: A basic introduction.* Oxford University Press. [WC]

Lappin, S. & Shieber, S. M. (2007) Machine learning theory and practice as a source of insight into universal grammar. *Journal of Linguistics* 43:1–34. [BMe]

Lasnik, H. (1989) *Essays on anaphora.* Kluwer. [LR]

Lasnik, H. & Saito, M. (1992) *Move alpha.* MIT Press. [LR]

Laughren, M. (2002) Syntactic constraints in a "free word order language." In: *Language universals and variation*, ed. M. Amberber, pp. 83–130. Ablex. [DP]

Lazareva, O. F. & Wasserman, E. A. (2008) Categories and concepts in animals. In: *Learning and memory: A comprehensive reference. Vol. 1: Learning theory and behavior*, ed. R. Menzel, pp. 197–226. (Series editor: J. Byrne). Elsevier. [BMc]

Legate, J. A. (2001) The configurational structure of a nonconfigurational language. In: *Linguistic Variation Yearbook, vol. 1*, ed. P. Pica, pp. 61–104. John Benjamins. Available at: http://www.ling.upenn.edu/~jlegate/livy.1.legate.pdf. [DP]

(2002) Warlpiri: Theoretical implications. Doctoral dissertation, Massachusetts Institute of Technology. Available at: http://www.ling.upenn.edu/~jlegate/main.pdf. [DP]

Legendre, G., Smolensky, P. & Wilson, C. (1998) When is less more? Faithfulness and minimal links in *wh*-chains. In: *Is the best good enough? Optimality and competition in syntax*, ed. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, pp. 249–89. MIT Press. [PS]

Lehmann, C. (1995) *Thoughts on grammaticalization, 2nd revised edition.* Lincom Europa. [BMe]

Lenneberg, E. H. (1964) The capacity for language acquisition. In: *The structure of language*, ed. J. A. Fodor & J. J. Katz, pp. 579–603. Prentice Hall. [PS]

(1967) *Biological foundations of language.* Wiley. [aNE]

Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B. & Kail, R. V. (2007) Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research* 50:408–28. [ELB]

Levelt, W. J. M. (2008) *Formal grammars in linguistics and psycholinguistics, vol.1-3.* Benjamins. [aNE]

Levin, B. (1993) *English verb classes and alternations: A preliminary investigation*. University of Chicago Press. [SP]

Levinson, S. C. (1987) Pragmatics and the grammar of anaphora. *Journal of Linguistics* 23:379–434. [rNE]

(2000) *Presumptive meanings: The theory of generalized conversational implicature.* MIT Press. [arNE]

(2003) *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press. [aNE]

(under review) Syntactic ergativity in Yélî Dnye, the Papuan language of Rossel Island, and its implications for typology. [aNE]

Levinson, S. C. & Jaisson, P., eds. (2006) *Evolution and culture. A Fyssen Foundation symposium*. MIT Press. [aNE]

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. [aNE]

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

Li, P. & Gleitman, L. R. (2002) Turning the tables: Language and spatial reasoning. *Cognition* 83(3):265–94. [aNE, SP]

Li, P., Hai Tan, L., Bates, E. & Tzeng, O. (2006) *The handbook of East Asian psycholinguistics, vol. 1: Chinese*. Cambridge University Press. [aNE]

Lickliter, R. & Honeycutt, H. (2003) Developmental dynamics: Towards a biologically plausible evolutionary psychology. *Psychological Bulletin* 129:819–35. [BMc]

Lieberman, P. (2006) *Toward an evolutionary biology of language*. Belknap/Harvard. [rNE]

Lieven, E. V. M., Behrens, H., Speares, J. & Tomasello, M. (2003) Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30:333–70. [ELB]

Liljencrants, J. & Lindblom, B. (1972) Numerical simulations of vowel quality systems: The role of perceptual contrast. *Language* 48:839–62. [rNE]

Lipkind, W. (1945) *Winnebago grammar.* King's Crown. [DH]

Loewenstein, J. & Gentner, D. (2005) Relational language and the development of relational mapping. *Cognitive Psychology* 50(4):315–53. [DCP]

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

Lucy, J. & Gaskins, S. (2001) Grammatical categories and the development of classification: it comparative approach. In: *Language acquisition and conceptual development*, ed. M. Bowerman & S. Levinson, pp. 257–83. Cambridge University Press. [aNE]

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* 261:117–21. [aNE]

MacLarnon, A. & Hewitt, G. (2004) Increased breathing control: another factor in the evolution of human language. *Evolutionary Anthropology* 13:181–97. [aNE]

MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, C. R., Suckling, J., Calvert, G. A. & Brammer, M. J. (2002) Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain* 125:1583–93. [aNE]

MacWhinney, B., ed. (1999) *The mechanisms of language acquisition*. Erlbaum. [ELB]

(2000) *The CHILDES Project: Tools for analyzing talk. Vol. 1: Transcription format and programs. Vol. 2: The Database.* Erlbaum. [HW]

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

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

(1984) *Patterns of sounds.* Cambridge University Press. [aNE]

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

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. [aNE]

Majid, A., Bowerman, M., Kita, S., Haun, D. B. & Levinson, S. C. (2004) Can language restructure cognition? The case for space. *Trends in Cognitive Science* 8(3):108–14. [aNE, DCP]

Manzini, M. R. & Wexler, K. (1986) Parameters, binding theory and learnability. *Linguistic Inquiry* 18:413–44. [LR]

Marantz, A. (1984) *On the nature of grammatical relations*. MIT Press. [MCB]

Marcus, G. F. (2001) *The algebraic mind: Reflections on connectionism and cognitive science.* MIT Press. [SP]

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 111–41. Erlbaum. [aNE]

Marler, P. (1970) Bird song and speech development: Could there be parallels? *American Scientist* 58:669–73. [BMe]

Marler, P. & Tamura, M. (1964) Culturally transmitted patterns of vocal behavior in sparrows. *Science* 146 (3650):1483–86. [DM]

Marsaja, I. G. (2008) *Desa kolok – A deaf village and its sign language in Bali, Indonesia*. Ishara Press. [aNE]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Master, A. (1946) The zero negative in Dravidian. *Transactions of the Philological Society* 1946:137–55. [aNE]
Matthews, P. H. (1981) *Syntax*. Cambridge University Press. [aNE]
(2007) *Syntactic relations: A critical survey.* Cambridge University Press. [aNE]
Mayr, E. (1975) *Evolution and the diversity of life*. Harvard University Press. [AEG]
McCarthy, J. J. (2002) *A thematic guide to optimality theory* (*Research Surveys in Linguistics*). Cambridge University Press. [rNE]
McCarthy, J. & Prince, A. (1986) *Prosodic morphology 1986. Report No. RuCCS-TR-32.* Rutgers University Center for Cognitive Science. Available at: http://ruccs.rutgers.edu/pub/papers/pm86all.pdf. [AN]
McCloskey, J. (1991) Clause structure, ellipsis and proper government in Irish. *Lingua* 85:259–302. [MCB]
McMahon, A. & McMahon, R. (2006) *Language classification by numbers.* Oxford University Press. [aNE]
McMurray, B., Aslin, R. N. & Toscano, J. (2009) Statistical learning of phonetic categories: Computational insights and limitations. *Developmental Science* 12(3):369–79. [BMc]
Meir, I., Sandler, W., Padden, C. & Aronoff, M. (in press) Emerging sign languages. In: *Oxford handbook of deaf studies, language, and education, vol. 2*, ed. M. Marschark & P. Spencer. Oxford University Press. [arNE]
Melčuk, I. (1988) *Dependency syntax: Theory and practice.* SUNY Press. [aNE]
Merker, B. (2005) The conformal motive in birdsong, music and language: An introduction. *Annals of the New York Academy of Sciences* 1060:17–28. [BMe]
(in press) The vocal learning constellation: Imitation, ritual culture, encephalization. In: *Music, language and human evolution*, ed. N. Bannan & S. Mithen. Oxford University Press. [BMe]
Merker, B. & Okanoya, K. (2007) The natural history of human language: Bridging the gaps without magic. In: *Emergence of communication and language*, ed. C. Lyon, L. Nehaniv & A. Cangelosi, pp. 403–20. Springer-Verlag. [BMe]
Mielke, J. (2007) *The emergence of distinctive features*. Oxford University Press. [aNE]
Milroy, J. (1992) *Linguistic variation and change.* Blackwell. [WC]
Mintz, T. H. (2006) Finding the verbs: Distributional clues to categories available to young learners. In: *Action meets words: How children learn verbs*, ed. K. Hirsh-Pasek & R. M. Golinkoff, pp. 31–63. Oxford University Press. [ELB]
Mithun, M. (1984) How to avoid subordination. *Berkeley Linguistic Society* 10:493–509. [aNE]
(1999) *The languages of native North America* (Cambridge Language Surveys). Cambridge University Press. [arNE, DP]
Moerk, E. L. (1992) *First language: Taught and learned*. Brookes. [ACC]
Molotsi, K. J. (1993) The characteristics of Southern Sotho ideophones. Master's thesis, University of Stellenbosch. [rNE]
Montgomery, J. W., Evans, J. L. & Gillam, R. B. (2009) Relation of auditory attention and complex sentence comprehension in children with specific language impairment: A preliminary study. *Applied Psycholinguistics* 30:123–51. [ELB]
Moreton, E. (2008) Analytic bias and phonological typology. *Phonology* 25(1):83–127. [IB, AN]
Moro, A. (2008) *The boundaries of Babel.* MIT Press. [LR]
Morsanyi, K. & Holyoak, K. J. (in press) Analogical reasoning ability in autistic and typically developing children. *Developmental Science.* [DCP]
Mufwene, S. (2001) *The ecology of language evolution.* Cambridge University Press. [WC]
(2008) *Language evolution: Contact, competition and change.* Continuum. [WC]
Müller, R.-A. (2009) Language universals in the brain: How linguistic are they? In: *Language universals*, ed. M. Christiansen, C. Collins & S. Edelman, pp. 224–52. Oxford University Press. [rNE]
Mycock, L. (2006) The typology of constituent questions: A lexical-functional grammar analysis of "Wh"-questions. Unpublished Dissertation. University of Manchester. [rNE]
Nakayama, R., Mazuka, R. & Shirai, Y. (2006) *Handbook of East Asian psycholinguistics, vol. 2: Japanese.* Cambridge University Press. [aNE]
Nettle, D. (1999) *Linguistic diversity*. Oxford University Press. [aNE]
Newell, A. (1980) Physical symbol systems. *Cognitive Science* 4:135–83. [DCP]
Newmeyer, F. J. (1986) *Linguistic theory in America, 2nd edition.* Academic Press. [aNE]
(2004) Typological evidence and universal grammar. *Studies in Language* 28(3):527–48. [arNE]
(2005) *Possible and probable languages: A generative perspective on linguistic typology*. Oxford University Press. [arNE]
Newport, E. L., Hauser, M. D., Spaepen, G. & Aslin, R. N. (2004) Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49:85–117. [BMc]
Nichols, J. (1992) *Language diversity in space and time*. University of Chicago Press. [aNE]
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. [aNE]
Nordlinger, R. (1998) *Constructive case. Evidence from Australian languages*. CSLI. [aNE]
(2006) Spearing the Emu drinking: Subordination and the adjoined relative clause in Wambaya. *Australian Journal of Linguistics* 26:5–29. [MTa]
Nordlinger, R. & Sadler, L. (2004) Nominal tense in crosslinguistic perspective. *Language* 80:776–806. [aNE]
Norman, J. (1988) *Chinese*. Cambridge University Press. [aNE]
Nottebohm, F. (1975) A zoologist's view of some language phenomena, with particular emphasis on vocal learning. In: *Foundations of language development*, ed. E. H. Lenneberg & E. Lenneberg, pp. 61–103. Academic Press. [BMe]
Noyer, R. (1992) Features, positions and affixes in autonomous morphological structure. Doctoral dissertation, Massachusetts Institute of Technology. [DH]
Odling-Smee, F. J., Laland, K. N. & Feldman, M. W. (2003) *Niche construction: The neglected process in evolution.* Princeton University Press. [aNE]
O'Donnell, T., Hauser, M. & Fitch, W. T. (2005) Using mathematical models of language experimentally. *Trends in Cognitive Science* 9:284–89. [aNE]
Okanoya, K. & Merker, B. (2007) Neural substrates for string-context mutual segmentation: A path to human language. In: *Emergence of communication and language*, ed. C. Lyon, L. Nehaniv & A. Cangelosi, pp. 421–34. Springer-Verlag. [BMe]
Onnis, L., Waterfall, H. R. & Edelman, S. (2008) Learn locally, act globally: Learning language from variation set cues. *Cognition* 109:423–30. [HW]
Osada, T. (1992) *A reference grammar of Mundari.* Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. [arNE, DP]
Padden, C. & Perlmutter, D. (1987) American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory* 5:335–75. [aNE]
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. [aNE]
Pagel, M., Atkinson, Q. D. & Meade, A. (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449:717–20. [aNE]
Papassotiropoulos, A., Stephan, D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V., Huynh, K.-D., Brunner, F., Corneveaux, J., Osborne, D., Wollmer, M. A., Aerni, A., Coluccia, D., Hänggi, J., Mondadori, C. R. A., Buchmann, A., Reiman, E. M., Caselli, R. J., Henke, K. & de Quervain, D. J.-F. (2006) Common Kibra alleles are associated with human memory performance. *Science* 314(5798):475–78. [rNE]
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. 239–46. World Scientific Press. [aNE]
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. [aNE]
Partee, B., ter Meulen, A. & Wall, R. (1990) *Mathematical methods in linguistics*. Kluwer. [aNE]
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. [aNE]
Pawley, A., Gi, S. P., Majnep, I. S. & Kias, J. (2000) Hunger acts on me: The grammar and semantics of bodily and mental process expressions in Kalam. In: *Grammatical analysis: Morphology, syntax and semantics: Studies in honor of Stan Starosta*, ed. V. P. De Guzman & B. W. Bender, pp. 153–85. University of Hawaii Press. [rNE]
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. 233–45. Chicago Linguistic Society. [aNE]
Penn, D. C., Holyoak, K. J. & Povinelli, D. J. (2008) Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences* 31(2):109–78. [DCP]
Penn, D. C. & Povinelli, D. J. (2007a) Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology* 58:97–118. [DCP]
(2007b) On the lack of evidence that non-human animals possess anything remotely resembling a "theory of mind." *Philosophical Transactions of the Royal Society B* 362:731–44. [DCP]
(in press) The comparative delusion: The "behavioristic"/ "mentalistic" dichotomy in comparative theory of mind research. In: *Oxford handbook of*

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

philosophy and cognitive science, ed. R. Samuels & S. P. Stich. Oxford University Press. [DCP]
Peperkamp, S., Le Calvez, R., Nadal, J. P. & Dupoux, E. (2006) The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101(B):31–41. [PS]
Perniss, P. M., Pfau, R. & Steinbach, M., eds. (2008) *Visible variation: Comparative studies on sign language structure. Trends in Linguistics 188.* Mouton de Gruyter. [aNE]
Perniss, P. M. & Zeshan, U., eds. (2008) *Possessive and existential constructions in sign languages. Sign Language Typology Series No. 2.* Ishara Press. [aNE]
Pesetsky, D. (2000) *Phrasal movement and its kin.* MIT Press. [LR]
Piattelli-Palmarini, M., ed. (1980) *Language and learning: The debate between Jean Piaget and Noam Chomsky.* Harvard University Press. [BMe]
Pierrehumbert, J. (2000) What people know about the sounds of language. *Linguistic Sciences* 29:111–20. [aNE]
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/Oxford University Press. [arNE]
Pinker, S. (1989) *Learnability and cognition: The acquisition of argument structure.* MIT Press. [SP]
(1994) *The language instinct.* W. Morrow. [aNE, MH, DCP]
(2007) *The stuff of thought: Language as a window into human nature*. Viking. [SP]
Pinker, S. & Bloom, P. (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–26. [aNE]
Pinker, S. & Jackendoff, R. (2005) The faculty of language: What's special about it? *Cognition* 95:201–36. [aNE, SP]
Pinker, S. & Ullman, M. T. (2002a) Combination and structure, not gradedness, is the issue. *Trends in Cognitive Sciences* 6(11):472–74. [SP]
(2002b) The past and future of the past tense. *Trends in Cognitive Sciences* 6(11):456–63. [SP]
(2003) Beyond one model per phenomenon. *Trends in Cognitive Science* 7(3):108–109. [SP]
Plank, F. (2001) Typology by the end of the 18th century. In: *History of the language sciences, vol. 2*, ed. S. Auroux, pp. 1399–1414. Mouton de Gruyter. [MH]
Pomerantz, J. R. & Kubovy, M. (1986) Theoretical approaches to perceptual organization: Simplicity and likelihood principles. In: *Handbook of perception and human performance, vol. 2: Cognitive processes and performance*, ed. K. R. Boff, L. Kaufman & J. P. Thomas, pp. 36:1–46. Wiley. [MHC]
Port, R. & Leary, A. (2005) Against formal phonology. *Language* 81:927–64. [aNE]
Postal, P. (1970) The method of universal grammar. In: *On method in linguistics*, ed. P. Garvin, pp. 113–31. Mouton de Gruyter. [aNE]
Povinelli, D. J. (2000) *Folk physics for apes: The chimpanzee's theory of how the world works.* Oxford University Press. [DCP]
Povinelli, D. J. & Vonk, J. (2003) Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences* 7(4):157–60. [DCP]
Prince, A. & Smolensky, P. (1993/2004) *Optimality theory: Constraint interaction in generative grammar*. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Rutgers Optimality Archive 537. [Revised version published by Blackwell, 2004.] [IB, PS]
(1997) Optimality: From neural networks to universal grammar. *Science* 275(5306):1604–10. [PS]
Pullum, G. K. & Rogers, J. (2006) *Animal pattern learning experiments: Some mathematical background.* Unpublished ms, University of Edinburgh. Available at: http://ling.ed.ac.uk/~gpullum/MonkeyMath.pdf. [GKP]
Pullum, G. K. & Scholz, B. C. (2002) Empirical assessment of stimulus poverty arguments. *Linguistic Review* 19:9–50. [MHC]
(2007) Systematicity and natural language syntax. *Croatian Journal of Philosophy* 7(21):375–402. [GKP]
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. [aNE]
Ramamurti, G. V. (1931) *A manual of the So:ra (or Savara) language.* Government Press. [rNE]
Reali, F. & Christiansen, M. H. (2005) Uncovering the richness of the stimulus: Structure dependence and indirect statistical evidence. *Cognitive Science* 29:1007–28. [MHC]
Reddy, M. J. (1979) The conduit metaphor – a case of frame conflict in our language about language. In: *Metaphor and thought*, ed. A. Ortony, pp. 284–324. Cambridge University Press. [ACC]
Reesink, G., Dunn, M. & Singer, R. (under review) Explaining the linguistic diversity of Sahul using population models. *PLoS*. [aNE]
Reinhart, T. (1976) The syntactic domain of anaphora. Doctoral dissertation, Massachusetts Institute of Technology. [LR]
Reuer, V. (2004) Book review of Falk, Yehuda N., *Lexical-functional grammar – an introduction to parallel constraint-based syntax. Lecture Notes No. 126 (CSLI-LN).* Center for the Study of Language and Information, Stanford, 2001, xv + 237 pages. *Machine Translation* 18.4:359–64. [rNE]
Rice, M. & Wexler, K. (1996) Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research* 39:1239–57. [ELB]
Rizzi, L. (1978/1982) Violations of the *wh*-island constraint in Italian and the subjacency condition. In: *Montreal Working Papers in Linguistics, vol. 11*, ed. C. Dubisson, D. Lightfoot & Y. C. Morin, pp. 155–90. Reprinted in L. Rizzi (1982) *Issues in Italian syntax*, pp. 49–76. Foris. [RF, LR]
(1990) *Relativized minimality.* MIT Press. [LR]
(1997) The fine structure of the left periphery. In: *Elements of grammar*, ed. L. Haegeman, pp. 281–337. Kluwer Academic. [DP]
(2006) Grammatically-based target-inconsistencies in child language. In: *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition – North America (GALANA)*, ed. K. U. Deen, J. Nomura, B. Schulz & B. D. Schwartz. (UCONN/MIT Working Papers in Linguistics). MIT Press. [LR]
Rogers, J. & Pullum, G. K. (2007) Aural pattern recognition experiments and the subregular hierarchy. Paper presented at the Mathematics of Language 10 Conference, UCLA, July 2007. Available at: http://ling.ed.ac.uk/~gpullum/MoL10paper.pdf. [GKP]
Ross, J. R. (1967) Constraints on variables in syntax. Unpublished doctoral dissertation, MIT. Published as *Infinite syntax*. Ablex. 1986. [RF]
Rost, G. & McMurray, B. (2009) Speaker variability augments phonological processing in early word learning. *Developmental Science* 12(2):339–49. [BMc]
Saffran, J. R., Aslin, R. N. & Newport, E. (1996) Statistical learning by 8-month-old infants. *Science* 274(5294):1926–28. [ELB, BMc]
Sag, I., Hofmeister, P. & Snider, N. (2007) Processing complexity in subjacency violations: The complex noun phrase constraint. *Chicago Linguistics Society* 43(1):219–29. [rNE]
Sandler, W. (1993) A sonority cycle in American Sign language. *Phonology* 10:242–79. [IB]
(2009) Symbiotic symbolization by hand and mouth in sign language. *Semiotica* 174(1/4):241–75. [arNE]
Sandler, W., Aronoff, M., Meir, I. & Padden, C. (2009) The gradual emergence of phonological form in a new language. Master's thesis, University of Haifa, State University of New York at Stony Brook, and University of California, San Diego. [rNE]
Sandler, W. & Lillo-Martin, D. C. (2006) *Sign language and linguistic universals*. Cambridge University Press. [IB, rNE]
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 USA* 102(7):2661–65. [aNE]
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. [aNE]
(1985) Parts-of-speech systems. In: *Language typology and syntactic description, vol. 1: Clause structure*, ed. T. Shopen, pp. 3–61. Cambridge University Press. [rNE]
Schultze-Berndt, E. (2000) Simple and complex verbs in Jaminjung: A study of event categorization in an Australian language. Doctoral dissertation, Radboud University, MPI Series in Psycholinguistics. [aNE]
Schwager, W. & Zeshan, U. (2008) Word classes in sign languages – criteria and classifications. *Studies in Language* 32(3):509–45. [aNE]
Senghas, A., Kita, S. & Özyürek, A. (2004) Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305(5691):1779–82. [aNE]
Skinner, B. F. (1957) *Verbal behavior.* Appleton-Century-Crofts. [BMc]
Slobin, D. I., ed. (1985a) *The cross linguistic study of language acquisition, vol. 1.* Erlbaum. [ELB]
(1985b) *The cross linguistic study of language acquisition, vol. 2.* Erlbaum. [ELB]
(1992) *The cross linguistic study of language acquisition, vol. 3.* Erlbaum. [ELB]
(1997a) *The crosslinguistic study of language acquisition, vol. 4.* Erlbaum. [ELB, aNE]
(1997b) *The cross linguistic study of language acquisition, vol. 5.* Erlbaum. [ELB]
Smith, J. L. (2005) *Phonological augmentation in prominent positions.* Routledge. [IB]
Smith, L. B. (1999) Children's noun learning: How general learning processes make specialized learning mechanisms. In: *The emergence of language*, ed. B. MacWhinney, pp. 277–303. Erlbaum. [ELB]
Smith, N. (1999) *Chomsky: Ideas and ideals*. Cambridge University Press. [GKP]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press

Smolensky, P. (2006) Optimality in phonology. II: Markedness, feature domains, and local constraint conjunction. In: *The harmonic mind: From neural computation to optimality-theoretic grammar, vol. 2: Linguistic and philosophical implications*, ed. P. Smolensky & G. Legendre, pp. 27–160. MIT Press. [IB]
Smolensky, P. & Legendre, G. (2006) *The harmonic mind. vol. 2.* MIT Press. [PS]
Soderstrom, M., Mathis, D. & Smolensky, P. (2006) Abstract genomic encoding of universal grammar in optimality theory. In: *The harmonic mind, vol. 2*, ed. P. Smolensky & G. Legendre, pp. 403–71. MIT Press. [PS]
Solan, Z., Horn, D., Ruppin, E. & Edelman, S. (2005) Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences USA* 102:11629–34. [HW]
Solan, Z., Ruppin, E., Horn, D. & Edelman, S. (2003) Unsupervised efficient learning and representation of language structure. In: *Proceedings of the 25th conference of the Cognitive Science Society*, ed. R. Alterman & D. Kirsh, pp. 1106–11. Erlbaum. [HW]
Sommer, B. (1970) An Australian language without CV syllables. *International Journal of American Linguistics* 36:57–58. [AN]
(1981) The shape of Kunjen syllables. In: *Phonology in the 1980's*, ed. D. L. Goyvaerts, pp. 231–44. E. Story-Scientia. [AN]
Starke, M. (2001) Move dissolves into merge. Doctoral dissertation, University of Geneva. [LR]
Steels, L. & Belpaeme, T. (2005) Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and Brain Sciences* 28:469–529. [aNE]
Stivers, T., Enfield, N., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E. & Levinson, S. C. (2009) Universals and cultural variation in turn taking in conversation. *Proceedings of the National Academy of Sciences USA* 106(26):10587–92. [rNE]
Swadesh, M. (1939) Nootka internal syntax. *International Journal of American Linguistics* 9:78–102. [MTa]
Szabolcsi, A. (2006) Strong vs. weak islands. In: *The Blackwell companion to syntax, vol. 4*, ed. M. Everaert & H. van Riemsdijk, pp. 479–532. Blackwell. [LR]
Talmy, L. (2000) *Towards a cognitive semantics*. MIT Press. [aNE]
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* 307:580–84. [aNE]
Terrace, H. S. (1975) Evidence of the innate basis of the hue dimension in the duckling. *Journal of the Experimental Analysis of Behavior* 24:79–87. [ACC]
Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S. & Weniger, D. (2002) Neural correlates for the acquisition of natural language syntax. *NeuroImage* 17:700–709. [LR]
Thiessen, E. D. (2007) The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language* 56(1):16–34. [BMc]
(2009) Statistical learning. In: *The Cambridge handbook of child language*, ed. E. L. Bavin, pp. 35–50. Cambridge University Press. [ELB]
Thomas, D. (1955) Three analyses of the Ilocano pronoun system. *Word* 11:204–208. [DH]
Thornton, R. (2008) Why continuity. *Natural Language and Linguistic Theory* 26(1):107–46. [LR]
Tomasello, M. (1995) Language is not an instinct. *Cognitive Development* 10:131–56. [aNE]
ed. (1998) *The new psychology of language: Cognitive and functional approaches to language structure.* Erlbaum. [WC]
(2000) *The cultural origins of human cognition.* Harvard University Press. [aNE]
(2003a) *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press. [ELB, MTo]
ed. (2003b) *The new psychology of language, vol. 2.* Erlbaum. [WC]
(2004) What kind of evidence could refute the UG hypothesis? Commentary on Wunderlich. *Studies in Language* 28(3):642–45. [AEG]
(2008) *The origins of human communication*. MIT Press. [arNE, DCP, MTo]
(2009) The usage-based theory of language acquisition. In: *The Cambridge handbook of child language*, ed. E .L. Bavin, pp. 69–87. Cambridge University Press. [ELB]
Topintzi, N. (2009) Onsets: An exploration of their suprasegmental and prosodic behaviour. Unpublished book manuscript, Aristotle University of Thessaloniki. [AN]
Tsai, W.-T. D. (1994) On economizing the theory of A-bar dependencies. Doctoral dissertation, Massachusetts Institute of Technology. [LR]
Valian, V. (2009) Innateness and learnability. In: *The Cambridge handbook of child language*, ed. E .L. Bavin, pp. 15–34. Cambridge University Press. [ELB]
Van Valin, R. D. (1998) The acquisition of WH-questions and the mechanisms of language acquisition. In: *The new psychology of language: Cognitive and functional approaches to language structure*, ed. M. Tomasello, pp. 221–49. Erlbaum. [rNE]
Van Valin, R. D. & LaPolla, R. (1997) *Syntax. Structure, meaning and function*. Cambridge University Press. [aNE]
Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., Alarcón, M., Oliver, P. L., Davies, K. E., Geschwind, D. H., Monaco, A. P. & Fisher, S. E. (2008) A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine* 359:2337–45. [arNE]
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* 134:501–20. [aNE]
Wall, J. D. & Kim S. K. (2007) Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genetics* 3(10):e175. [aNE]
Wasserman, E. A. & Zentall, T. R., eds. (2006) *Comparative cognition: Experimental explorations of animal intelligence.* Oxford University Press. [BMc]
Waterfall, H. R. (2006) A little change is a good thing: Feature theory, language acquisition and variation sets. Doctoral dissertation, University of Chicago. [HW]
(under review) A little change is a good thing: The relation of variation sets to children's noun, verb and verb-frame development. [HW]
Waterfall, H. R., Sandbank, B., Onnis, L. & Edelman, S. (under review) An empirical generative framework for computational modeling of language acquisition. [HW]
Werker J. & Curtin, S. (2005) PRIMPAR: A developmental framework of early speech processing. *Language Learning and Development* 1:197–234. [ELB]
Werker, J. F. & Tees, R. C. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7:49–63. [ELB]
(2005) Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* 46(3):233–51. [aNE]
West-Eberhard, M. J. (2003) *Developmental plasticity and evolution*. Oxford University Press. [rNE]
Whaley, L. (1997) *Introduction to typology.* Sage. [aNE]
Widmann, T. & Bakker, D. (2006) Does sampling matter: A test in replicability, concerning numerals. *Linguistic Typology* 10(1):83–95. [aNE]
Wierzbicka, A. (1982) Why can you *have a drink* when you can't **have an eat*? *Language* 58:753–99. [WC]
Wilson, C. (2006) Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30:945–82. [IB]
Wittgenstein, L. (1953) *Philosophical investigations*. Blackwell. [MHC]
Zeshan, U. (2002) Indo-Pakistani sign language grammar: A typological outline. *Sign Language Studies* 3(2):157–212. [aNE]
ed.(2006a) *Interrogative and negative constructions in sign languages* (Sign Language Typology Series No. 1). Ishara Press. [arNE]
(2006b) Sign languages of the world. In: *Encyclopedia of language and linguistics, 2nd editon*, ed. K. Brown, pp. 358–65. Elsevier. [arNE]
Zuidema, W. (2003) How the poverty of the stimulus solves the poverty of the stimulus. In: *Advances in neural information processing systems 15 (Proceedings of NIPS'02)*, ed. S. Becker, S. Thrun & K. Obermayer, pp. 51–58. MIT Press. [BMe]
Zuidema, W. & De Boer, B. (2009) The evolution of combinatorial phonology. *Journal of Phonetics* 37(2):125–44. [rNE]

https://doi.org/10.1017/S0140525X09990537 Published online by Cambridge University Press