

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 *mp^war* and *tanjk*, 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: *mp^wárem*, "is making," versus *i.k^wént*, "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 *tanjk* 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

^aCognitive Evolution Group, University of Louisiana at Lafayette, New Iberia, LA, 70560; ^bDepartment 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

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 explanandum 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

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

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}

^aDepartment of Psychology, Harvard University, Cambridge, MA 02138;

^bCenter for Cognitive Studies, Tufts University, Medford, MA 02155; and

^cSanta 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-style “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 lumpers’) 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

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 putty-nose 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:

$$a(c^*dc^*dc^*)^*a + b(c^*dc^*dc^*)^*b$$

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*, *accdeccddccccda*; *bdecccdcb*; *acdeccddccddda* . . .) 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

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

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

^aDepartment of Cognitive Science, Johns Hopkins University, Baltimore, MD 21218-2685; ^bLaboratoire 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.

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 architectural-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/elli/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 quuʔas-ʔi
work-3s:INDIC man-the
"The man is working."
2. quuʔas-maa mamuuk-ʔi
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 -ʔi, 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)aḥ
work-PAST-1s.INDIC
"I was working."

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. ʔahʔaaʔaʔ qahʃiʔ-ʔa ʔ 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)ah
person-PAST-1s.INDIC
“I was a man.”
7. ʔuunuʔ 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-vi *tjaa* kunka-(ng)ku *pukutjurrka* lhayi ngwi-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 ngiv-a lurrqbanyi 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

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

^aDepartment of Psychology, Cornell University, Ithaca, NY 14853, and Department of Psychology, University of Chicago, Chicago, IL 60637;

^bDepartment 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

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 이빠 누가 제일 이빠 지원이 제일 이빠 맞아
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

^aDepartment of Linguistics, Research School of Asian and Pacific Studies, Australian National University, ACT 0200, Australia; ^bMax 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

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

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 ground 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” (**Nejns**).

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 **Nejns’** 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

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 stochastic 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

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 **Neivins** 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 **Neivins** 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 \pm 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 Jiwari (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, subadjacency, 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

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.

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

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 Yéfi 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

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

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 Yéli 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 [\pm 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 [\pm 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 Biniñ 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

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 morphological 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 [anómə], [anóm], [nómə], or [nóm]), 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. ŋk); (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 signaling 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

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 subject 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.

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, HPSC, 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 Kratochvíl’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 proto-language, 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 the 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., Ketz, 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*,

- 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. Zubizarreta, 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-Bemema, 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., Dahrymple, 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. Zubizarreta, 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. [BMe]
- 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 (Ojibwa). *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]

- 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 to 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. & Ruppín, 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 with Suffixaufnahme*, ed. F. Plank, pp. 396–430. Oxford University Press. [aNE]
- (2002) The true status of grammatical object affixes: Evidence from Biniñ Gun-wok. In: *Problems of polysynthesis*, ed. N. Evans & H.-J. Sasse, pp. 15–50. Akademie. [aNE]
- (2003a) *Biniñ Gun-wok: A pan-dialectal grammar of Mayali, Kunwinjku and Kane. 2 vols*. Pacific Linguistics. [arNE]

- (2003b) Context, culture and structuration in the languages of Australia. *Annual Review of Anthropology* 32:13–40. [aNE]
- (2004) Experienter 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. [aNE]
- 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. [aNE]
- 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. [aNE, 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 Mwothlap, 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 Vincennes. [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. Pulvel, 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]

- 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. 1*, 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 Zinacantan, 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 (Zinacantepec) 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.pdf> [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 (Abzui dialect). *Transactions of the Philological Society* 77(1): 211–38. [rNE]
- Himmelman, 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. Thaeheksa. [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 1998). [PS]
- Hudson, R. (1993) Recent developments in dependency theory. In: *Syntax: An international handbook of contemporary research*, ed. J. Jacobs, A. von Stechow, W. Sternefeld & 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]. *Behavioral 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 neuro-constructivist 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. *Sprachtypologie 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]

- Knecht, S., Deppe, M., Dräger, B., Bohe, 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 Yéfi 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 Yéfi 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]

- 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]
- Mel'čuk, 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]
- Mycok, 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*

- 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. Aurox, 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 subadjacency 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 subadjacency 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]

- 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, MTa]
- 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, MTa]
- (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 edition*, 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]