Contents lists available at ScienceDirect

Lingua

journal homepage: www.elsevier.com/locate/lingua

Time for a sea-change in linguistics: Response to comments on 'The Myth of Language Universals'

Stephen C. Levinson [a,b,*], Nicholas Evans [c]

[a] Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands
[b] Donders Institute for Brain, Cognition and Behavior, and Radboud University, PB 9102, 6500 HC, Nijmegen, The Netherlands
[c] Department of Linguistics, College of Asia/Pacific, Australian National University, Canberra, ACT 0200, Australia

ARTICLE INFO

ABSTRACT

This paper argues that the language sciences are on the brink of major changes in primary data, methods and theory. Reactions to 'The myth of language universals' (Evans and Levinson, 2009a,b) divide in response to these new challenges. Chomskyan-inspired 'C-linguists' defend a status quo, based on intuitive data and disparate universalizing abstract frameworks, reflecting 30 years of changing models. Linguists driven by interests in richer data and linguistic diversity, 'D-linguists', though more responsive to the new developments, have tended to lack an integrating framework. Here we outline such an integrative framework of the kind we were presupposing in 'Myth', namely a coevolutionary model of the interaction between mind and cultural linguistic traditions which puts variation central at all levels – a model that offers the right kind of response to the new challenges. In doing so we traverse the fundamental questions raised by the commentary in this special issue: What constitutes the data, what is the place of formal representations, how should linguistic comparison be done, what counts as explanation, what is the source of design in language?

Radical changes in data, methods and theory are upon us. The future of the discipline will depend on responses to these changes: either the field turns in on itself and atrophies, or it modernizes, and tries to capitalize on the way language lies at the intersection of all the disciplines interested in human nature.

© 2010 Elsevier B.V. All rights reserved.

1. The general issues

This special issue replays many of the quarrels that have bedeviled the discipline over the last fifty years. Most of these disputes are undignified and ill-informed, and do not serve the discipline well (see e.g. Newmeyer, 1983, 1999, 2005). Nevertheless, there is much at stake – in particular the hearts and minds of young scholars whom we hope will advance linguistics. The new generation have excellent instincts for changes in the academic landscape, and they will look for fresh approaches with a clear sense of mission. Discussions of the sort in this special issue are not the place then to make cheap debating points. We all have a sense of a tipping point in the language sciences, and the question is: What is the way forward?

We believe that linguistics is on the brink of major changes in data, methods and theory. This was the message we tried to get across in the BBS paper and perhaps especially in our response to the 23 earlier comments in BBS – a response, we suspect, that many of the commentators here have failed to study.

* Corresponding author. Tel.: +31 243521276; fax: +31 243521300.
E-mail addresses: stephen.levinson@mpi.nl (S.C. Levinson), nicholas.evans@anu.edu.au (N. Evans).

In the response here we try to locate the central issues in this collection of eighteen further responses to our BBS paper. As before, about half are pro and half are con. Here, drawing support from the pros we try to sharpen our own position, while focussing on the cons, to see exactly where the fault lines lie. The issues matter enormously to a discipline which, if it makes the right moves, can regain its centrality as the study of what makes us human, thus linking the humanities, the social sciences and biology in a way that will guarantee its long-term disciplinary success.

The cons are (as with the first round of responses in BBS) universally from the Chomskyan wing of the generative movement – theorists in e.g. LFG and RRG are not hostile to our line of thinking. However, the critics do not form a united front. Chomskyan theory has moved considerably in the last thirty years, through a series of different models each with their virtues but also with a series of unreconciled different notions of Universal Grammar. Many scholars still operate in the P&P model of the 1980s, not the SMT of the 2000s. Consequently, what one generative commentator will attack, another will defend – indeed it would be easy to show how they effectively shoot each other (see in these comments the variable status of notions like 'subject', 'parameters' or 'UG'). But our purpose here is more serious: it is to try to understand what really divides us.

One issue needs to be resolved before proceeding. Recollect that the BBS article's prime function was to draw the attention of cognitive scientists to the sheer diversity of languages, a fact that has been obscured to the outside world by the very terminology of the Chomskyan tradition (Universal Grammar in particular). The degree of interest in our paper outside linguistics shows that we were right in thinking this diversity constituted news for many cognitive scientists (see e.g. Kenneally, 2010). As mentioned, the opposition to the programme we sketched belongs almost entirely to the Chomskyan camp. It is not, as **the editors**[1] presume, a quarrel between generativists in the broad sense (inclusive of LFG, RRG, etc.) and the rest. Rather, it is a rift between those linguistic approaches that put linguistic diversity central (which includes LFG, RRG, descriptivists, functionalists and typologists) and those that treat diversity as a distraction from the main enterprise (as in the SMT, most clearly stated in Chomsky, 2010). Now this division between the Chomskyans, whom we will call the C-linguists (think of universal Constituency as a mnemonic), and the rest, may seem to leave a motley crew as the residue. But this crew, whom we will refer to as the D-linguists, do subscribe to a number of essential tenets: their theories are Diversity-driven in the sense that they centrally aim to encompass variation, they are Data-driven and more surfacey (and thus more falsifiable) than those of the C-linguists, and they are Darwinian, in the sense that they are interested in variation at every level and the historical and cognitive processes involved in generating and limiting diversity. The great divide then is between the C-Linguists and the D-linguists, and we'll use these labels as a shorthand in what follows.

Here is what we think are the central issues that divide the two approaches:

(1) **The nature of linguistic data**: C-linguists draw on a very small subset of the data – especially, intuitions about complex clauses. Meanwhile the available data types (corpora, typological databases, multimedia records), and the range of data over the languages of the world, has vastly increased in recent years, as has the scientific treatment of grammatical intuitions.

(2) **Models and Theory**: C-linguists presume, on the basis of strong universal assumptions, that the structural analysis of one language can be imported directly into the analysis of another. D-linguists prefer Boasian 'methodological relativism' – first analyze a language in its own terms, then compare. C-linguists equate the metalanguage used in the study of one language with the metalanguage required for comparison across languages, D-linguists don't.

(3) **UG**: C-linguists presume an innate set of principles exclusive to language, D-linguists are at best agnostic. Consequently C-linguists are interested in internal, structure-based explanations, which often seem circular to outsiders, while D-linguists are more attracted to explanations that invoke functional, 'performance' factors, and historical development. Déjà vu and rehash? Indeed, so far, there is nothing new here that has not been discussed in earlier commentaries on the linguistic camps (see e.g. Newmeyer, 1998, 2005). But the external intellectual landscape has in recent years fundamentally changed, radically shifting the balance in ways that the C-linguists do not seem to grasp. The new elements are:

(4) the **vastly increased quantity, quality and types of data** now available to the descriptive and comparative linguist, as will be elaborated below.

(5) **New methods**, drawing on the biological and statistical sciences, which allow us to see pattern in masses of material drawn partly from older scholarship and partly from the new databases. In addition, new computer power makes possible the modelling of language evolution and language change in increasingly realistic ways. It also allows us to model the effects of cognitive factors on these processes, in close connection to data from real human performance in the lab.

(6) In particular, new models of **culture-biology coevolution** allow us to explore how much design must be in the organism and how much design can evolve on purely cultural grounds. Appreciating the power of cultural evolution to produce elaborate design without a designer is crucial to understanding how to apportion pattern between biology and history. It also explains why language generalizations are always statistical in character, since the distribution of attested structural types across the design space reflects the likelihoods of the evolutionary pathways that would engender them, rather than specific constraints on the structures themselves.

---

[1] **Boldface** indicates the author of a commentary in this issue; the full citation will be found in the References under that name (here Editors).

(7) **Interactions with other fields**: These new methods applied to extended data make it possible for linguistics to engage with other disciplines in much more intimate and fruitful ways. Neuroscience and genetics, for example, do not deal with invariants – they deal with phenomena that distribute across the population according to the laws of population biology. But language data, both within and across languages, distribute in exactly the same manner, allowing unification of methods and direct connections to biology. In all these fields, variation is the lever used to discover underlying systematics: you can't find a gene without phenotypic variation, and you can't find a fundamental building block for language without comparing languages.

These developments are why linguistics now lies at a crossroads. It can either seize the new opportunities, requiring a new kind of interdisciplinary training for our graduate students. Or it can remain an armchair discipline, where all the interesting new developments will be made by invaders from other disciplines working on our turf. Consider the fate of cultural anthropology: faced with many of the same opportunities, it has turned inward into an exclusively humanities subject without a future-cross-cultural empirical research is now done largely by biologists, psychologists, behavioural ecologists and other social scientists.

Our own position has been caricatured by our critics, and we'll correct the impression en passant. Suffice it to say for now that of course we believe in looking at languages as systems, of course we are interested in developing a general cross-linguistic framework of analysis, we do not believe in limitless variation, and we are not averse to abstractions or to formal linguistics of a kind (like LFG or autosegmental phonology) that is responsive to diversity.

We turn now to take up these central issues in more detail, using the commentaries as cues to what needs further explication.

2. The nature of linguistic data

The C-linguists take linguistic intuitions of grammaticality and ambiguity as the primary data. Often, the crucial sentences distinguishing different theories are complex embedded structures of a kind that would hardly if ever appear in a corpus.[2] It seems often presupposed that these data give direct or at least best access to iLanguage.

Critiques of intuitive data go back early in the history of generative grammar – Labov (1975) showed that people will deny saying what they actually say. But there has recently been a resurgence of empirical work on the solidity of linguistic intuitions (Schütze, 1996; Tremblay, 2005; Dąbrowska, 2010). The upshot is that intuitions prove fairly wobbly both across speakers and within speakers across time.

Nevertheless, all this can be quantified and studied. Psychologists aim for an **explicit theory of primary data**,[3] on the grounds that there are only a handful of basic data types (Coombs, 1964), each of which makes available a restricted range of models of psychological competence. Levelt ([1974] 2008) applied this approach to intuitive judgements of grammaticality and cohesion (as used by Chomsky, 1965, and C-linguists generally since). He showed that using different kinds of judgement from a good range of speakers, some very interesting conclusions about English structure can be derived – for example, the judgements only support the least hierarchical of plausible trees, and they prefer a dependency over a constituency model of the language (see **Hudson** 2010)! If linguists could agree on a notion of data even remotely similar to that employed in psychology, with its systematic relation between primary observation and analysis, we would not be in the impasse we now find ourselves in.

However, linguistic intuitions are just one kind of data type. There is no particular reason to think that they provide a hot-line to iLanguage. Rather, they are off-line metalinguistic judgements, and do not directly reflect online processes and representations actually involved in speaking and understanding. They should just be seen as another kind of behavioural, performance data – arguably furthest removed from the knowledge of language actually mobilized in language use. Notice that while intuitions about grammaticality have this off-line character, the same does not always apply to that other crucial generativist staple, ambiguity, which does occur in natural corpora and surfaces overtly when an interlocutor interprets an utterance in a way unintended by the speaker, who corrects the understanding (Schegloff, 1988), as well as in jokes and puns.

Other kinds of data include corpus data, experimental production and comprehension data, acquisition data, typological feature data, historical corpora, semantic elicitation data, etc. In recent years all of these data sources have increased massively, so that they are now more representative and cover more aspects of language and its use – and above all, cover a much greater range of languages. There are now 4 times as many grammars appearing per annum as in 1957, there are massive investments (unique in the humanities) into language documentation resulting in archives of up to 60TB of data (as at MPI Nijmegen), there are organized databases of a kind parallel to that in biology (see e.g. WALS; Haspelmath et al., 2005; Reesink et al., 2009) and there is rapid growth of data-mining and data-visualization techniques that make it practical to explore this vast information. With large corpora of video-taped interaction, we can now explore the relation, for example, between questions and their answers (including gaze, nods, gestures) in ten languages from five continents generalizing over many thousands of utterances (see e.g. Stivers et al., 2009). We are in a completely different position as regards the understanding of linguistic diversity, compared to where we were even 10 years ago.

Comparable developments have taken place in the study of language learning, so crucial for theories of UG. We now have 'dense' databases of child language acquisition that capture 10% of everything that goes in and out of a kid in the first years

---

[2] See e.g. Karlsson (2007) on the vanishingly rare occurrence of center-embedded clauses in large corpora.
[3] This remark was prompted by Pim Levelt, who asked: Why can't linguists have a theory of data, just like we have had for half a century in psychology?

(Maslen et al., 2004). Paralleling these changes in the data are new methods borrowed from statistics, psychology and biology for modelling child language development (Vogt and Lieven, 2010). These methods will have a big impact on linguistic theorizing. For example, Bannard et al. (2009) show how unsupervised Bayesian grammar learning applied to actual corpus data can predict a child's output at each successive stage – the induced grammars start off item-based, and only slowly acquire abstract categories like, first, noun and then verb by age 3. No initial abstract categories have to be supposed, contrary to UG.

These statistical methods are proving equally exciting in extracting patterns that allow us to reconstruct language history (see e.g. Dunn et al., 2008). We now know – thanks to Bayesian reconstruction of ancestral states in language families – that many Greenbergian universals are perhaps only an accident of history, and further that syntactic properties are immensely slow to change on average, and show remarkably little linkage to other syntactic properties (Dunn et al., submitted for publication).

For reasons already mentioned, systematic variation is the key to understanding the emergence of structure. This variation can be pursued both within languages (sociolinguistics, dialect studies, and Bybeean functionalist studies[4]) and across them (as in typology). Another crucial variant is the modality of expression, a point made by all the commentaries on sign language (**Sandler**; **Malaia and Wilbur**; **Cormier et al.**). As the commentaries suggest, it makes sense to think of human communication as a modality general affair, with special use of the hand and mouth: what sign languages do is reverse the loading of the different channels. By the same token, speech and gesture should be studied together, as reflexes of the one system.

The quality of data and respect for the different types of data lie at the heart of any science. We also think the relation between data and theory should be as transparent as possible. The generative models fare poorly in this respect. In his BBS response Freidin (2009) cited Chomsky in support of the idea that one cannot possibly produce a counterexample to a principle like subjacency – only a (partial) grammar can be so refuted (Chomsky, 1977:74). Similarly **Duffield** here argues that we have made a "category mistake: no matter how hard one digs into mature grammatical systems, there is no logical reason to expect that one will excavate UG in any recognizable form", while **Abels and Neeleman** argue that "a counterexample to a universal can, in fact, be a confirmation of the theoretical assumptions used to explain it".

The problem with this position is that it is never possible to adduce the knock-down counterexample, the carefully planned experiment that will prove or disprove the theory – it is always possible to invoke escape clauses, without changing the theory. This method of proceeding will not yield a body of generally accepted results, and this is exactly what we find – even within the generativist camp represented here there is a wide diversity of opinion. Equally, the tortuous, non-determinate nature of the representations it sets up makes it extremely difficult to take advantage of the plethora of cross-linguistic data we now have in a direct, even-handed way. A theory should be responsible for a wide range of predictions across data types, and it should be possible to disconfirm it with primary data.

3. Models and theory

A number of the responses both here (**The editors**; **Reuland and Everaert**; **Longobardi and Roberts** and others) and in BBS misread the target paper as some kind of radical anti-theoretical tract. This misunderstanding presumably results from reading the BBS article in the context of the unfortunate dichotomisation of linguists into formalists and functionalists that has been endemic in the field. Newmeyer (2005:Ch. 4), for example, even supposes that given functionalist views on 'emergent grammar' and the like, it is necessary to argue for the very existence of linguistic systems, whereas in fact all linguistic description actually presupposes structuralist notions of system of one kind or another.

In this section, in the expectation that quite a few of the criticisms will thereby be deflated, we spell out our position by clarifying what we see as the proper relation between the basic kinds of linguistic undertaking:

(a) individual language descriptions, presented in a relatively informal way, that capture the structural properties of each language on its own terms (*descriptive linguistics*).
(b) the devising and use of a range of formal tools to characterize the abstract properties of individual languages (*formal linguistics*).
(c) the task of mapping the world's diverse linguistic structures, developing cross-linguistically viable ontologies, and of keeping track of what phenomena occur in which languages – what is where, what correlates with what, and how do we define the phenomena in ways that permit cross-linguistic comparison? (*linguistic typology*).
(d) explanatory theories that account for the distribution of attested types across the design space – what occurs, what doesn't, what is rare, what is common, and why? (*explanatory linguistics*).

The character of these different levels of endeavour can be illustrated by considering their application to the nature of linguistic categories or form classes (like noun and verb), the units over which rule systems operate. This is a microcosm in which the dilemmas of the four levels above can be rehearsed. At level (a), the description of languages, as **Kemmerer and Eggleston** remind us, there are only language-specific classes, and moreover within each language substitution tests will establish hundreds, perhaps thousands, of micro-classes, as the rigorous procedures of American structuralism long ago discovered (Pullum and Scholz, 2007). Formal models at level (b) raise questions for language-specific categories that are

---

[4] Here, the editors remind us, C-linguists have also recently played an important role by systematically mapping microvariation in the syntax of European languages like Dutch or Italian; see e.g. Barbiers et al. (2008).

rarely faced squarely – notions like DP presuppose both categories like D and clear criteria for heads. Yet exponents of D are always language-specific with their own properties, so that the capital letters mask massive equivocations. Language-specificity of categories raises problems especially for levels (c) and (d), as the typologists have become increasingly aware (Dryer, 1997; Haspelmath, 2007, in press). It doesn't follow that comparison is impossible, only that it has to be undertaken in an auxiliary language designed to generalize over language-specific categories.

A number of approaches to this are currently being explored by typologists – one is to use prototypes of e.g. noun and verb categories, as **Kemmerer and Eggleston** suggest, following Croft (2001), and another is to use canonical definitions designed to efficiently partition the design space a bit like cardinal vowels (Corbett, 2006). Yet another is to develop a special auxiliary set of 'comparative concepts' (Haspelmath, 2007), and a fourth is to use a much-finer grained mesh of descriptive concepts that allows diversity to be captured with minimal idealization and data reduction (Bickel, 2009). Once we have comparative concepts, of whichever type, we can go on to ask as in (d), why noun-like categories tend to attract gender marking while verb-like categories tend to attract tense-marking (though things can be reversed).

We now take up the four kinds of linguistic exercise consecutively:

(a) **Individual language descriptions** are the fundamental starting point for any generalizing account of the nature of human language.[5]

Since Boas (1911), descriptivists have followed the fundamental methodological heuristic that a cultural or linguistic system should in the first instance be understood in its own terms. Although such a perspective makes sense in many branches of the natural sciences, biology for example, it has especial relevance for the human sciences, where the apparently same behaviour can have a totally different significance for one cultural group or another. If Hindu widows wear white, while Christian widows wear black and are married in white, this suggests that a primary analysis of the colour symbolism of different cultures should be in local terms. By the same token an /e/ or a word-class or an inversion rule should in the first instance be understood against the local system of phonological oppositions, parts of speech or word-order specifications. The principle is elementary: first do the primary description, then later on the generalizing can get done.

As a matter of practice, initial descriptions of languages tend to use the terminology that has built up through the recurrent practice of grammar writing, including both the categories of traditional grammar (*nominalization*, *adjective*, *subject*, *passive*, etc.), the standard extensions for 'exotic' languages (*ergative*, *antipassive*, *logophoric*, *switch reference*, etc.) and recent terms of art like *inverse system*, *conjunct-disjunct system*, *mirative*, etc. Although these terms have the advantages of flexibility, durability and accessibility to readers of different times and backgrounds, across primary descriptions they have only a family-resemblance anchorage: *all categories should in fact be set up using language-internal justification*. In this way, grammar writing and text annotation have followed the Boasian lead.

Relatively few primary descriptions of a language have been couched in theoretical formalisms, and when they have been expressed for example in tagmemic or early transformational rule systems, this has greatly discouraged later use. Where tree diagrams or other structural representations are used, successful grammars use these more for diagrammatic clarity than for theoretical sophistication. Even for the best studied languages, informed by the widest generative and non-generative research, the best grammars use only minimal formalism, e.g. simple tree diagrams (see e.g. Huddleston and Pullum, 2002). Our call for 'theory-neutral' presentation[6] does not therefore imply lack of theoretical orientation – a host of inquiries in generativist frameworks have opened our eyes to interconnections between structures that we would never otherwise have thought to explore in the description of a new language (see e.g. Rice, 2006).

In sum, primary description has been, and should remain, couched in 'theory-neutral' (framework-independent) terms as much as possible, while at the same time the aspects of grammar explored or texts collected and annotated should be informed by the widest theoretical questions. What has substantially changed however in the last decade is the possibility of putting primary data (audio and video recordings) online, allowing others to come to their own descriptive conclusions using the latest acoustic or symbolic search technology (see e.g. DOBES programme).

(b) **the tool-box for formal description (formal linguistics)**

Every language, we have said, is a thing unto itself. This makes it a challenge to compare languages – a problem we discuss in detail in (c). But even when describing a single language, using natural language metalanguage leaves much to be desired in the way of precision. As Evans and Dench (2006:6) put it:

'The rigorous model-building that formal systems allow makes it possible to test exhaustively the interactions between the many different rules, of phonology, morphology, syntax, etc., that might be included in the average grammatical description'.

---

[5] A corollary of (a), which we return to in section 5 on coevolution, is that descriptions can in principle be applied either to the language produced by an individual or by a social group (at varying degrees of aggregation), and that with respect to an individual this may apply at one or more stages in their development. This allows us to include e.g. grammars of a child at a particular age (which can then be compared with caregiver input, or the results of processing experiments), but also e.g. a pan-dialectal grammar that includes statements applicable to subparts of the population. This move recognises the dual status of language as social convention and individual representation.

[6] We put 'theory neutral' in scare quotes because we all recognize of course there is no such thing; but there is a mode of description that minimizes commitment to controversial analyses and passing fads. See Dixon (2009) for a developed framework of concepts.

Linguistics has acquired a formidable toolbox of formal, more or less algebraic, representation systems, thanks especially to the generative enterprise in the widest sense. This has been augmented by innovations from e.g. autosegmental phonology (such as linked representational tiers), from information science (such as attribute-value pairings) and from logic (such as lambda calculus). Although the application of each of these may make theoretical commitments, it is a category mistake to conflate 'theory' and 'formalisation', if only because rival formalizations can make identical generalizations, or fail to make any interesting generalizations at all.

C-linguists seem to think that D-linguists are anti-formalism, a view that colours many of the responses to our BBS article. Formal representations only have real power as part of a formal system that allows theorems to be derived or rule-interactions to be checked, but few linguistic representations are actually used in this rigorous way, outside computational linguistics (where engineering expediency is more important than cognitive plausibility). Certainly C-linguists rarely if ever present computationally complete rule systems. Outside these rigorous uses, formal representations have their uses – they can help us grasp the nature, e.g. of linguistic ambiguities, where essentially they add the clarity and insight of diagrams (as in many uses of tree diagrams). They can also help to minimize equivocation on terms, and prompt the generation of counterexamples. Simple formal models may also allow the entire system to be visualized, as e.g. with distinctive features (which like tree diagrams long predate the generative paradigms). In general, formalisms have their virtues, rarely properly exploited, but they notoriously date, making interesting observations inaccessible to later generations of scholars, and hence should obviously be used with caution in any primary description. But apart from these slightly deflationary remarks, we doubt that any D-linguists dispute the utility of formalisms in their proper place. They sharpen our thinking and make precise our understandings of phenomena.

We have inherited a rich formal toolbox from the logicians, philosophers, computational theorists and linguists who have tried to model linguistic representations. The tools should be treated as what they actually are, namely representational devices for checking out the properties that that particular tool is designed for. Rule systems are good for checking out predictions of productivity, and good for thinking through how one linguistic structure might be related to, or feed another. Hierarchical representations are good for testing structural units and dependencies, feature representations for thinking through inheritance from major categories, logical systems for checking out the predicted inferences. All of these have their many variants, autosegmental tiers, constituency grammars of myriad sorts, dependency grammars, feature unification systems, logics of a thousand kinds. Often the variants are more or less inter-translatable, so not too much store can be put on any one of them. As we argued in the BBS paper, this variety of multiple representational mechanisms (such as dependency formalisms as alternatives to constituency-based formulations), which can be tried for best fit against language-specific facts, is one of the cumulative triumphs of linguistic research. But formalisms should be used eclectically for the descriptive job at hand. Just as a whole logic is designed to model just one little aspect of natural language – the connectives (propositional calculus), the quantifiers (predicate calculus), the modals (e.g. S5), say – so each of these representational systems has its specialized use.

We note below that many terms like subject, head, or secondary predicate have, in most uses, loose family-resemblance definitions. They also of course have formalization-dependent definitions, which will be much more restrictive, and much confusion results from not clarifying the use in question. It is also natural that different theoretical schools will employ different representational tools for the same language-specific phenomena – leading to debates e.g. about whether to represent a particular syntactic characteristic in terms of constituency or dependency (see **Hudson**'s commentary). This is a normal part of science, as the fit of different models to the data is tested. What is more problematic for progress in the field is the use of particular formalizations as markers of group identity in linguistic wars, so inhibiting the standardisation of terms and evaluation of alternative models, and making impenetrable barriers to researchers from other disciplines.

Where we differ radically from many of our commentators is our opposition to the conflation of the metalanguage with the object of description (see our BBS Response, p. 474), a Chomkyan move that seems to us a fundamental category mistake, like confusing mathematics with the objects it describes – a nautilus shell exhibits a logarithmic spiral, but it doesn't use logarithms to produce its shell. Even the strongest universal patterns (e.g. rule constraints like 'formal universals', if such can be found) could emerge from multiple constraints in populations of language users and learners, without residing in the heads of speakers. This conflation was exemplified by commentators on the target BBS article such as Berent ('the brains of all speakers represent a shared set of grammatical categories') and Nevins ('formal universals in phonology are constituted by the analytic elements that human minds employ in constructing representations of sound structure'). By avoiding this move – and the misleading baggage of the term 'formal universals' – and by recognizing that formal tools are simply part of the linguists' metalinguistic toolbox, it becomes possible to introduce new formalisms to accommodate newly discovered facts, without burdening the neonate mind.

(c) systematising cross-linguistic data (linguistic typology)

The fact that every language works in terms of its own internal categories constitutes a fundamental problem for building a cross-linguistic metalanguage that is at once true to language specificities and gives us a means to generalize across them. When illustrating with the problem of morphosyntactic categories, we mentioned four tacks that have been taken – use of prototype categories (Croft, 2000), use of canonical types (Corbett, 2006), devising a set of generalized categories (Haspelmath, 2007), or use of a much-finer feature grid that is able to capture both the language idiosyncrasies and the generalizations (Bickel, 2009). We shall refer to this last kind of cross-linguistic metalanguage as a Low-Level Feature

Metalanguage (LLFM). This approach looks especially promising, because it allows rigorous definition of a concept like 'noun' in terms, e.g. of a polythetic set of many complex features, much as 'subject' is now treated in most traditions as a composite notion. These larger feature sets will allow new, unanticipated generalizations to emerge from comparative research – e.g. new kinds of morphosyntactic cross-linguistic categories.

There is a fifth approach, of course, which is to use off-the-shelf categories arising from specific grammatical traditions, and foist them on all languages. This is essentially the method that C-linguistics has employed. As outlined in our BBS article, typologists and descriptivists deplore it not only because it does procrustean violence to the basic data, but also because it pre-empts the discovery of new categories and new patterns.

Yet a further alternative approach is sketched by **Keenan and Stabler** (2010) which abstracts right out of the constructional details. It relies on two moves: a semantic move that defines the phenomenon of interest without recourse to structural definition; and a structural move that characterizes the nature of inherited structures – unmodified, or modified, as the case may be. So the particular claim that they advance amounts to this: In all languages you should be able to locate Anaphors semantically, and you will then find that (in quite different ways) anaphoric constructions are structure preserving and have a fixed structural relation to their 'antecedents' (no order relation implicit here).

There is a similar way to try to achieve a common framework for comparison. This is again to exploit the semantic basis for our comparative notions, as long recognized (e.g. by Greenberg) as the basis for comparative morphosyntax. Two methods have emerged here. One is the method of semantic maps, encompassing the prototype approach mentioned above (Croft, 2001). The second method uses e.g. visual stimuli designed to systematically explore the nature of a semantic domain, like the nature of local notions encoded in adpositions (Levinson and Meira, 2003) or positional verbs (Ameka and Levinson, 2007) or the nature of concepts encoded in natural language reciprocals (Evans et al., in press). Both methods yield a *tertium comparationis* or 'etic grid' allowing controlled comparison.

In this programme, what happens to all the comparative findings produced by nearly 50 years of generative linguistics? Does the entire baby go out with the bathwater? Of course not. Let us suppose we have a good set of generalizations about constituency, c-command, reflexives and reciprocals initially derived from a set of European languages. Then to the degree that a language has the constituency required to define c-command, these 'Binding Conditions' make predictions both about the distribution of morphemes and about the interpretations of pronouns. It then becomes a further question whether these conditions are descriptive or explanatory (or e.g. side-effects of pragmatic principles as sketched in Levinson, 2000).

In the same way, where formalisms help as compact statements or diagrammatic clarifications, or (as they rarely do) theorem-deriving methods, they too can be based on LLFM categories, or on a higher level vocabulary defined on top of it. LLFM is neutral between, but compatible with, rival schemes here. In this way, the tiers of autosegmental phonology, or the ranked constraints of OT, may continue to perform the illuminating role they have played in formulating cross-linguistic generalizations without the mentalistic, universalizing overtones.

The applicability of such concepts to a given language always needs to be overtly demonstrated, and squared with local heuristics. They can be reanalyzed in more specific LLFM features, and this would allow us to 'translate' e.g. c-command into a corresponding dependency notion and go on and explore the predictions in languages without the necessary constituency. We know there are counterexamples to the Binding Conditions (see our discussion in (d) below), but we could now go on to try and pin down exactly which languages (categorized in LLFM terms) counterexemplify and why. The point is this: Most of the generalizations that linguists care about can be expressed and tested in LLFM. We are not going backwards, but we expect every one of the generative generalizations to hold for only a proper subset of languages, and the interesting question then becomes – why this subset? Are these features that are historically inherited within language families, or do they have the character of 'stable design properties' characteristic of emergent solutions to perennial problems?

What now needs doing is to launch a major new research effort for world-wide linguistic typology. In practice, current typology proceeds by using somewhat loose family-resemblance categories, like 'verb', 'subject', 'basic word order' and so forth (see e.g. WALS, and the extended explanation of each of these notions). Like Bickel (2009) we think these family-resemblance notions can be given firmer foundations by a much larger, more explicit set of underlying features which will be quite detailed in character. Replacing family-resemblance metalanguage with a more precise and augmentable LLFM would finally allow the construction of a database of the major structural features of all the languages in the world. The majority of linguists follow the classic humanities method of working with minimal resources and alone, without a team. But an effort of the kind we are proposing here requires a completely different level of ambition and a different order of finance. Most linguists will feel this is completely out of reach, but in fact 7000 languages is not an insurmountable goal – it pales into insignificance against the achievements of comparative biology (there are about as many amphibians, frog and toad species, alone). Given the importance of language in human existence and evolution, and the keen interest shown by natural scientists and journals like *Science*, *Nature* and *PNAS* in efforts to do this kind of thing (see e.g. Dunn et al., 2005; Gray and Atkinson, 2003; Pagel, 2009; Reesink et al., 2009), there is little doubt that the funding for big programs of this sort could in fact be found. But efforts of this kind are going to need serious database and computational skills, both to organize the data collection and mine the masses of data for the fundamental insights into the nature of language that they will hold.

(d) **explanatory theories**

The editors hold that 'both sides are part of the scientific enterprise but the generativists' attempt at theory construction manifests additional, explanatory, ambitions'. The division between explanation-seeking C-linguists and merely taxonomic,

butterfly-collecting typologists and descriptivists is a favourite generativist trope. But we think that few if any D-linguists would agree with this imputed lack of explanatory ambition (see Dryer, 2006). Indeed, they share most of the very same explanatory goals, differing primarily in the preferred mode of explanation.

The goal of linguistics is, the textbooks endlessly repeat, to explain language universals. The BBS paper, in questioning the existence of such universals, moves the goal posts in fundamental ways. We would rather say: the goal is to explain why languages have the properties they do, including properties that they may tend to share. If these are the goals, it is useful to follow Newmeyer's (1998) distinction between internal and external explanation. C-linguists are locked exclusively into internal explanation, explaining linguistic observations in terms of language-internal systematics.

For example, if the cross-linguistic data-gathering enterprise in (c) confirms Greenberg's generalization that, while there are languages with O_{pron} V and V O_{nom}, there are no languages with O_{nom} V and V O_{pron}, this calls for explanation. One can formulate either structure-based explanations (such as that proposed by **Abels and Neeleman**) or function-based explanations – e.g. that old information is ordered to come first, that pronouns represent old information, and that full NPs represent new information, so we would expect pronominal objects to grammaticalize into earlier positions than full NP objects. Rather than presume that only Chomskyan generativists seek explanatory theories, it is more useful to develop an integrated set of rules for argumentation that allow us to evaluate functional vs. structural accounts, along the lines of Newmeyer's argumentation.

Chomsky's famous metric of observational, descriptive and explanatory adequacy was intended to discriminate between competing ways of generating the grammatical strings of the language. This has had little practical impact on the field, because we have rarely had the pleasure of deciding between two rival models equally capable of generating the data, let alone generating the structural intuitions that go with them. However, in so far as hanging on to the postulates of UG is thought to amount to achieving explanatory adequacy, the scene is set for damaging circularity: some analysis is to be preferred over another because it would be consistent with those postulates.

This helps to explain the kind of desperate ruse proposed by our critics – such as **Reuland and Everaert's** postulation of invisible anaphors for Fijian, or **Safir**'s argument that counterexamples to subjacency can be dealt with by postulating covert resumptive pronouns, and that there is underlying recursion in Kayardild which we just can't see because it will always be morphologically blocked, or **Longobardi and Roberts**'s suggestion that free-word order can be generated from fixed constituency-generated orders by applying scrambling rules. It is against analyses like these that our point about taking diversity at face value is intended: not that entertaining abstract entities should never be done, but that it is only justified where the language-internal facts warrant it. If we don't adopt this methodological constraint on the postulation of structure, it is all too easy to render putative features of universal grammar unfalsifiable.

The classical goals of linguistics – to explain why languages have the properties they do – clearly allow, indeed require, both internal and external explanations. These were the directions we sketched in our Response to the earlier commentaries in BBS. Essentially, what we suggested was that the systemic properties of languages are to be explained by the processes of cultural evolution interacting with cognitive and neurophysiological constraints. No one perspective will ever explain what is common, what is rare, the directions of language change or the possibility space that languages inhabit. There is no chance, for example, of unearthing innate constraints on syntactic form without a larger model of the interactions between population effects, historical forces, memory and learning constraints, and so forth. Such a model has the essential characteristics of population genetics: methods for understanding how individuals produce the population-level characteristics. We take up these issues again in section 5, and consider any such venture thoroughly theoretical, although closely tuned to empirical material at every level.

We recognize that C-linguists mean something quite different by theory. Specifically, they generally mean an internal systems model that relates different levels of representation to one another (e.g. phonology to syntax to semantics). Such a model is just a part of the larger population-genetics model that we have in mind, just as demographers use models of family structure to generate out the expected population structure. Our position is that these systems models are going to be largely language specific, while recognizing that there are basic architectural constants, like the triad of semantics, morphosyntax and phonology, as in Jackendoff, 2003. However, like Jackendoff, and indeed most D-linguists, we doubt the true 'autonomy' of the levels, or the restriction of generativity to the syntax. **The editors**, for example, defend the idea that "phonological and semantic rules can refer to syntactic information, but syntactic rules cannot refer to phonological or semantic information", a C-linguistic credo that caused the linguistic wars of the 1970s, which were won only by attrition (see Newmeyer, 1986).[7] These architectural issues are of considerable interest, but they have been deeply clouded alas by ideological positions (see e. g. Seuren, 2004 vs. Chomsky, 1995 on the semantics/syntax interface).

4. Universal grammar

Formulations of Universal Grammar within the Chomskyan tradition have varied dramatically over the years, from the Aspects views about substantive and formal universals, through the P&P model which tried to build all the seeds of possible language variation into UG, to the stripped down model of SMT where UG is almost or even totally vacuous. The exact

---

[7] Anyone defending the independence of syntax from phonology must handle the new data coming in all the time about linguistic variation, for example "literal alliterative agreement" as found in Atlantic languages, where agreement is marked by productive copying of the first syllable of the controller item (Dobrin, 1995).

content of what it contains has thus waxed and waned over time, with respect both to formal and to substantive universals. In the early phase of theory development the concept was open to testing against the data of language diversity:

> The main task of linguistic theory must be to develop an account of linguistic universals that, on the one hand, *will not be falsified by the actual diversity of languages* [italics ours] and, on the other, will be sufficiently rich and explicit to account for the rapidity and uniformity of language learning, and the remarkable complexity and range of the generative grammars that are the product of language learning. (Chomsky, 1965:278)

> We must postulate an innate structure that is rich enough to account for the disparity between experience and knowledge, one that can account for the construction of the empirically justified generative grammars within the given limitations of time and access to data. At the same time, this postulated innate mental structure *must not be so rich and restrictive as to exclude certain known languages*. (Chomsky, 1968:79)

> These quotes make the empirical orientation of early conceptions of UG quite clear. The substantial content claimed for UG reached its peak during the Principles & Parameters phase during the 1980s. At the heyday of the P&P model, UG included the entire parameter set, effectively demarcating the set of all possible languages:

> What we 'know innately' are the principles of the various subsystems of S_0 [the initial state of the language faculty] and the manner of their interaction, and the parameters associated with these principles. What we learn are the values of these parameters and the elements of the periphery ... (Chomsky, 1986:150–151)

Some of the empirical counterexamples we offered in our BBS article were aimed at that phase – such as claims about limitations in the distribution of bound anaphors, falsified as we show there by a range of languages. (See also Evans and Levinson, 2009a:433, 2009b:476 and references therein for arguments against regarding binary phonological features as formal universals.) **Keenan and Stabler**'s article adds further counterexamples to those we adduce there, showing "the failure of Principle A with its attendant c-command requirement on antecedents of anaphors" (see also Keenan, 2009). The fact that **Reuland and Everaert** attempt to defend it in their paper, through the sorts of analytic moves whose defensibility we argue against in Appendix A, illustrates how little consensus there is on what substantive universals remain standing.

The survival of so many UG hold-outs from the expansive GB phase is particularly striking at a time when Chomsky himself has retreated to a much narrower view of the empirical base of UG, confining it to Merge or even just recursion in his recent formulations. Within the SMT model, language variation is essentially a matter of the accidents of 'externalization' (Chomsky, 2010:61). Increasingly, he tends to adopt the narrowest view of UG – see e.g. Chomsky, 2010:51, where as non-linguistic cognition progressively turns out to explain linguistic patterns, so "the hypothesized role of UG declines" (see also **the editors**, 2010).

Different C-linguists, then, have gotten off this roller coaster at different stations, often vacillating between the attractions of one or the other. Chomsky (2007, 2010) himself in his latest writings often seems to want to be in two places at once (P&P parameters vs. SMT vacuity), leaving valiant commentators like Golumbia (2010) struggling to make sense of it.

This fractionation of the generativist enterprise across different phases presents us with a great diversity of frameworks and rival analytical presumptions. This makes it impossible for antagonists to formulate a single coherent response, just as when Hercules took on the Hydra and every head cut off turned into two new ones. Instead, we find a landscape of half-abandoned generativist models, from which defenders may pick and choose analytical tools. It is the coexistence of these variants that allow our critics to accuse us sometimes of naïvely supplying putative counterexamples to claims that cannot be tested in this direct way, while at other times they propose alternative analyses of phenomena we cite which purport to remove them as counterexamples.

The editors in their editorial demonstrate the typical slippage in the use of the term UG. They distinguishe $UG_{definitional}$ from $UG_{empirical}$ but thereafter fail to tell us which sense of UG they have in mind. So when they conclude "It is strange that some accuse UG of being unfalsifiable, others of being false. It can't be both", they are wrong: $UG_{empirical}$ can be false, but $UG_{definitional}$ cannot be! Even this distinction between definitional and empirical forms of UG doesn't capture all mutations of the claim. Features attributed to the neonate mind are hypothetical, not definitional, though much more difficult to test directly, as several commentators have pointed out (e.g. **Duffield**, **Crain et al.**, see below).

There is one further, crucial ambiguity within current use of the term UG which is the source of much confusion. On the one hand it can denote (as we pointed out in the BBS article) no more than whatever it is that the child brings to bear on initial language learning (including e.g. the necessary general memory and pattern learning abilities). Everyone has to subscribe to UG in this sense of course – although we think the term is utterly misleading, as indicated below. On the other hand, it can refer to a domain-specific innate faculty of language, which many D-linguists are skeptical of. In the terminology of Hauser et al. (2002), there are two interpretations of UG: UG corresponding to all the infrastructure that yields FLB (faculty of language in the broad sense), or UG corresponding only to FLN (faculty of language, narrowly understood to refer to domain-specific principles). As **Sandler** (2010) notes, many but not all responses to the BBS article took it strictly in the latter sense (UG/FLN), possibly creating more heat than light.

Despite the clear tendency in Chomsky's work to assume UG/FLN, in the paper with Hauser et al. (2002), he freely entertained that FLB may be all there is, and FLN may be vacuous (as **the editors** concede). We took the line that FLN is, as a

working presumption, best considered vacuous (although we excepted the speech system, including the vocal learning unique among the primates, which though confined by Chomsky to the sensory-motor 'interface', is the only clear innate adaptation for language). It follows that, in our view, there is no Faculty of Language in the Fodorean sense of Faculty Psychology (no module or modules wholly specific to language processing), and UG is an unfortunate misnomer, because there is nothing essentially grammatical about the capacities an infant uses to acquire language. To make this crystal clear: UG/FLB must include all the infrastructure for language including the neuroanatomy for speech and cognition, theory of mind, statistical learning capacities, and all the communicative and cooperative motivations and interactional abilities special to the species (see Enfield and Levinson, 2006 for a compilation of views).

If all this is agreed, it is time to change the terminology and avoid all the theory-laden terms used so far. We would urge our colleagues simply to talk about the human capacity for language (including language-special aspects of cognition, if any) and the ingredients or elements that may contribute to that – the rest carries too much baggage, as the field advances rapidly beyond the creaky old baggage train.

5. The dividends of a coevolutionary approach

Our paper advocated the explicit adoption of a coevolutionary approach to linguistics, that is, understanding language within the context of the interactions between culture and human biology that have been the source of the peculiar success of our species – we have invaded every niche on the planet just through using culture as our special mode of adaptation. Cultural adaptation requires large brains and complex communication skills, so we have evolved a system where our abilities are a curious mix of cultural and biological evolution. We have, as it were, partially 'outsourced' our adaptive systems into the environment, where processes of cultural evolution hone languages into the marvelous instruments they are. The theory of twin track evolution (aka dual inheritance systems), of how culture and biology interact, is now an advanced topic of investigation (Durham, 1991; Richerson and Boyd, 2004; Odling-Smee et al., 2003). This ongoing process is reflected in a genome under constant selective pressure, with new alleles peaking just in the last few thousand years (Hawks et al., 2007).

Why should linguists have any interest in these larger frameworks, which may seem completely irrelevant to their ordinary preoccupations? The reason is that stepping back a little can resolve a lot of the background puzzles that bedevil the discipline. What is the ontological status of language – is it a thing in the head or in the community? In what sense could language be innate, and how could it get, as it were, into the genes? Where does the design in language come from, and the intricate patterning unique to particular languages? Why are there so many languages? Why are there nevertheless striking similarities across them? Why are some languages more complex than others, having, say, more phonemes or grammatical categories than others? How and why did language evolve? These are all the sorts of question that that one might expect a discipline dedicated to language to be able to answer, but they will be answered only within a coevolutionary framework.

Here we spell out just a few of the dividends that this larger framework brings us, noting the consequent shifts away from the kinds of position espoused by the C-linguists.

5.1. Language evolution

Coevolutionary approaches are particularly useful for modelling the early emergence of language as a multi-step process. This is because the models allow advances in each interacting system to egg the other on, much as computer software and hardware have fed the other in a spiraling coevolution over the last decades.

Chomsky is well known for his anti-adaptational view on language origins, imagining some sudden emergence of a "Prometheus" complete with Merge in the last 50–100,000 years (Chomsky, 2010). In fact, all the evidence is against any such saltational view of language origins, and increasingly points to the gradual assemblages of a number of elements (cf. Lieberman, 2006). We now know that genetically we were nearly identical to Neanderthals (only half a dozen genes separate us, mostly associated with skin and bone development). We thus shared essential features with the common ancestor who goes back at least 400,000 years (Green et al., 2010). By that time all the most crucial elements of the system of speech and audition were already fully evolved (Kay et al., 1998; MacLarnon and Hewitt, 1999; Martinez et al., 2004), including the modern form of the FOXP2 gene contributing finer motor control over speech (Krause et al., 2007). There are still some residual questions about how the shape of the supralaryngeal tract in fossils of this age may have impacted the vowel space (Boë et al., 2002; Lieberman and McCarthy, 2008; de Boer and Fitch, 2010), but there can be no doubt that pre-modern humans were talking.

Thus just as the anatomy of the hand has coevolved with tool using, it seems clear that our vocal apparatus has coevolved with language (or at least speech) over some half-million years. Moreover, the interaction of genes and language continue to evolve, as shown by the recent finding that language change has been channeled by population genetics in the last few thousands of years (Dediu and Ladd, 2007; Nettle, 2007). There are also fundamental differences in the way language is implemented in brains (Catani et al., 2007), reflected in considerable individual differences in language performance, providing the raw material for ongoing evolution. If we are at all interested in language diversity and language variation, or the instantiation of language in the brain, we will need to bear a coevolutionary model of language in mind.

Coevolution presupposes that the cultural track can maintain a tradition over great time periods, a supposition amply verified by archaeological work on e.g. stone tool traditions. In the last decade a flurry of applications of bioinformatic techniques to cultural materials have demonstrated strongly tree-like branching in cultural traditions in diverse materials,

from Turkmen carpets to Polynesian canoes, tool-kits to bride wealth (see e.g. Mace et al., 2005). But language stands out as the 'model organism' for such endeavors, and treating historical linguistics as if it was a branch of genetics has yielded fundamental insights from the deep age of Indo-European (Gray and Atkinson, 2003), to the frequency basis for ancient preserved vocabulary (Pagel et al., 2007), to the possibility of structural phylogenetics (Dunn et al., 2005). Scepticism based on the extent of borrowing and horizontal transfer turns out to be unwarranted (Greenhill et al., 2009). It is notable that the key figures in these developments are mostly not linguists (excepting Dunn et al., 2008; McMahon and McMahon, 2006), and that linguists have mostly taken little (unless hostile) interest in these developments even when they promise to revolutionize the field.

5.2. Evolved design

Cultural evolution is hardly ever mentioned in theoretical discussions of language, yet it is clearly central to any coherent theory of language diversity and design, and has been the subject of substantial research (see e.g. Mace et al., 2005; Boyd and Richerson, 1985; Henrich et al., 2008; Pagel, 2009).

Cultural evolution is capable of evolving highly intricate systems by the slow accumulation of (mostly) tiny increments over generations – witness mathematical or musical systems, or the nature of the economy. Sometimes there is not the slightest trace of individual agency or individual teleology in such systems, as for example in the intricate design of kinship systems (clearly phylogenetically inherited, see Jordan et al., 2009), the most complex of which often elude the understanding of individuals (you may well know what to call your mother's brother's son, but not understand how that is part of a systematic system tailored to a pattern of marriage exchange). A good example is the emergence of a paradigmatically structured eight-category subsection system in many Aboriginal languages (Evans, 2003; McConvell, 1985). Historically, these arose from the interaction of two four-section systems in neighbouring languages, which got integrated as bilingual speakers begin to treat it as a single interwoven system. The crucial thing is that it could begin to function as an eight-class system – as indicated by the rules for recruitment into one class from another via marriage and matrifiliation – before the stage at which any individual needed to represent it as such.

Complex design without a designer can thus be achieved by cultural as well as by biological evolution, and indeed every linguist wonders at the miracles of design in the particular languages they study.[8] Take for example the complex paradigms familiar from Latin, Greek, Bininj Gun-wok, Yélî Dnye or any number of other languages described recently for the first time. These can be so complex that they can be hard to grasp, and are readily misanalyzed – a nice simple example is the discovery of the logic of augmentation in Rembarrnga pronouns (McKay, 1978). Once you see the logic, you get a genuine 'Eureka' experience. No-one designed a paradigm. No paradigm is innate. There is no rival explanation. A paradigm is a clear example of design through cultural evolution, honed through generation after generation of language learners.

Paradigms are a simple example of the power of cultural evolution. But of course whole grammars are subject to the same shaping forces. When people talk about cultural evolution they often fixate on rapid processes of change like the arms race. But languages are able to retain a signature of lineal descent almost as long as genes (see Pagel, 2009 on 20,000 year old words, or Dunn et al., 2005 on ancient structure dating back right through the Holocene). It will be obvious to anyone inspecting a whole conservative language family (say Dravidian or Mayan) that often many properties (distinctive phonological features, main word order, parts of speech) have stayed the same through each of the descendant branches of the family, each branching representing separate millennia of descent. The structural properties of language change *on average* very slowly indeed, on the order of tens of thousands of years (Dunn et al., 2008, submitted for publication).

Combine this staying power with the fact that the modern human species only left Africa within the last 100,000 years, and we see immediately that the patterns of language diversity are as much a matter of historical accident as they are of the systematic exploration of the design space of possible languages – a point we made in response to Pinker and Jackendoff's commentary on the target article. Moreover, the fact that apparently unrelated languages may seem similar in design is no longer surprising – they may indeed be distant cousins (see e.g. Vajda, 2008 on an unexpected inter-continental link). This perspective again shifts the goal posts – we cannot search for language universals or do typology as if each language was an independent experiment: they share structure through both descent and hybridization (language contact). Historical, evolutionary processes – that is, descent with modification – lie behind each and every grammar.

This is one reason why sign languages (especially the minor, indigenous ones) have such an importance for linguistic theory. Indigenous sign languages are the only languages without ancestors, the only ones that actually test *de novo* the limits of the design space for languages, which on current research looks surprisingly expansive (Meir et al., 2010; and **Sandler**; **Malaia and Wilbur**; **Cormier et al.**, 2010).

5.3. Ontology: iLanguage and eLanguage, the individual and the speech community

The move from Saussure (or structuralism generally) to Chomsky involved a shift from viewing language as a social institution (*langue*) from which individual representations were derived, to viewing it as a phenomenon in individual minds (initially, 'competence', later 'i-language'). For Saussure, individual knowledge was secondary and derivative, conceived of as a series of identical copies deposited in the minds of each speaker – individual variation was relegated to *parole*, rather than

---

[8] Of course, languages also display horrible kluges and inefficiencies just like naturally evolved organisms (Marcus, 2008).

characterizing individual mental representations[9]. For Chomsky, language as a social phenomenon was achieved, also secondarily and derivatively, by the opposite move – assuming an idealized, homogeneous speech-community replicating the representations of individual speaker-hearers.

The inadequacy of each of these positions has, however, become increasingly clear. There is manifest non-identity between the individual and social guises of language, obvious in the case of the lexicon (no single speaker knows the entire contents of even a modest modern dictionary), and haunting the study of syntax (as **Longobardi and Roberts** indicate, talking of the many dialects of Italian). This has prompted developments like stochastic models of syntax which attempt to bridge the gap by relating variation in the speech community to individual grammaticality judgements (see Bresnan et al., 2007). Sociolinguists have grappled with these problems from the beginning, given the structured variation within a speech community that is constituted by individual grammars but reaches beyond individual overview. This led to Weinreich's (1954) concept of a *diasystem*, and Bickerton's *panlectal grammars* (Bickerton, 1976).

This is another reason why we need coevolutionary approaches. They allow us to produce a more complex and satisfying view of language – and culture – as something that is simultaneously a psychological and a social phenomenon. If I am the last speaker of a language, which has never been recorded and which I never speak any more for lack of conversational partners, the language exists only in my head. At the other extreme, a language known only through surviving documents and not spoken by anyone – say, Tangut – survives as a (written representation of) a social phenomenon only. Between these two abnormal extremes, language lives a double life, passing back-and-forth between individual and social guises, and many scientific advantages accrue to models that recognize this explicitly.

Individuals develop their own mental representations of language through exposure to the social phenomenon of language, as heard from (usually) many people in their environment, with each then producing utterances on the basis of their own individual representations. This sets up a perpetual cycling back-and-forth between the social and the individual/ psychological, which is the mechanism for language change, involving the selection of variants produced by the operation of various functional factors – e.g. articulatory shortcuts, reductions, pragmatic enrichments, abductions, cognitive constraints on what is easy to represent, or syntactic reanalyses. A whole host of variants are rattling around in any system and language-acquiring children make their own selection from among these in building their individual linguistic representations. The endless cycling between the individual and social is also a process which has some quite surprising properties, which we turn to next.

5.4. Emergence

There has been a great deal of research in the last two decades on how complex structural properties can emerge from this sort of coevolutionary back-and-forth, thus offering alternatives to the view that these properties are directly part of our innate endowment, or somehow part of a fixed architecture for language. Recent work has shown, for example, how

(i) duality of patterning is an emergent property of such recycling of systems through individual heads (Del Giudice et al., 2010);
(ii) emergent discrete combinatoriality also drives the CV template in phonology (Zuidema and De Boer, 2009);
(iii) arbitrary, Saussurean signs arise out of such iterated learning (Kirby, 2001; Kirby and Hurford, 2002);
(iv) phrase-structure grammars emerge in unsupervised learning from raw corpus data (Waterfall and Edelman, 2009);
(v) quasi-recursive abilities comparable to actual human performance emerge through probabilistic learning (Christiansen and Chater, 2008).

This is an extraordinary catalogue of many of the features that have been thought to be definitive for human language, now shown to be potentially emergent. A general property of all these simulations is the amplifying effects of repeated learning biases – as Chater and Christiansen (2009) put it, "through prior generations of cultural selection, the form itself will have been optimized to embody whatever inductive biases the learner may have".

Many widespread cross-linguistic tendencies can then be accounted for as emerging from funneling and amplification processes of this type, without needing to be located in 'individual knowledge of language'. A number of the commentators touch on just such emergent tendencies, and attempt to revive them as candidate universals. **The editors** for example suggest that all languages tend to follow the Keenan-Comrie accessibility hierarchy,[10] presuming therefore the hierarchy must be in every child's head – "an internalist perspective is inevitable". But emergent properties are everywhere, and this is almost certainly one of them. We think that "an internalist perspective" is exactly in error here. As we pointed out in connection with OT approaches:

---

[9] 'La langue existe dans la collectivité sous la forme d'une somme d'empreintes déposées dans chaque cerveau, à peu près comme un dictionnaire dont tous les exemplaires, identiques, seraient répartis entre les individus[…] C'est donc quelque chose qui est dans chacun d'eux, tout en étant commun à tous et placé en dehors de la volonté des dépositaires… Il n'y a donc rien de collectif dans la parole; les manifestations en sont individuelles et momentanées.' (de Saussure, 1979:38) [Language exists in the form of a sum of impressions deposited in the brain of each member of a community, almost like a dictionary of which identical copies have been distributed to each individual. Language exists in each individual, yet is common to all… Speaking is thus not a collective instrument; its manifestations are individual and momentary' (Baskin transl., p. 19)].

[10] Note that this is posed in terms of grammatical relations like subject and object that other generativists like **Crain et al.**, say are entirely obsolete; see Newmeyer, 2005:21ff for earlier generative attacks on this hierarchy.

'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'. (Evans and Levinson, 2009b:474)

The counter-hypothesis, as we have sketched above, is that the funneling properties of coevolutionary selection generate distinct sets of relativisation rules across languages which, when taken together, can be organized into the implicational pattern known as the Keenan-Comrie hierarchy. This is not something that any speaker ever needs to know in order to learn the grammar of an individual language.

The property of 'structure-dependence' (brought up by **Crain et al.**, and **the editors**) is another example. Structure dependence may be the strong tendency, but many kinds of exceptions exist, from 'heavy NP shift', to the counting of surface units sometimes involved in stress-rules (Goedemans and van der Hulst, 2008) or morphology, or the systematic grammatical effects of animacy or definiteness hierarchies.[11] The rarity of 'structure-independent' rules can be accounted for in a coevolutionary framework through the operation of a 'diachronic filter': there are just very few functionally motivated diachronic pathways that would lead to their emergence. Alternate structures (the nearest analogue to structure-dependent transformations in a non-transformational model) arise by reinterpretations of existing structures through reanalysis. To get alternations corresponding to 'structure-independent rules', a pathway would need to be found that generates variants by moving around chunks that aren't represented as units in the individual representations of speakers.[12]

Crain et al. do not give actual examples of structure independence posing learning problems, but **the editors** mention the example of the savant Christopher and his difficulty learning a grammar of the made-up language Epun with its structure-independent rules. However, the interpretation of this data is so uncertain that it cannot be considered as definitive evidence either way – cognitively normal speakers, for example, can indeed learn Epun.[13] Further, analysis of child input together with learning simulations focused exactly on this issue, suggest that there is no reason to presume that structure-dependence is part of the initial state of the child learner (Reali and Christiansen, 2005) – these analyses come close to meeting **Crain et al.**,'s "a simple refutation of UG would be the finding that a human child can learn a structure-independent rule".

6. Hockettian design features

Several commentators (**Bolender**; **Longobardi and Roberts**; **Rothstein and Treves**), and **the editors**, find it odd that we exclude Hockettian 'design features' from our evaluation of what universals remain. Here we simply followed Joseph Greenberg (and the Konstanz Universals Archive)[14]: if questions of what is universal in language are to have any empirical content, they cannot at the same time be definitional of what language is, or their presence is merely tautological – a bit like 'discovering' that it is universal for even numbers to be divisible by two (of course, **the editors**' term $UG_{definitional}$ could encompass this by stipulation, but we think it is clear that the attraction of linguistic universals to the broader cognitive science community would be that they have actual empirical content).

The most mentioned property that we are accused of sweeping aside as a 'design feature' by commentators is discrete combinatoriality. We certainly agree with our commentators that every language exhibits discrete combinatoriality to a significant extent.[15] But then so do many animal communication systems, including those of many oscine and non-oscine birds (Leger, 2005; Bradbury and Vehrencamp, 1998). Even humpback whales do it (Suzuki et al., 2006). This is why Hockett's original breakdown into discreteness and productivity is more useful, together with his separation of duality of patterning. The commentators confuse many of these, together with recursion, under the label 'discrete combinatoriality'.

---

[11] Chomsky's (1968:61–62) original formulation was: 'grammatical transformations are invariably structure-dependent' – restated by **Crain et al.**, as constraints on types of movement rules. Since so many broadly generative approaches to syntax have developed alternatives to movement rules and transformations, the original formulation then has no application. However, interpreting it as broadly as possible, in terms of correspondence rules between semantically isomorphic structures, structure-dependency is violated by a range of grammatical phenomena such as the placement of heavy constituents in right position (Hawkins, 2004) or the ordering of elements by the animacy hierarchy (so that e.g. the grammaticality of Navajo morphology and word-order is determined by the relative animacy of arguments – see Croft, 1990:114). In the first case the determinant of placement is relative size (in terms of numbers of words or syllables), in the second it is the animacy of the referent (not a structural characteristic). Another, loose formulation of 'structure dependence' is that 'grammars don't count'. Though broadly true, there are interesting counterexamples, such as the retroflex *l*-allomorphs of the Warlpiri locative and ergative case suffixes after stems of three or more syllables as opposed to two or fewer (cf. wati-ngki 'man-ERG', maliki-rli 'dog-ERG'). To wave away such counterexamples, the formulation 'three or more' could be re-analyzed by constructing feet on alternating syllables, generating a stray third syllable which serves as the conditioning factor for the *l*-allomorph – but such an analysis comes at the price of developing a complex apparatus whose epicycles largely appear motivated by the need to evade falsification.

[12] Of course in a coevolutionary model, it becomes evident that there are really two claims here – (i) that such systems don't exist, translatable into diachronic terms as there being no possible pathways that could lead to their evolution, and (ii) that they would be unlearnable if they did exist, which is the mentalist side of the claim.

[13] See Newmeyer (2005:6–8) for discussion, who thinks it telling that different parts of the brain were involved in processing Epun structures (Smith and Tsimpli, 1995). Actually, it is well known that different languages, from Chinese to ASL, involve differential neural activation (see Valaki et al., 2004; Simos et al., 2005; Emmorey, 2002).

[14] "First, although the dividing line is sometimes difficult to draw, to qualify for inclusion in the archive universals ought to be substantive rather than being 'design features', whose universality is definitional (such as vocal-auditory channel, evanescence, duality of patterning; see Hockett, 1961: "lacking them, a communication system, instead of being considered a counterexample, would not be called a natural 'language'." (Plank and Filimonova, 2000).

[15] Though at the same time including significant portions of morphosyntax which run against this, from fusional case or verb-inflectional paradigms to idioms to non-compositional constructions.

Nonetheless, we agree that it may now be helpful to revisit some of these design properties, particularly in the light of the dynamic field of simulated evolutionary linguistics, since separating out these features allows us to simulate more plausible, gradualist, evolutionary scenarios where coevolving sociocultural systems (languages) and their biological platforms (human brains) spur each other on. Indeed we have already mentioned, under the rubric of emergent properties, simulations or learning experiments in which most of the Hockettian design features have been shown to evolve naturally under specific conditions. For example, it is quite plausible that the development of sophisticated intention-attribution allowed the development of a greatly expanded system of arbitrary signs – many initially holophrastic – from which discrete combinatoriality then developed as the emerging language system had greater expressive demands made on it (Jackendoff, 2002). In line with this gradualist move, in the BBS article we cited recent work by Zuidema and De Boer (2009) on how discrete combinatorial phonology – and thus duality of patterning – can emerge from iterative learning of a system that starts out non-discrete, under selective pressure to optimize perceptual distinctiveness.

Once a coevolutionary approach is adopted, as outlined in section 5 above, the presence of any universal feature calls for explanation in terms of two interacting components: the evolving semiotic system, and the evolving capacities of its users (including of course the neurocognition, as **Rothstein and Treves** remind us). Zuidema and Verhagen (2010) stress the importance of close modelling of transitions across stages, as structural properties are funneled back-and-forth between external sets of conventions (their E-language) and individual representations (their I-grammars). They point out (p. 59) that "compositionality in I-grammars can more easily evolve if some form of compositionality in the E-language has already been established for other reasons".

7. Linguistics and other fields

Linguistics, we have pointed out, occupies a crucial crossroads between the humanities, the social sciences and the biological sciences. Exploiting this position will be essential to the future of the discipline. **Abels and Neeleman** in their commentary offer us exactly the wrong way of doing this. They argue that any attempt to unify linguistics and cognitive science must follow the principle of 'methodological modularity'. (It is odd to find Boas's point about 'methodological relativity' of language systems, which A&N clearly do not subscribe to, recycled in this new guise.) The idea is that we must first have a developed linguistic science and an independent developed psychology and only then could we ask how the two articulate or can be unified. But this is quite odd: every linguistic principle must obey every relevant psychological law – language ability just *is* part of our psychological ability. We hardly need to unify what is already necessarily intrinsically unified.

In general there is a quite misplaced dismissive attitude discernable in generative linguistics: There have been cases where eminent C-linguists have been heard to say to top brain scientists "you are wrong, this is how it must work in the brain", or to leading geneticists "there must be a gene for Merge". Language is too fundamental a part of human biology, psychic and social life to be owned wholly by linguistics. It will be essential to try to understand what the other contributing sciences are trying to tell us – for example, that there are no (exclusively) language areas in the brain (Fuster, 2003), or that there can be no gene for grammar (Fisher and Marcus, 2006).

Linguistics has a unique position between disciplines for a number of reasons. As we argued in section 5 on Coevolution, language bridges the mental and the social, the psychological and the historical, the ideational and the behavioural. For that reason it touches on surrounding fields as diverse as psychology, the brain sciences, genetics, history, archaeology, anthropology and geography. Linguistic expertise can input important data and analysis into each of these fields (not to mention others from medicine to forensics). But to play this central role, the discipline needs to have a unifying framework in which we can articulate linguistic findings from the different realms, cognitive, behavioural and historical. That framework, we submit, is the coevolutionary framework sketched above.

The coevolutionary framework clearly enlists linguistics in the study of prehistory: it offers new tools and ways of thinking to historical linguistics (McMahon and McMahon, 2006), merges typology with historical linguistics to make a new field of evolutionary language structure (Dunn et al., 2008), and pushes the time barriers on linguistic reconstruction (Dunn et al., 2005; Pagel, 2009). But it also opens many new doors – for example, the possibility of studying grammaticalization or semantic change phylogenetically. Since evolutionary processes operate on variation, the kinds of datasets that sociolinguists and dialectologists have collected suddenly become of intense interest. Historical linguistics is the oldest and most conservative branch of linguistics, but so amplified, it would become even more of a key player, along with archaeology and anthropology, in the historical sciences.

C-linguistics has prided itself on its relation to the cognitive sciences. But it has been a difficult partner, as mentioned above: dismissive of alternative approaches, prone to internal theoretical dismantlements, obsessed with metalinguistic intuitions as data, and without serious interest in understanding linguistic diversity. D-linguistics, with its feet more firmly on the empirical ground, its interest in online experimental data and corpora, and its focus on variation, offers a better alternative, but it has lacked a unifying framework – again, a coevolutionary perspective offers exactly what is needed.

As we pointed out in our BBS Target Paper and Response (p. 480), the cognitive sciences are under radical revision. The idealization made at the outset of the movement, namely that the job is to characterize an invariant cognitive architecture (*the* human mind), has now gone out the window. Spurred by the breakthroughs in human genetics, the focus is now on variation in cognitive phenotypes which will allow us to pinpoint the genes that build our brains (see e.g. Papassotiropoulos et al., 2006). There is significant variation in language ability, as every teacher knows. The sources will be genetic, socio-

economic, multilingual backgrounds, and so forth. Superimpose on that substrate the strong social pressure for conformist language behaviour and it is clear that different individuals will use different acquisition, comprehension and production strategies to achieve passably uniform behaviour. Now superimpose on that the different kinds of parsing system required by different kinds of language – left-branching vs. right-branching, inflecting vs. isolating, constituency-based vs. free-word order – and we can see that we are dealing with variable cultural targets overlaying individual variation (see e.g. Valaki et al., 2004). Exactly the kind of material for coevolutionary investigations, and the central subject matter of what we can call the new psycholinguistics.

A number of commentators (**Bolender**; **Rothstein and Treves**; **Kemmerer and Eggleston**) make suggestions based on brain science. It is often thought by those outside the field that whatever happens in the brain happens by virtue of an innate, fixed neural circuitry. In fact, the assumption in the brain sciences is rather the opposite (Edelman, 1987), namely that brains are sculpted by experience, so that new experiences are reflected in a brain that remains remarkably plastic into adulthood. Consider reading: a new cultural art rewires the brain, leading to much increased white matter in the corpus callosum and other areas of the brain (Petersson et al., 2007; Carreiras et al., 2009). The reasons seem to be that the pre-existing connectivity of the brain 'attracts' letter-recognition to a specific area of the temporal lobe well-suited to small-scale pattern recognition (Dehaene, 2009). A purely visual task would handle the two visual fields in both hemispheres, but character recognition ends up lateralized in order to connect to previously lateralized language – hence the enlarged highway for data transmission across the hemispheres. Hooking up linguistic generalizations to brain circuitry requires in-depth knowledge of how both systems work, an appreciation of brain plasticity, and a willingness to revise opinions in the light of evidence from other disciplines.

Thus when **Bolender** suggests identifying UG with specific brain tissue he is completely off track. This would be a fundamental category mistake, mistaking the algorithm for the implementation as it were. A way to see this is to consider that we now know that individual humans process the same language in very different ways – for example, right-handed males tend to have strongly left-lateralized circuitry for language, but right-handed females are in two thirds of cases partially or strongly bilateralized (Catani et al., 2007). When we start to look at left-handers, and non-Western non-undergraduates, the circuitry patterns will show ever increasing diversity. In the other direction, we know that different languages tend to activate different brain circuitry, for example that Chinese comprehension is more bilateral than English or Spanish (Valaki et al., 2004), and the same is true for sign language (Emmorey, 2002; Newman et al., 2002).

Linguistic diversity offers wonderful opportunities to investigate this brain plasticity. As the most practiced human skill, one should expect to find many small but significant differences in brain structure and function across speech communities. The correlation of absolute pitch with speakers of tone languages is already a pointer in this direction (Deutsch et al., 2006). **Kemmerer and Eggleston** join us in calling for a new neuroscience that exploits this natural laboratory of human variation.

8. The future shape of linguistics

When **Reuland and Everaert** say dismissively "we have been here before", they are of course partly correct (see e.g. the sharp exchange between Comrie and Coopmans in the mid 1980s, nicely reviewed by Newmeyer, 2005:20–27). But as we said at the outset, the intellectual landscape has meanwhile fundamentally changed, and the question now is how linguistics as a discipline should respond to these new challenges and new opportunities.

We think it will prove highly detrimental to the discipline, even fatal in the case of many departments, if linguistics chooses to remain a wholly armchair discipline rooted exclusively in the humanities, with the central school based on linguistic intuitions about the familiar languages. All the evidence points to diminishing returns, and a scholastic fractionation of ideas, from this way of working. The alternative is to claim the high ground as the only discipline fully devoted to the study of language, and to try and encompass all the new lines of data, the new methods and the new insights coming from all the other disciplines that have increasing interest and expertise in the study of language. This requires fundamentally rethinking the training in our graduate programs, so that experiment, field work, statistical and computational expertise are all part of the normal package.

Take the example of the biologists who have moved into historical linguistics – they have demonstrated how powerful is the combination of explicit databases and new software developed to explore phylogeny and hybridization or contact (Gray and Atkinson, 2003; Pagel, 2009). A few linguists have now jumped at the opportunities opened up (McMahon and McMahon, 2006; Dunn et al., 2005, 2008; Reesink et al., 2009; Malchukov et al., in preparation), but most historical linguists retain a luddite stance on these developments even though they could easily pitch in and improve them. The question is whether linguistics is going to stand still while the biologists and psychologists invest in the great databases required to do this kind of work, or are we too going to view our discipline as based on cumulative, explicit data and the development of increasingly sophisticated tools to exploit it? Where linguists have made these investments, as with the WALS database, the explosion of new research based on new questions has been spectacular.

Moving forward requires a greater willingness to agree on common ground, so we can project as a discipline a consensual body of achievement. Of course there will be fierce disagreement about all sorts of issues, but we have a great body of work that we all take for granted – nobody doubts the value of cognates or shared innovations for establishing language relatedness, or the tests for immediate constituency, or the phonemic principle, even though we are all familiar with problems at the margins. What we need to project as a discipline is animated debate, but set within the context of solid, established, cumulative science. Otherwise, like cats squabbling over tidbits, we will lose out again and again in the

competition for big funding in the research councils. Stabilizing terminology would be a big step forward,[16] and so would abandoning terms (like UG) that are misleading to outsiders and have been polluted by too many divergent interpretations. More fundamentally, a greater willingness to engage with the surrounding disciplines would enormously help the fortunes of linguistics. There are many potential partners, looking for collaborators with deep linguistic expertise – in archaeology, psychology, brain sciences, anthropology, genetics. We hope, if only for the continued existence of the discipline, that the next generation of linguists will not disappoint them.

Acknowledgments

Appendix A. Further comments on the individual papers

We have dealt with many of the objections raised to the BBS paper in the main body of our paper above. However, some points could not be accommodated within that structure so we deal with them in this appendix.

A.1. *Comments on the editorial*

The exercise represented by this special issue would have had little point if there was not some modicum of greater light shed on the nature of language and the ongoing quarrels about that nature. We think there is some collective progress and we thank the editors and contributors for some enlightening discussion.

A.1.1. Overt vs. covert marking

The editors suggest that overt marking never constitutes a problem for the learner, and consequently overt patterns never need be coded in UG. This criterion would eject binding principles, grammatical categories, structure-dependence and much else once held dear by nativists from the garden of Eden, and indeed as referenced above many such patterns are emergent or extractable by statistical learning. One of the central lessons of the new information about linguistic diversity is that almost the whole of the vast structure of a language has somehow to be induced by the child – not a trivial task. While we agree that an overt pattern poses an overt puzzle for the learner (see Clark, 1993 on the principle of contrast), the puzzle may not be easy, as in inverse marking or complex paradigms. Here reference may usefully be made to another line of linguistic theorizing, which suggests that different levels of linguistic patterning are under differential conscious awareness (Silverstein, 1981). But it is important to remember that what is overt in one language may be covert in another. Who would think for example that languages would exist (like Hai//om) that mark their questions by an *absence* of an assertoric or imperative marker? Some languages mark their agents with case, others by order, others do neither. Some languages mark information structure with overt morphemes, others by position, yet others leave it to context.

A.1.2. Evolution

The editors suggest that the distinction FLB/FLN is basic, and moreover that FLN should consist only of Competence (i-Language properties). We really think it would be helpful to escape these Chomskyan dichotomies, which run counter to more productive ways of thinking. Take the issue of the genetic endowment for language. So far, the evidence for this is confined to the input/output systems, specifically to whatever genes code for the vocal apparatus and vocal learning (human alleles for FOXP2 appear to facilitate fine motor control for example). This is a central finding: it should form the heart of FLN/UG instead of being ruled out of court as mere sensory-motoric interface, or even Performance or e-Language. Put it another way: in communication, process is king; structure is a byproduct of processual factors. Why does our skeleton have the shape it does? Due to an interaction of the processes of inheritance and the processes of locomotion and action (see West-Eberhard, 2003 on the case of the two-legged goat). What is incompatible with a Darwinian perspective is to think that abstract structures have some kind of priority, or that knowledge exists for something other than use.

A.1.3. Recursion and minimalism

The editors suggest that recursion may well be inherited from the language of thought (LoT), in which case FLN is empty. We agree (see target BBS article, p. 443), while amplifying the reasons for locating it elsewhere by pointing out the existence of recursively embedded pair-parts in conversation, but distributed across speech-partners rather than within the syntax of one speaker (see response to **Koschmann** article below). Nonetheless, we would add a word of warning about the idea that

---

[16] Thanks are due to the lexicographers of the discipline, especially Crystal, Trask and Matthews, whose dictionaries of grammatical terminology have provided important anchors.

LoT is entirely sui generis. There are many reasons to think that some aspects of conceptual structure are inherited from language-specific categories (see Levinson, 2003a,b). And those who have studied humans deprived of early language acquisition (e.g. home signers) have found that abilities in understanding the recursion of embedded beliefs is at least partially due to the acquisition of the structure of embedding in syntax (De Villiers, 2000; Pyers and Senghas, 2009). The story is going to be complex: there is no fully specified universal mentalese that maps onto different languages, but rational thought and action is nevertheless available to all.

A.2. Comments on the other individual papers

A.2.1. **Duffield**

Duffield compares the BBS paper to some mixture of a freak show and a nature broadcast. But language diversity is indeed wondrous – there is nothing remotely like it elsewhere in the natural world – and a proper object of scientific enquiry. Duffield rightly takes us to task for talking as if grammatical relations might be primitives in a generativist model – they are defined of course over constituent structure. Yet the notion that they play a role in Universal Grammar is explicitly mentioned in Chomsky (1968), even if they are defined in terms of other substantive universals:

> 'Suppose that we define the relations subject-of as the relation holding between a noun phrase and a sentence of which it is an immediate constituent and the relation predicate-of as holding between a verb phrase and a sentence of which it is an immediate constituent and the relation predicate-of as holding between a verb phrase and a sentence of which it is an immediate constituent' (Chomsky, 1968:137)

> '[i]f the notions "noun phrase," "verb phrase," "sentence," "verb," can receive a language-independent characterization within universal grammar, then the grammatical relations defined above [SL&NE: sc. subject-of, predicate-of] will also receive a universal characterization.' (Chomsky, 1968:139)

But, as the BBS paper made clear, all these primitive notions out of which grammatical relations might be secondarily defined, such as verb phrase, and immediate constituency, are highly problematic from a cross-linguistic point of view. Thus grammatical relations either have to be taken as primitives (as in Relational Grammar) or abandoned as universal notions.

Duffield argues that UG, as the initial innate state of the faculty of language, need not be reflected in any way directly in the grammars of specific languages, and that we have therefore made a Rylean category mistake in thinking that grammars of existing languages can have any direct bearing on the content of UG. Freidin's comment in the BBS issue made a similar point about substantive universals: you can't falsify a constraint on rules by exhibiting any number of sentences, only by producing a grammar of rules that benefits from breaking it. We're like two bulls charging around a china shop while the matador is actually upstairs. From the point of view of Chomskyan exegesis, we think the comment is probably correct. But this immunity to counter-evidence is exactly what irks the non-Chomskyan.

Duffield goes on to make another theoretical point, namely, that the more diverse the grammars of the world, the more powerful the learning device (aka UG) must be. The identical point has been made in other learning domains and is generally valid, although caution is in order, because, as Christiansen and Chater (2008) point out, languages are adapted to our pre-existing learning abilities. However, our point all along was not that the human mind is a blank slate, but simply that there is little evidence (beyond the speech system and vocal learning) for a domain-specific learning machine for language. In modern terms, if FLN has no language-only componentry, the entire idea of a fixed architecture intrinsic to language would collapse.

A.2.2. **Crain, Khlentzos and Thornton**

In the conception of these authors (CK&T), UG says nothing about the constructions in a particular language, and has no account for "irregularities and exceptions", or diversity in general. It only characterizes the initial state of the child learner. Amongst the initial assumptions, suggest CK&T, are structure-dependence, but as we discussed above, statistical learning algorithms can reproduce structure-dependence without any special intrinsic machinery. The particular case brought up by CK&T, auxiliary fronting in English polar questions, has been the precise target of one such demonstration, where given samples of actual child input, a learning algorithm correctly simulated the output (Reali and Christiansen, 2005).

Another property that CK&T impute to UG is downwards entailment, the reversal of the inference from 'John bought a Prius' → 'John bought a car', in negative or conditional contexts, as in 'John didn't buy a car' → 'He didn't buy a Prius'. This brings up the whole question of the relation of logical or rational inference and language. We take it as evident that, for a language to serve its speakers well enough as an instrument of communication so that their culture survives and prospers, it should be consistent with the nature of the external world, for which logic is curiously a useful guide (as a matter of fact, no car, no Prius). Downwards entailment is a property of logic, not language – so it is no more surprising that all languages tend to reflect logical principles than it is to find that there are no languages whose number systems defy the rules of standard addition.

This is not the same as saying that there is a single, determinate solution to the problems of evolving linguistic structures to encode (or rather partially reflect) logic or mathematics. We can illustrate this with the very case of Japanese disjunction which CK&T take to support a parameter-based UG account (a model that is now completely abandoned by many if not most

C-linguists). For a start, conjunction/disjunction in Japanese differs from English (and most other languages of the world, here including Russian) not simply in whether disjunction means exclusive vs. inclusive 'or', but also on a dimension of conjunctive and disjunctive exhaustivity: 'John and Mary' in English 'John and Mary came' will be translated into Japanese as *John to Mary* if it refers just to John and Mary, but as *John ya Mary* if others are also included though not mentioned. Exhaustive vs. non-exhaustive disjunction can be expressed in English by prosody but is not lexicalized – cf. the well-known example 'Do you want tea or coffee ↑? [i.e. or nothing, or anything else]' vs. 'Do you want tea, or coffee↓?' [i.e. a reading which exhausts the choice set].

CK&T. state that 'disjunction words always mean inclusive-or'. But this is not an accurate translation of *ka* – which in fact is an optional reduction of the fuller construction X *ka* Y *ka* which is more comparable in its semantics to English *either . . . or*, as its fuller bipartite structure suggests. Though CK&T dismiss the possibility that this means exclusive-'or', their characterization differs from standard reference works on Japanese, e.g. Makino and Tsutsui (1986:164), whose examples are consistently translated with 'either . . . or'. There are many ways of accounting for the differences between Japanese children and adults in the conclusions they draw from negative clauses with *ka* disjuncts, including the possibility that exclusive disjunction (at least if not accompanied by prosodic or segmental emphasis on both alternatives) is a more semantically marked or derivative category (as reflected in the fact that it is not taken as a primitive in logic, whereas inclusive 'or' is). To attribute the differences to UG-style parameter setting is quite unnecessary – there is no possibility whatsoever that parameters of this level of specification could be encoded in the genome. Besides, as Newmeyer (2005:126) succinctly puts the current majority view, "It is not the job of generative grammar to account for typological generalizations. Attempts to do so have been failures."

A.2.3. **Reuland and Everaert**

R&E think we "strongly oppose" any search for regularities. That is a complete misreading of our enterprise. What we said was "In short, it has proven extremely hard to come up with even quite abstract generalizations that don't run afoul of the cross-linguistic facts. This doesn't mean that such generalizations won't ultimately be found. . .", and indeed we could hardly disagree with the proposition that science is in the business of generalizing. The crucial question is where the generalizations are to be found. What we suggested was that, just as in evolutionary biology, the generalizations are likely to be found in the processes that generate structures, rather than the structures themselves, with complex interactions between systems (biological, cognitive, cultural) over time – and not in a dedicated faculty of language.

R&E accuse us of being out of date with respect to developments in Binding theory. Quite possibly, but the account they outline is complex and quite unconvincing. And once again, here we are back in the world of detailed parameters, which no longer looks like an attractive explanatory methodology to most linguists. We have already mentioned that we don't buy invisible pronoun doubling or other impositions of structure on data which look designed simply to preserve a theory. In addition, there are pragmatic accounts (as in Levinson, 2000, mentioned only as a data source) which seem to offer a very straightforward account of why recurrent binding patterns would occur in the languages of the world, and how they could be learnt by the child.

All the complex talk about parametric variation, invented to deal with the linguistic diversity of course, is intended to preserve the underlying theory from counterexamples. But the underlying theory, that of the Binding Principles, is itself not of the right kind to have universal application. The Binding Principles presume universal constituency – they require a structure in which subjects can command objects in a tree structure. But we argued that there are numerous languages where this structural dominance will not be found – first, in languages without constituency as the major organizing principle of clause structure, second in the much larger class of languages without a VP constituent, and third, in languages where the dominance relation of subject to anaphor is reversed. Since we wrote the BBS article, Keenan (2009, and see his response here with Stabler) has produced some well attested examples of languages where the anaphors command their coreferential NPs, which are deeply problematic for the standard theory.

The issues of binding straddle the syntax/semantics border, and they cannot be satisfactorily resolved by arguing about a few examples taken from grammars, or even examining intuitions. Reflexivity and reciprocity is actually a grammatical domain that is reasonably easy to investigate empirically – see e.g. Evans et al. (in press), where a series of video stimuli were used to investigate the structures and meanings of reciprocal expressions across twenty languages, an exercise that verified that some languages really don't have anaphors, and also flushed out constructions which have not been discussed in the literature. The way forward is to have much more serious data to constrain the theoretical imagination.

A.2.4. **Safir**

We welcome Safir's relatively jargon-free rehearsal of the kinds of observations that motivate constraints on rules like subjacency, which forbid movement over two or more bounding nodes of specified kinds. Following Rizzi, he suggests that the exceptions to subjacency in Italian are easily explained by positing a hidden resumptive pronoun (resumptive pronouns have long been noted to promiscuously license long-range movement). For example, in English the following sentence may be judged ungrammatical without the bold pronoun and perhaps marginally acceptable with it – in the same way all Italian counterexamples would be held to actually have a similar *covert* pronoun:

"Do you remember the men who Mary said the dinner which (**they**) were eating was disgusting?"

Safir goes on to suggest that we (E&L) find ourselves in the horns of a dilemma. Either we reject this elegant analysis on the grounds that it is too abstract, or we would have to buy into the whole Minimalist apparatus as part of our innate linguistic endowment. In fact, we don't necessarily feel obliged to take either horn. Following Newmeyer (2005), we think

that the language-specific instantiations of subjacency required to handle even a handful of languages undermine the elegance of the whole intellectual move. There's no point in positing a universal rule subject to myriad language-specific exemptions. In particular, the invisible resumptive pronoun argument has given rise to a whole range of very peculiar claims about other languages. Take the fact that Palauan provides no evidence whatsoever of subjacency or any island constraints on movement – Georgopoulos therefore argues that every variable (pronominal or gap) in Palauan is a (frequently covert) resumptive pronoun (Georgopoulos, 1991)[17] Frankly, this simply looks like an escape hatch for a theory in trouble.

Perhaps even more disturbing is the way Kayardild is dismissed as a counterexample to unbounded recursion. In our BBS paper we cited Kayardild as an example of a language where embedding is confined to one level, and cited an interesting morphological reason for this – that subordinate clauses require the 'complementizing oblique', but that, since this cannot be followed by further case marking (against the general tendency in Kayardild to allow stacking of case suffixes), this caps the level of complexity at two clauses. Safir goes on to suggest that "it would be thoroughly compatible with their analysis of the facts to posit a role for operations that generate recursion in this language, but the effects of it are obscured by independent factors". Moves like this are deeply revealing of the casualness with which claims are jollied out of the way of falsifiability by UG-practitioners. Returning to the Boasian emic postulate, we cannot see what is gained by postulating mechanisms that are never detectable.

A.2.5. *Keenan and Stabler*

As mentioned above, the approach K&S advocate has kinship with the Greenbergian idea that the basis for typological comparison is ultimately semantic. A crucial question for this approach to Binding and Anaphors is the kind of example that **Reuland and Everaert** adduce, namely the structure in Fijian (glossing 'Mike criticized him') which allows, but does not require, a reflexive reading, or similarly with verbal derivations which are general over reflexive or reciprocal or other readings. K&S's argument is complex, and if we understand them correctly, we think that in these cases K&S must posit both semantic and structural ambiguities, because their universals do not seem to allow them to capture semantic generality. Nevertheless, many natural language expressions simply are semantically general (cf. *aunt* general over mother's brother's wife and mother's sister, etc.), and are resolved pragmatically, e.g. by Gricean means. This makes it entirely conceivable that some languages handle reflexivity in this way (see Levinson, 2000 for Gricean applications in the reflexive domain).

We thank K&S for usefully pointing out that the failures of the Chomskyan approaches need not be blamed on over-abstraction, for such approaches may in some respects be not nearly abstract enough to capture the observable linguistic diversity.

A.2.6. *Abels and Neeleman*

A&L equate progress with the degree to which the possibility space for languages is restricted. This elementary Popperianism sounds fine, until one actually tries to interpret it. Suppose we could build a model of grammar that permits just and only the existent natural languages, which we know to be a poor sample of pre-colonial diversity. The model would then be unable to capture the rest of the half-million odd languages that have preceded them. Moreover, the theory might turn out to be just as much a model of cultural evolution as a model of psychic capacity, because, as we pointed out in the BBS response (p. 477), all spoken languages must ultimately be presumed to share common descent. If, as seems not unlikely, that in 100 years accelerating language loss leaves us with just, say, 200 languages, the exercise by then would clearly be absurd.

If we are interested in *cognitive constraints* on possible languages, we will find these in such plodding performance factors as short-term memory, long-term memory, attention and executive control (see e.g. Hawkins, 2004) – the sorts of factors central in psycholinguistics and the brain sciences. Spoken languages have a peculiar nature: they are linear in an evanescent medium. This puts tremendous pressure on both articulation and comprehension – since the latter can run about four times faster than production, one universal of language (but not a universal of linguistic structure) is that there will always be more left to inference than what is coded (Levinson, 2000:28).

A.2.7. *Bolender*

Bolender proposes to resuscitate UG by identifying it with Hockettian design principles, like duality of patterning. Moreover, he suggests, this particulate recombinatorial system may lie behind the higher cognitive capacities unique to humans. He sheds some doubt on the significance of the apparent lack of duality of patterning in the Al Sayyid sign language, suggesting that it might result from lack of exposure in a very small community. For the record, this does not seem to be the case: the language has been around for three or more generations, has c. 150 deaf users, and perhaps as much as a thousand hearing second-language signers (Sandler et al., 2005).

We agree that the interrelations between higher cognition and language ability offer intriguing possibilities for research. A number of lines of research all have important bearing on this issue: (a) the neo-Whorfian exploration of language-specific categories on human thinking (see e.g. Gentner and Goldin-Meadow, 2003; Levinson, 2003a,b; Boroditsky, 2009), mentioned by Bolender; (b) the possible limits on human thinking imposed by the linguistic limitations of recently developed languages and home-sign (see Schick et al., 2007; Pyers and Senghas, 2009; Goldin-Meadow, 2003); (c) the mental capacities of pre-

---

[17] As a reviewer of Georgopoulos' book pointed out (Durie, 1993) this doctoral thesis was overtaken by a change in the theory of Barriers that undermines the model anyway.

linguistic infants (see e.g. Liszkowski et al., 2009); (d) the mental capacities of non-linguistic species (see e.g. Call and Tomasello, 2008). Clearly this is going to be a chicken-and-egg story: infants need a rich endowment of cooperative and mind-reading abilities to bootstrap themselves into any human linguistic system; on the other hand, the particular system they are surrounded by may have significant effects on their cognition.

A.2.8. *Lee, Lee, Gordon and Hendrik*

LLG&H provide a useful reminder that linguistic diversity in form predicts processing difficulties of different kinds, according to the nature of the language. Nevertheless, they argue, the underlying cognitive bottlenecks are the same: the costs of processing ambiguity on the one hand, and holding NPs in memory before unification with the verb they are arguments of, on the other. They illustrate this in detail by contrasting the processing difficulties of relative clauses in English, Korean and Chinese, summarising findings of great significance for linguists interested in the differences between languages.

It is not entirely clear what sin LLG&H think E&L have committed, except to have been overly concerned that psycholinguistics might be led astray by the current preoccupations of generative linguistics. Historically, LLG&H are partly right: Psycholinguistics and generative theory parted company in all matters of detail, first during the 1960s when transformational models failed to predict complexity of processing, and decisively during the Linguistic Wars (Newmeyer, 1986), when it became hard to know whose models to follow. But Chomsky's overall abstract programme has nevertheless been taken as a guiding framework by many cognitive scientists, to the extant that we have become recently concerned by the vast efforts devoted to tracking down recursive abilities in man and beast – we very much doubt this will shed much light on language. Apart from this kind of distraction, we agree that psycholinguistics as an empirical discipline soldiers on, usually in glorious detachment from linguistic theory.

But is this a satisfactory state of affairs? Isn't it a pity that, apart from the kind of work that LLG&H report on, there is so little systematic use of the vast variety of languages to test the kinds of theories that psycholinguists make? That is the point we made in the BBS article (e.g. Box 3 p. 447). It is so striking that no sustained effort has been put into understanding the processing differences between, for example, verb-final and verb-initial languages – verb-final languages pushing new information crucial to resolution to the very end of the clause, while verb-initial languages put it all up front. Here one may be trading off production vs. comprehension costs: verb-initial languages require the entire clause to be worked out before speaking, but offer smooth comprehension, verb-final languages allow the speaker to make it up as he goes along, but impose severe memory loads on comprehenders.

In short, we think a new union between linguistics and psycholinguistics could be born by a joint undertaking which exploits the natural variation between languages to explore the range of processing mechanisms and the trade-offs that languages make between them. That was why we talked about the search for "stable engineering solutions satisfying multiple design constraints", and the possibility of explaining both the current linguistic diversity and the nature of future and past language change in terms of such constraints.

A.2.9. *O'Grady*

Both Chomsky and Greenberg discount external functional factors in the search for universals – Greenberg because these require no real search for systems universals, and Chomsky because such external factors would not motivate a language-special innate mechanism. So for neither theorist would the fact that language users have ways to request actions or specific information be of much interest, for example.

O'Grady makes the important point that functional factors, if brought into the picture, might motivate across the board strong universals. This gives useful flesh to the BBS response by Haspelmath. He gives two examples: (1) voiceless stops deplete oral pressure, so can only be maintained at a cost – they are the marked option for physical reasons, (2) filler gap-dependencies exercise working memory, so the longer the distance, the greater the cognitive load. These should not be thought about, however, as laws of language – rather, they are facts whose origin lies outside the linguistic system proper, but constrain its properties. Languages cannot circumvent them, they can only push the limits to varying degrees.

A.2.10. *Hudson*

Hudson argues that anything you can do with constituency, you can do just as well with dependency. Moreover, he continues, dependency analyses have a striking advantage: they allow a direct relation between words, without unnecessary mediation of structure, and this direct relation between entities is manifested everywhere else in cognition (for example in kin relations). This renders basic clausal structure in a familiar language like English into garden variety cognitive relations, and obviates the need for language-only structures and relations.

Hudson goes on to claim that dependency relations are, by virtue of their domain-general nature, universally available, and are always employed by languages. But he is at pains to make clear that this does not amount to a reincarnation of UG based on replacement of constituency with dependency, because dependency has none of the domain specificity claimed for constituency. He sees one merit of the dependency approach in the natural account it provides of the cognitive costs of long-distance dependencies, echoing here exactly some of the thoughts voiced by O'Grady.

Replacing constituency with dependency models even for languages like English is quite radical – it breaks with the best part of a hundred years of American structuralism, with its emphasis on substitution classes. Unlike Hudson, we would not want to use dependency to replace constituency-based models everywhere. Firstly, there are some phenomena which are difficult to deal with by dependency (such as conjunction, as Hudson admits). Secondly, getting rid of constituency

altogether makes it difficult to account for phenomena like the 'coalescing first constituent' in Warlpiri – where independent elements of a phrase have the option of being grouped together in first position and just in such cases count as a single constituent (most notably for purposes of auxiliary placement), even though most of Warlpiri grammar is better captured by dependency. (In LFG representations this means that the ratio of f-structure to c-structure rules is much higher in such languages.) Finally, having both these representational methods allows us to capture an important dimension of cross-linguistic variation which would be concealed if dependency methods were generalized across the board.

Meanwhile, this radical revision of the foundations of syntactic theory has much to recommend it, and Hudson is not the only protagonist who argues for it, as he shows.

A.2.11. **Sandler**

Sign languages have come to form a specialism in linguistics, and experts on spoken language often know very little about such languages. We therefore welcome the three comments from this angle. Sandler argues two points; First, there are striking uniformities in the organization of both spoken and signed languages on an architectural level – in e.g. having organized timing units like syllables[18] and syntactic embedding; Second, on nearly every mode of implementation, phonological and grammatical properties contrast markedly across the two modalities. She thinks we are in danger of throwing out the baby (language uniformities) with the bathwater (a previous undervaluation of diversity).

Sandler defines UG in the restrictive sense, of a domain-specific innate knowledge of language, and argues that sign languages thoroughly undermine this notion, since sign languages clearly argue for an 'embodied' cognition centred around the visual/motor modality. Nevertheless, she argues, there are striking commonalities between spoken and signed languages, pointing to the existence of a phonology with similar abstract organization in both spoken and signed modalities. The comment ends however with the observation that new sign languages after four or five generations may still not exhibit such fundamental phonological properties as 'duality of patterning'. These new languages raise the question of how far phonology is an emergent, self-organizing property of languages. Here reference should be made (as cited above) to work that has simulated exactly this kind of emergence both in terms of computational models of populations evolving shared communication systems, and in terms of experimental recursive learning by humans of novel communication systems. This simulation and experimental work suggests that commonalities in languages may arise from the nature of shared communication systems passing through minds that seek pattern and regularity – little more need be presupposed.

A.2.12. **Cormier, Schembri and Woll**

CS&W join **Sandler** in arguing that sign languages are crucial in assessing the evidence for strong innate language universals. They suggest that the UG-coloured spectacles through which sign languages have often been viewed is a natural historical concomitant of the struggle to have sign languages viewed as proper, fully expressive languages, equivalent to spoken ones. They take issue with a couple of remarks in the BBS paper, one in which we suggested that pronouns in sign may be gestural, and another in which we (inadvertently) implied that sign language classifiers are the same kind of thing as classifiers in spoken languages. As regards pronouns, they are surely correct that though gestural in origin pronouns may have lexicalized and grammaticalized to varying degrees in different sign languages (see also comments by **Malaia and Wilbur**). Classifiers in sign languages are clearly sui generis, and although most sign languages exhibit something like them, some village sign languages seem not to (see e.g. Nyst, 2007). The issue here is the same as that made by **Kemmerer and Eggleston** – namely that distribution classes of any category are not only language-specific, but construction-specific within languages. Linguistic diversity is fractal, all the way down.

An important point made by CS&W is that language and gesture should be seen as a single package, since they intertwine in intricate ways both in spoken but especially sign languages. Indeed, we pointed out on p. 433 of the target article a likely common evolutionary origin: "An alternative view is that language evolved from a modality-hybrid communication system in which hand and mouth both participated, as they do today in both spoken and signed languages (cf. Sandler, 2009)". We believe this is an important corrective perspective, although there is little doubt that in hearing populations the spoken language becomes the dominant partner.

A.2.13. **Malaia and Wilbur**

M&W make many of the same points made by **Sandler** and **CS&W**, namely that examination of specific signs shows that they are more intricately grammaticalized than at first appears. They go on to argue for 'abstract constructs' from generative approaches as offering both cross-linguistic generalizations and hooks to neuroscience. But the example they parade, about 'change of state' marking by abrupt deceleration in sign does not trace its pedigree to generative approaches. Rather, it derives from the Vendler tradition rooted in natural language philosophy, and plays a role in much descriptive work on grammar and semantics (see e.g. Kita, 2006). Descriptive approaches to gesture (never studied in the generative tradition) are also fully sensitive to accelerative features, since this is what isolates the so-called 'stroke' of a gesture. Behind the remarks lurks a misinterpretation of the descriptive enterprise as hooked only to 'minimal constituent units' – abstract generalizations are just as available to these analytic approaches, but the preference is for inductive emergence of the categories from the data rather than application of pre-existing categories from a generative model.

---

[18] Though we do not endorse her view that syllables are 'another language universal' – see Hyman (submitted for publication) on Gokana as a language where the syllable is extraordinarily hard to justify as a unit.

A.2.14. *Longobardi and Roberts*

L&R echo a number of critiques that have come up elsewhere (and were partly responded to in our BBS response), but are worth responding to here directly.

(1) Why aren't markedness universals real and interesting universals? L&R give the example of nasal vowels implying oral vowels, and dual implying plural. Intuition suggests markedness universals have a plodding character – does the following sound like an interesting biological universal to you: all animals that may move fast, may also move slowly? Or the following an important insight into mathematical cognition: Anyone who can add 131 and 493 can also add 2 + 2? Put simply, the possession of the more complex, more secondary trait, implies the possession of the simpler, more primary trait. Interestingly, if the sample is large enough, these generalizations are usually counter-exemplified, as we showed with the Nen counterexample to the dual → plural inference, and as shown by the existence of systems with only voiceless stops (see **O'Grady's** comment; see also Haspelmath, 2006, on the diverse meanings of markedness).
(2) E&L allegedly reject abstract analyses. What we actually said, and L&R quote us, is that there are attendant dangers. We have no quarrel with abstract analyses per se, but we would like to see these arise inductively, and not be derived deductively from a model based on English and familiar languages.
(3) Constituency. It is of course trivial to generate a free-word order language from a fixed base with some scrambling or movement rules (although not without the cost of massively multiplying language-specific parameters, as suggested here). The point we wanted to make is that there is no psychological evidence for such an imposition, and it won't help you parse the sentence. Parsing will have to be based on entirely separate principles. The history of generative grammar is littered with implausible underlying word orders proposed for many languages, invented solely to preserve a theory that doesn't match the facts (Newmeyer, 2005). More plausible then to go the dependency route, as sketched in the commentary by **Hudson**.
(4) Diversity and parameters. L&R think we were right to emphasize diversity, but wrong to think that it cannot be handled by a sophisticated theory of parameters. At the risk of going over the same ground again, let us just point to the critique of parametric models by Newmeyer (2005). His point is simply that by the time we have micro-parameterized the lexicon in a language, we have done nothing to help the child learn it – if the child can figure all that out, it can figure out the rest without UG. Most D-linguists assume that parameters have gone the way of generalized transformations, the A-over-A principle, and countless other technical tools, onto the ever increasing scrap heap of generative grammar.

A.2.15. *Koschmann*

Koschmann notes that recursion can be found in discourse structure, as in:

Q: How much is that?

    Q: You mean this big one?

    A: Yes

A: Twenty six ninety nine

This important point was also made by Levinson (1981, 2006), who points out that such recursive embedding is typical only of cooperative interaction systems. What is striking is that in discourse one can easily find centre-embeddings four or more applications deep (see Levinson, 1983:305–306), a degree of centre-embedding that is vanishingly rare in naturally occurring language sentences. This is clearly an example of the computational advantages of 'distributed cognition', a task divided over two minds. We agree with Koschmann that embedded recursion in dialogue, which is almost certainly universal, is an interesting and relevant phenomenon. Notice that we can do this non-linguistically: If you are standing near the coat rack, I can hold out my hand requesting my coat, you can lift a candidate coat waiting for agreement, and I can nod my agreement, getting my coat. Further centre-embedded recursions are even possible: when you lift up the candidate coat, I indicate that I can't quite make out the crucial buttons, so you hold those up for me to check, and so on. Recursion is thus not wholly a linguistic affair, and certainly not confined to the gymnastics of the clause.

A.2.16. *Rothstein and Treves*

R&T point correctly to the central importance of compositionality and ask what are the underlying computational problems and how are they implemented in neural circuitry. As it happens, there is an existing well worked out model by Hagoort and colleagues which talks to the neural base of these operations under the rubric of 'unification' (Hagoort, 2005). They suggest a model where tree fragments are recovered from the temporal lobe and shunted to (roughly) Broca's area, where the fragments are unified in parallel with meaning composition in the same area (the centre of gravity for syntactic unification and semantic unification are just slightly different). Hagoort and his colleagues have found that information in all channels, including gesture, is incrementally 'unified' with a presumed semantic structure almost instantly, indicating that a running discourse model rather than a clause-level process is involved. These procedures look rather different than those presumed by many linguists: there's a 'lexicon' of tree fragments on the syntactic side (as in e.g. tree-adjoining grammar), and no preference for clause-bound material over discourse and gesture information on the semantic side. The modality-neutral nature of the unification suggests, as in the case of reading, that such circuitry is built in infancy and attracted to Broca's area by nature of the strong connections that it has to other areas involved in language processing.

A.2.17. **Kemmerer and Eggleston**

K&E elaborate a point we made in the BBS article, and explore the implications of language diversity for cognitive neuroscience. They take up the issue of categories, and specifically whether Noun vs. Verb constitute universal grammatical categories. They argue cogently that not only will no languages coincide in their grammatical categories, but also that, within a language, grammatical categories always internally fractionate into myriads of detailed substitution classes. Nevertheless, they suggest (following Croft, 2001) that there are prototype attractors in the distinctions between objects and actions on the one hand, and reference and predication on the other. Moreover, there is clear specialization for objects and actions in temporal cortex and motor cortex respectively. However, how these cortical regions play out in less prototoypical members of categories is not known. And what is known is based almost entirely on familiar languages. This severely limits any general claims that the neurocognition of language is able to make about the underlying, language-independent neural representation of word classes. K&E join us in calling for a 'new neurocognition of language' that is typologically informed.

References

Abels, K., Neeleman, A., 2010. Nihilism masquerading as progress. Lingua 120, 2657–2660.

Ameka, F., Levinson, S.C. (Eds.), 2007. The Typology and Semantics of Locative Predicates: Posturals, Positionals and other Beasts. Positionals. The State of the Art Special Issue of Linguistics 45 (5/6).

Bannard, C., Lieven, E., Tomasello, M., 2009. Modeling children's early grammatical knowledge. Proceedings of the National Academy of Sciences 106 (41), 17284–17289.

Barbiers, S., Auwera, J., van der, Bennis, H., Boef, E., de Vogelaer, G., van der Ham, M. (Eds.), 2008. Syntactische atlas van de Nederlandse dialecten, deel 1. Amsterdam University Press, Amsterdam.

Bickel, B., 2009. Typological patterns and hidden diversity. In: Plenary Talk, 8th Association for Linguistic Typology Conference, Berkeley, CA, http://www.unileipzig.de/~bickel/research/presentations/alt2009bickel-plenary.html.

Bickerton, D., 1976. Dynamics of a Creole System. Cambridge University Press, New York.

Boas, F., 1911. Handbook of American Indian languages, vol. 1. Bureau of American Ethnology, Bulletin 40. Government Print Office, Washington.

Boë, L.-J., Heim, J.L., Honda, K., Maeda, S., 2002. The potential Neanderthal vowel space was as large as that of modern humans. Journal of Phonetics 30 (3), 465–484.

de Boer, B., Fitch, W.T., 2010. Computer models of vocal tract evolution: an overview and critique. Adaptive Behavior 18, 36–47, doi:10.1177/1059712309350972.

Bolender, 2010. Universal grammar as more than a programmatic label. Lingua 120, 2661–2663.

Boroditsky, L., 2009. How does our language shape the way we think? In: Brockman, M. (Ed.), What's Next? Dispatches on the Future of Science. Vintage Press, New York.

Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. Chicago University Press, Chicago.

Bradbury, J.W., Vehrencamp, S.L., 1998. Principles of Animal Communication. Sinauer Associates, Sunderland, MA.

Bresnan, J., Cueni, A., Nikitina, T., Baayen, R.H., 2007. Predicting the dative alternation. In: Bouma, G., Kraemer, I., Zwarts, J. (Eds.), Cognitive Foundations of Interpretation. Royal Netherlands Academy of Arts and Sciences, Amsterdam, pp. 69–94.

Call, J., Tomasello, M., 2008. Does the chimpanzee have a theory of mind? 30 years later. Trends in Cognitive Science 12, 187–192.

Carreiras, M., Seghier, M.L., Baquero, S., Estévez, A., Lozano, A., Devlin, J.T., Price, C.J., 2009. An anatomical signature for literacy. Nature 461, 983–986, doi:10.1038/nature08461.

Catani, M., Allin, P.G.A., Husain, M., Pugliese, L., Mesulam, M.M., Murray, R.M., Jones, D.K., 2007. Symmetries in human brain pathways predict verbal recall. Proceedings of the National Academy of Sciences 104 (43), 17163–17168.

Chater, N., Christiansen, M.H., 2009. Language acquisition meets language evolution. Cognitive Science 1–27, doi:10.1111/j.1551-6709.2009.01049.x.

Chomsky, N., 1965. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA.

Chomsky, N., 1968. Language and Mind. Harcourt Brace & World, Inc., New York.

Chomsky, N., 1977. Essays on Form and Interpretation. North-Holland/Elsevier, Amsterdam.

Chomsky, N., 1986. Knowledge of Language. Praeger, New York.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.

Chomsky, N., 2007. Of minds and language. Biolinguistics 1, 1009–1027.

Chomsky, N., 2010. Some simple evo devo theses. How true might they be for language? In: Larson, R.K., Deprez, V.M., Yamakido, H. (Eds.), Approaches to the Evolution of Language. Cambridge University Press, Cambridge, pp. 45–62.

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

Clark, E., 1993. The Lexicon in Acquisition. Cambridge University Press, Cambridge.

Coombs, C., 1964. A Theory of Data. John Wiley, New York.

Corbett, G., 2006. Agreement. Cambridge University Press, New York.

Cormier, K., Schembri, A., Woll, B., 2010. Diversity across sign languages and spoken languages: implications for language universals. Lingua 120, 2664–2667.

Crain, S., Khlentzos, D., Thornton, D., 2010. Universal Grammar versus language diversity. Lingua 120, 2668–2672.

Croft, W.A., 1990. Typology and Universals. Cambridge University Press, Cambridge.

Croft, W.A., 2000. Explaining Language Change. Longmans, London.

Croft, W.A., 2001. Radical Construction Grammar: Syntactic Theory in Typological Perspective. Oxford University Press, Oxford.

Dąbrowska, E., 2010. Naive vs. expert intuitions: an empirical study of acceptability judgments. The Linguistic Review 27 (1), 1–23.

Saussure, F. de, 1979. Cours de linguistique générale. Paris: Payot. Édition critique préparée par Tullio de Mauro. Payot, Paris.

De Villiers, J., 2000. Language and theory of mind: what are the developmental relationships? In: Baron-Cohen, S., Tager-Flusberg, H., Cohen, D.J. (Eds.), Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience. 2nd ed. Oxford University Press, New York, pp. 83–123.

Dediu, D., Ladd, 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 104, 10944–10949.

Dehaene, S., 2009. Reading in the Brain. Viking, New York.

Del Giudice, A., Kirby, S., Padden, C., 2010. Recreating duality of patterning in the laboratory: a new experimental paradigm for studying emergence of sublexical structure. EVOLANG 2010.

Deutsch, D., Henthorn, T., Marvin, E., Xu, H.-S., 2006. Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. Journal of the Acoustical Society of America 119 (2), 719–722.

Dixon, R.M.W., 2009. Basic Linguistic Theory Volume 1: Methodology. Oxford University Press, Oxford.

Dobrin, L.M., 1995. Theoretical consequences of literal alliterative concord. In: Papers from the 31st Regional Meeting of the Chicago Linguistic Society. Chicago Linguistic Society, Chicago.

Dryer, M., 1997. Are grammatical relations universal? In: Bybee, J., Haiman, J., Thompson, S.A. (Eds.), Essays on Language Function and Language Type. Benjamins, Amsterdam, pp. 115–143.

Dryer, M., 2006. Descriptive theories, explanatory theories, and Basic Linguistic Theory. In: Ameka, F., Dench, A., Evans, N. (Eds.), Catching Language: The Standing Challenge of Grammar Writing. Mouton de Gruyter, Berlin, pp. 207–234.

Duffield, N., 2010. Roll up for the Mystery Tour! Lingua 120, 2673–2675.

Durie, M., 1993. Review of Carol Georgopoulos: syntactic variables: resumptive pronouns and A'-binding in Palauan (Studies in Natural Language and Linguistic Theory 24). Journal of Linguistics 29 (2), 512–517.

Dunn, M., Greenhill, S.J., Levinson, S.C., Gray, R.D., submitted for publication. Phylogenetic trees reveal lineage specific trends in the evolved structure of language.

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

Dunn, M., Terrill, A., Reesink, G., Foley, R.A., Levinson, S.C., 2005. Structural phylogenetics and the reconstruction of ancient language history. Science 309, 2072–2075.

Durham, W., 1991. Coevolution. Stanford University Press, Stanford, CA.

Edelman, G., 1987. Neural Darwinism – The Theory of Neuronal Group Selection. Basic Books, New York.

Editors, 2010. Editorial introduction to the special issue of Lingua on Evans & Levinson's 'The myth of language universals'. Lingua 120, 2651–2656

Emmorey, K., 2002. Language, Cognition, and the Brain: Insights from Sign Language Research. Lawrence Erlbaum, Mahwah, NJ.

Enfield, N., Levinson, S.C. (Eds.), 2006. Roots of Human Sociality. Culture, Cognition and Human Interaction. Berg Publishers, Oxford.

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

Evans, N., Dench, A., 2006. Introduction: catching language. In: Ameka, F., Dench, A., Evans, N. (Eds.), Catching Language: The Standing Challenge of Grammar Writing. Mouton de Gruyter, Berlin, pp. 1–40.

Evans, N., Levinson, S.C., 2009a. The myth of language universals: language diversity and its importance for cognitive Science. Behavioral and Brain Sciences 32 (5), 429–492.

Evans, N., Levinson, S.C., 2009b. With diversity in mind: freeing the language sciences from Universal Grammar. Behavioral and Brain Sciences 32 (5), 472–484.

Evans, N., Gaby, A., Levinson, S.C., Majid, A. (Eds.), in press. Reciprocals and Semantic Typology. John Benjamins, Amsterdam.

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

Freidin, R., 2009. A note on methodology in linguistics. Commentary on Evans & Levinson, The myth of language universals. Brain and Behavioral Sciences 32 (5), 454.

Fuster, J.M., 2003. Cortex and Mind. Oxford University Press, New York.

Gentner, D., Goldin-Meadow, S. (Eds.), 2003. Language in Mind: Advances in the Study of Language and Thought. The MIT Press, Cambridge, MA.

Georgopoulos, C., 1991. Syntactic Variables: Resumptive Pronouns and A'-binding in Palauan. Kluwer, Dordrecht.

Goedemans, R., van der Hulst, H., 2008. Fixed stress locations. In: Haspelmath, M., Dryer, M.S., Gil, D., Comrie, B. (Eds.), The World Atlas of Language Structures Online. Max Planck Digital Library, Munich, Chapter 14.

Goldin-Meadow, S., 2003. The resilience of language: what gesture creation in deaf children can tell us about how all children learn language. In: Werker, J., Wellman, H. (Eds.), The Essays in Developmental Psychology Series. Psychology Press, New York.

Golumbia, D., 2010. Minimalism is functionalism. Language Sciences 32, 28–42.

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

Green, R.E., et al., 2010. A draft sequence of the Neandertal genome. Science 328 (5979), 710–722.

Greenhill, S.J., Currie, T.E., Gray, R.D., 2009. Does horizontal transmission invalidate cultural phylogenies? Proceedings of the Royal Society B 276, 2299–2306.

Hagoort, P., 2005. On Broca, brain and binding: a new framework. Trends in Cognitive Science 9, 416–423.

Haspelmath, M., 2006. Against markedness (and what to replace it with). Journal of Linguistics 42 (1), 25–70.

Haspelmath, M., 2007. Pre-established categories don't exist: consequences for language description and typology. Linguistic Typology 11 (1), 119–132.

Haspelmath, M., in press. Framework-free grammatical theory. In: Heine, B., Narrog, H. (Eds.), The Oxford Handbook of Grammatical Analysis. Oxford University Press, Oxford.

Haspelmath, M., Dryer, M.S., Gil, D., Comrie, B. (Eds.), 2005. The World Atlas of Language Structure. Oxford University Press (online: http://wals.info/ published by Max Planck Digital Library, Munich).

Hauser, M., Chomsky, N., Fitch, T.W., 2002. The faculty of language: what is it, who has it, and how did it evolve? Science 298, 1569–1579.

Hawkins, J.A., 2004. Efficiency and Complexity in Grammars. Oxford University Press, Oxford.

Hawks, J., Wang, E.T., Cochran, G.M., Harpending, H.C., Moyzis, R.K., 2007. Recent acceleration of human adaptive evolution. Proceedings of the National Academy of Sciences 104 (52), 20753–20758.

Henrich, J., Boyd, R., Richerson, P., 2008. Five misunderstandings about cultural evolution. Human Nature 19, 119–137.

Hockett, C.F., 1961. The problem of universals in language. In: Greenberg, J. (Ed.), Universals of Language. MIT Press, Cambridge, MA, pp. 1–29.

Huddleston, R., Pullum, G., 2002. The Cambridge Grammar of the English Language. Cambridge University Press, Cambridge.

Hudson, R., 2010. Reaction to 'The Myth of Language Universals and cognitive science'. Lingua 120, 2676–2679.

Hyman, L., submitted for publication. Does Gokana really have syllables? Or: what's so great about being universal. Phonology.

Jackendoff, R., 2002. Foundations of Language. Oxford University Press, Oxford.

Jackendoff, R., 2003. Toward better mutual understanding (response to peer commentaries). Behavioral and Brain Sciences 26, 695–702.

Jordan, F., Gray, R., Greenhill, S., Mace, R., 2009. Matrilocal residence is ancestral in Austronesian societies. Proceedings of the Royal Society of London Series B 276 (1664), 1957–1964., In: http://www.mpi.nl/publications/escidoc-102532/@@popup.

Karlsson, F., 2007. Constraints on multiple center-embedding of clauses. Journal of Linguistics 43, 365–392.

Kay, R.F., Cartmill, M., Balow, M., 1998. The Hypoglossal Canal and the origin of human vocal behavior. Proceedings of the National Academy of Sciences 95 (9), 5417–5419.

Keenan, E.L., 2009. Voice determines co-argument anaphora in W. Austronesian. In: Chung, S., Finer, D., Paul, I., Potsdam, E. (Eds.), Proceedings of the XVIth Annual Meetings of the Austronesian Formal Linguistics Association (AFLA XVI). pp. 77–91.

Keenan, E.L., Stabler, E.P., 2010. Language variation and linguistic invariants. Lingua 120, 2680–2685.

Kemmerer, D., Eggleston, A., 2010. Nouns and verbs in the brain: implications of linguistic typology for cognitive neuroscience. Lingua 120, 2686–2690.

Kenneally, C., 2010. Language lessons: you are what you speak. New Scientist 2762, 32–35.

Kirby, S., 2001. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. IEEE Transactions on Evolutionary Computations 5, 102–110.

Kirby, S., Hurford, J., 2002. The emergence of linguistic structure: an overview of the iterated learning model. In: Cangelosi, A., Parisi, D. (Eds.), Simulating the Evolution of Language. Springer-Verlag, London, pp. 121–148.

Kita, S., 2006. A grammar of space in Japanese. In: Levinson, S.C., Wilkins, D. (Eds.), Grammars of Space. Cambridge University Press, Cambridge, pp. 437–476.

Koschmann, T., 2010. On the universality of recursion. Lingua 120, 2691–2694.

Krause, J., et al., 2007. The derived FOXP2 variant of modern humans was shared with Neandertals. Current Biology 17, 1908–1912.

Labov, W., 1975. What is a Linguistic Fact? Peter de Ridder Press, Lisse.

Lee, Y., Lee, E., Gordon, P.C., Hendrick, R., 2010. Commentary on Evans and Levinson. The myth of language universals. Lingua 120, 2695–2698.

Leger, D.W., 2005. First documentation of combinatorial song syntax in a suboscine passerine species. The Condor 107 (4), 765–774.

Levelt, W.J.M., 2008 [1974]. Formal Grammars in Linguistics and Psycholinguistics. Benjamins, Amsterdam.

Levinson, S.C., 1981. Some pre-observations on the modelling of dialogue. Discourse Processes 4 (2), 93–116.

Levinson, S.C., 1983. Pragmatics. Cambridge Textbooks in Linguistics, Cambridge University Press, Cambridge.

Levinson, S.C., 2000. Presumptive Meanings: The Theory of Generalized Conversational Implicature. MIT Press, Cambridge, MA.

Levinson, S.C., 2003a. Space in Language and Cognition: Explorations in Cognitive Diversity. Cambridge University Press, Cambridge.

Levinson, S.C., 2003b. Language and mind: let's get the issues straight! In: Gentner, D., Goldin-Meadow, S. (Eds.), Language in Mind: Advances in the Study of Language and Cognition. MIT Press, Cambridge, MA, pp. 25–46.

Levinson, S.C., 2006. On the human 'interactional engine'. In: Enfield, N., Levinson, S.C. (Eds.), Roots of Human Sociality. Culture, Cognition and Human Interaction. Berg Publishers, Oxford, pp. 39–69.

Levinson, S.C., Meira, S., The Language & Cognition Group, 2003. 'Natural concepts' in the spatial typological domain – adpositional meanings in cross-linguistic perspective: an exercise in semantic typology. Language 79 (3), 485–516.

Lieberman, P., 2006. Toward an Evolutionary Biology of Language. Harvard University Press, Cambridge, MA.

Lieberman, P., McCarthy, R., 2008. Tracking the evolution of language and speech. Expedition 49 (2), 15–20.

Liszkowski, U., Schäfer, M., Carpenter, M., Tomasello, M., 2009. Prelinguistic infants, but not chimpanzees, communicate about absent entities. Psychological Science 20, 654–660.

Longobardi, G., Roberts, I., 2010. Universals, diversity and change in the science of language: reaction to 'The Myth of Language Universals and Cognitive Science'. Lingua 120, 2699–2703.

Mace, R., Holden, C.J., Shennan, S. (Eds.), 2005. The Evolution of Cultural Diversity. A Phylogenetic Approach. UCL Press, London, pp. 67–83.

MacLarnon, A.M., Hewitt, G.P., 1999. The evolution of human speech: the role of enhanced breathing control. American Journal of Physical Anthropology 109, 341–363.

Makino, S., Tsutsui, M., 1986. A Dictionary of Basic Japanese Grammar. The Japanese Times, Tokyo.

Malaia, E., Wilbur, R., 2010. Sign languages: contribution to neurolinguistics from cross-modal research. Lingua 120, 2704–2706.

Malchukov, A., Cysouw, M., Haspelmath, M. (Eds.), in preparation. Semantic maps: methods and applications. Special Issue of Linguistic Discovery.

Marcus, G., 2008. Kluge: The Haphazard Construction of the Human Mind. Houghton Mifflin Company, Boston.

Martinez, I., Arsuaga, J.L., Quam, R., Carretero, J.R., Gracia, A., Rodriguez, L., 2004. Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. Proceedings of the National Academy of Sciences 101 (27), 9976–9981.

Maslen, R., Theakston, A., Lieven, E., Tomasello, M., 2004. A dense corpus study of past tense and plural overgeneralizations in English. Journal of Speech, Language and Hearing Research 47, 1319–1333.

McConvell, P., 1985. The origin of subsections in Northern Australia. Oceania 56, 1–33.

McKay, G.R., 1978. Pronominal person and number categories in Rembarrnga and Djeebbana. Oceanic Linguistics 17 (1), 27–37.

McMahon, A., McMahon, R., 2006. Language Classification by Numbers. Oxford University Press, Oxford.

Meir, I., Sandler, W., Padden, C., Aronoff, M., 2010. Emerging sign languages. In: Marschark, M., Spencer, P. (Eds.), Oxford Handbook of Deaf Studies, Languages, and Education, vol. 2. Oxford University Press, Oxford.

Nettle, D., 2007. Language and genes: a new perspective on the origins of human cultural diversity. Proceedings of the National Academy of Sciences 104, 10755–10756.

Newman, A.J., Bavelier, D., Corina, D., Jezzard, P., Neville, H.J., 2002. A critical period for right hemisphere recruitment in American Sign Language processing. Nature Neuroscience 5, 76–80.

Newmeyer, F.J., 1983. Grammatical Theory: Its Limits and Its Possibilities. Chicago University Press, Chicago.

Newmeyer, F.J., 1986. Linguistic Theory in America. Academic Press, Cambridge, MA.

Newmeyer, F.J., 1998. Language Form and Language Function. MIT Press, Cambridge, MA.

Newmeyer, F.J., 1999. Adaptation, optimality, and functional explanation: two serious problems (commentary on Martin Haspelmath 'Optimality and Diachronic Adaptation'). Zeitschrift für Sprachwissenschaft 18, 235–237.

Newmeyer, F.J., 2005. Possible and Probable Languages: A Generative Perspective on Linguistic Typology. Oxford University Press, Oxford.

Nyst, V., 2007. A Descriptive Analysis of Adamorobe Sign Language. LOT Dissertation Series, Utrecht.

Odling-Smee, F.J., Laland, K.N., Feldman, M.W., 2003. Niche Construction: The Neglected Process in Evolution. Princeton University Press, Princeton, NJ.

O'Grady, W., 2010. Fundamental universals of language. Lingua 120, 2707–2712.

Pagel, M., 2009. Human language as a culturally transmitted replicator. Nature Reviews Genetics 10, 405–415.

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

Papassotiropoulos, A., et al., 2006. Common Kibra alleles are associated with human memory performance. Science 314 (5798), 475–478, doi:10.1126/science.1129837.

Petersson, K.M., Silva, C., Castro-Caldas, A., Ingvar, M., Reis, A., 2007. Literacy: a cultural influence on functional left–right differences in the inferior parietal cortex. European Journal of Neuroscience 26 (3), 791–799.

Plank, F., Filimonova, E., 2000. The universals archive: a brief introduction for prospective users. Sprachtypologie und Universalienforschung 53, 109–123.

Pullum, G., Scholz, B.C., 2007. Systematicity and natural language syntax. Croatian Journal of Philosophy 7 (21), 375–402.

Pyers, J., Senghas, A., 2009. Language promotes false-belief understanding: evidence from Nicaraguan Sign Language. Psychological Science 20, 805–812.

Reali, F., Christiansen, M.H., 2005. Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. Cognitive Science 29, 1007–1028.

Reesink, G., Singer, R., Dunn, M., 2009. Explaining the linguistic diversity of Sahul using population models. PLoS Biology 7 (11), e1000241, doi:10.1371/journal.pbio.1000241.

Reuland, E., Everaert, M., 2010. Reaction to: the myth of language universals and cognitive science' – Evans and Levinson's cabinet of curiosities: should we pay the fee? Lingua 120, 2713–2716.

Rice, K., 2006. Let the language tell its story? The role of linguistic theory in writing grammars. In: Ameka, F., Dench, A., Evans, E. (Eds.), Catching Language: The Standing Challenge of Grammar-writing. Mouton de Gruyter, Berlin, pp. 235–268.

Richerson, P.J., Boyd, R., 2004. Not by Genes Alone: How Culture Transformed Human Evolution. Chicago University Press, Chicago.

Rothstein, S., Treves, A., 2010. Computational constraints on compositional interpretation: refocusing the debate on language universals. Lingua 120, 2717–2722.

Safir, K., 2010. How not to find linguistic universals. Lingua 120, 2723–2726.

Sandler, W., 2009. Symbiotic symbolization by hand and mouth in sign language. Semiotica 174 (1/4), 241–275.

Sandler, W., 2010. The uniformity and diversity of language: evidence from sign language. Lingua 120, 2727–2732.

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 102 (7), 2661–2665.

Schegloff, E., 1988. Presequences and indirection: applying speech act theory to ordinary conversation. Journal of Pragmatics 12 (1988), 55–62.

Schick, B., de Villiers, P., de Villiers, J., Hoffmeister, R., 2007. Language and theory of mind: a study of deaf children. Child Development 78, 376–396.

Schütze, C.T., 1996. The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology. The University of Chicago, Chicago.

Seuren, P.A.M., 2004. Chomsky's Minimalism. Oxford University Press, New York.

Silverstein, M., 1981. The limits of awareness. Sociolinguistic Working Paper Nr. 84. Available online from ERIC.

Simos, P.G., Billingsley-Marshall, R.L., Sarkari, S., Pataraia, E., Papanicolaou, A.C., 2005. Brain mechanisms supporting distinct languages. Learning Disabilities Research and Practice 20 (1), 31–38.

Smith, N.V., Tsimpli, I.M., 1995. The Mind of a Savant: Language Learning and Modularity. Blackwell, London.

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 106 (26), 10587–10592.

Suzuki, R., Buck, J.R., Tyack, P.L., 2006. Information entropy of humpback whale songs. Journal of the Acoustical Society of America 119 (3), 1849–1866.

Tremblay, A., 2005. Theoretical and methodological perspectives on the use of grammaticality judgment tasks in linguistic theory. Second Language Studies 24, 129–167.

Vajda, 2008. A Siberian link with Na-Dene languages. In: Dene-Yeniseic Symposium, Alaska, Fairbanks.

Valaki, C.E., Maestu, F., Simos, P.G., et al., 2004. Cortical organization for receptive language functions in Chinese, English, and Spanish: a cross-linguistic MEG study. Neuropsychologia 42 (7), 967–979.

Vogt, P., Lieven, E., 2010. Verifying theories of language acquisition using computer models of language evolution. Adaptive Behavior 18 (1), 21–35.

Waterfall, H., Edelman, S., 2009. The neglected universals: learnability constraints and discourse cues. A commentary on Evans & Levinson, The myth of language universals: language diversity and its importance for cognitive science. Behavioral and Brain Sciences 32, 471–472.

Weinreich, U., 1954. Is a structural dialectology possible? Word 14, 388–400.

West-Eberhard, M., 2003. Developmental Plasticity and Evolution. Oxford University Press, New York.

Zuidema, W., de Boer, B., 2009. The evolution of combinatorial phonology. Journal of Phonetics 37 (2), 125–144.

Zuidema, W., Verhagen, A., 2010. What are the unique design features of language? Formal tools for comparative claims. Adaptive Behavior 18 (1), 48–65.