With diversity in mind

Abstract (183 words)

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

R1. Introduction

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

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

We have organized the response as follows.

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

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

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

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

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

Section R2. Incompatible evaluation metrics reflect different paradigms

It was never our intention to engage in mud-slinging with our generative colleagues, but as Tomasello predicted there was a certain inevitability that familiar sibling quarrels would be rerun. Most of the criticisms from the generative camp reflect deep differences between generative and typological/functionalist approaches in their overall assumptions about many issues. Where do we locate causal explanations? Where do we seek the general unifying laws behind surface diversity – in structure or in process? Do we use only discrete mathematical models (favouring regularised representations), or do we bring in continuous and stochastic models as well (favouring representations sticking closer to surface variety)? Should generalisations purporting to directly represent mental reality, or are they modelling the summed information of thousands of different coevolutionary products shaped by multiple selective factors? Should we adopt essentialising categorisations (as ‘formal universals’), or abandon this as misleading and adopt a strategy that measures surface diversity directly so as not to lose data that is useful for evaluating the fit of models?
Generative and typological/functionalist accounts will give different answers to each of these questions, and it is this difference in overall scientific paradigm that accounts for the seemingly irreconcilable conflict between generativist commentators like Freidin and Pesetsky, who see our proposals as so imprecise as to be unfalsifiable, and psychologists like Tomasello and Margoliash & Nusbaum, for whom it is the generative approach that has moved away from falsifiability.

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

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

We now take up these issues one at a time. A further criticism, that we may have overstated the range of diversity by ignoring the fact that languages all lie within a certain possibility space (Pinker & Jackendoff, Tallerman) will be dealt with separately in R3 below.
R.2.1 What kind of theory?

Smolensky & Dupoux and Freidin complain that we did not offer a fully articulated theory with precise predictions about sentential structure. But that was not what we set out to do. Our goal was to survey our current understanding of language variation, explain its import for the cognitive sciences, and outline a fertile area for future research. We sketched the kind of biological models into which this variation neatly seems to fit, and which invite future development in a number of directions. A lot of these materials and ideas have not been sufficiently taken into account, we felt, by researchers in the cognitive sciences. We were gently suggesting that the time has come for a paradigm change, and at the end of this response we will say a little more.

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

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

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

R.2.2. Cognition, behavior, and representation

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

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

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

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

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

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

This conflation of the metalanguage with the object of description is a peculiar trick
of the generative tradition. By going down this path, it has opened up a huge gap between
theory and the behavioral data that would verify it. The complex representational structures
look undermotivated, and covert processes proliferate where alternative models deftly avoid
them (see the discussion of Subjacency and covert movement in R6.8).
A biologist does not assume that a snail maintains an internalized representation of the mathematical equations which describe the helical growth of its shell. Even for the internal characterization of a mental faculty the strategy is odd: computer scientists interested in characterizing the properties of programming languages use a more general auxiliary language to describe them, as in Scott-Strachey denotational semantics. Once explanatory theories hook external factors (e.g. psycholinguistic or evolutionary factors) into the account, this conflation of cognition and metalanguage can be dropped.

Smolensky & Dupoux’s aphorism “Generative grammar merits a central place in cognitive science because its topic is cognition and its method is science”, then, will not find universal approval: other branches of linguistics are much more in tune with psychological reality as reflected in language acquisition, production and comprehension. Nor has generative grammar of the Chomskyan variety been particularly successful as an explicit theory of linguistic representation. Many other representational formats like HPSG, and LFG have had greater uptake in computational circles (see e.g. Butt et al 2006, Reuer 2004). LFG, for example, adopts a parallel constraint-based architecture that includes dependency as well as constituency relations. This allows for the direct representation of crucial types of variability discussed in §5 of our target article, while avoiding the need for movement rules or large numbers of empty nodes (see §R6.8 for further discussion of how this works for subjacency). These formats, which represent different outgrowths from the same generative roots, show that precise, testable, computationally tractable models of language can be developed that reflect cross-linguistic diversity much more directly in their architecture.

R2.3 Abstractness and universal generalizations

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

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

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

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

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

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

R2.5 Non-abstract representations preserve information
Rizzi suggests that the typologist’s strategy of using an ‘extremely impoverished, non-abstract descriptive apparatus’ that takes diversity at face value in representing phenomena will have less success than the generative program in establishing universal patterns. Yet as the burden of explanation for cross-linguistic patterning is moved out of the prewired mind and into the evolution of individual language systems under selection from the sorts of factors outlined above, the most appropriate mathematical models employ stochastical and continuous methods rather than the discrete methods that have characterised the generative tradition (Pierrehumbert et al. 2000). And once we employ these methods, there are positive benefits in ‘directly measuring the variation, instead of reducing it’ (Bickel 2009): any other strategy risks degrading the information on which the methods are based.

Take the question of how perceptual discriminability and articulatory ease interact in the domain of vowel dispersion over the formant space to favour the emergence of some vowel systems over others. The classic study by Liljencrants & Lindblom (1972) simulated the evolution of vowel systems over time under these twin selectional pressures and compared the results to the distribution of attested vowel inventories. The insights yielded by their model would not have been possible if the descriptions of vowel systems had been in terms of discrete binary features such as [front] and [round] rather than in terms of position in a continuous three-dimensional space based on formant frequencies.

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

R2.6 Neglect of ‘formal universals’

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

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

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

**R2.7 An underanalyzed Wunderkammer of variation**

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

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

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

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

**Pesetsky** points out that Tlingit may after all have an initial slot into which constituents can be systematically shifted (we would need to know actually what can go there, and if that is actually predicted by a constituency analysis).

But he is wrong in presenting Warlpiri as the “free word order language par excellence” – it is well known that Warlpiri places its auxiliary after the first constituent, and that when words are grouped together into a contiguous NP only the last word needs to carry case, instead of the usual patterning of inflecting every word. Neither of these properties,
however, are found in Jiwarli (Austin & Bresnan 1996), which is why we chose it as our example.

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

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

The seven general charges we have discussed capture, we think, most of the sources of disagreement. Freidin’s commentary in particular indicates the deep rift in contemporary linguistics between Chomskyans and the rest, which ultimately rests on different judgements about the interlocking of theory and evidence. This is regrettable, as generative grammar has served to open up the ‘deep unconscious’ of language as it were, showing how languages are organized with far greater complexity than had hitherto been imagined. While Chomskyans have presumed that these complexities must be innate, we have argued that there are two blind watch-makers: cultural evolution acting over deep time, and genetic infrastructure, which for the most part of course will not be specific to language.
Finally, let us point out that Chomsky’s own point of view makes it clear that the generative enterprise simply has a different target than the program we are trying to promote, namely (in our case) working out the implications of language diversity for theories of cognition and human evolution. The following recent quote makes this clear:

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

On this view, UG primarily constrains the “language of thought”, not the details of its external expression. The same conclusion was stoically reached by Newmeyer (2004, p.545):

 Typological generalizations are therefore phenomena whose explanation is not the task of grammatical theory. If such a conclusion is correct, then the explanatory domain of Universal Grammar is considerably smaller than has been assumed in much work in the Principles-and-Parameters approach and Chomsky (2007, p. 18) seems in part to concur (“Diversity of language provides an upper bound on what may be attributed to UG”).

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

Section R3. How much of the design space is populated?

Pinker and Jackendoff point out correctly that the possible design space for human
languages is much greater than the space actually explored by existing languages. Two basic
questions arise: (i) What exactly are the dimensions of the possible design space, of which
just one corner is actually occupied? (ii) What exactly does this sequestration in a small
corner imply?

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

We have to presume that most likely all the languages we have now are descended by
cultural evolution from a single ancestral tongue (it would take an event of total spoken
language loss to be otherwise - not impossible, but requiring a highly unlikely scenario, like
an isolated lineage descended from a deaf couple). Now consider the following surprising
fact. The structural properties of language change on a near-glacial time scale. In an ongoing
study using Bayesian phylogenetics, Dunn et al. (forthcoming) have found that taken
individually, a structural feature within a single language large family like Austronesian
changes on average just once per c. 50,000 years.5 What that implies of course is that all the
languages we now sample from are within structural spitting distance of the ancestral
It is quite surprising in this light that typologists have been able to catalogue so much linguistic variation. Once again, a coevolutionary perspective is an essential corrective to the enterprise.

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

Let’s turn now to the properties of the design space. Pinker & Jackendoff point out that we set aside a rich set of functional universals on the grounds that they are definitional of language (a move we borrowed directly from Greenberg). Of course it is not trivial that these seem shared by all human groups (although very little empirical work has actually been done to establish this - but see e.g. Stivers et al. 2009). We think that there is a clear biological infrastructure for language, which is distinct from the structural properties of language. This consists of two key elements: the vocal apparatus and the capacity for vocal learning on the one hand (both biological properties unique in our immediate biological family the Hominidae), and a rich set of pragmatic universals (communicative intention recognition among them) on the other. This is the platform from which languages arise by cultural evolution, and yes, it limits the design space, like our general cognitive and learning capacities (Christiansen & Chater). We emphasized that those interested in the evolution of the biological preconditions for language have been looking in the wrong place - instead of looking at the input-output system (as Philip Lieberman has been urging for years - see e.g. Lieberman 2006), or the pragmatics of communicative exchange, we’ve been focussed on the syntax and combinatorics, the least determined part of the system, as demonstrated by linguistic typology.
A functional perspective has been a long running undercurrent in typological and descriptive linguistics, as Croft and Goldberg remind us. Goldberg suggests that the design space is highly constrained by systems motivations, for example, pressures to keep forms distinct while staying within the normal sound patterns of a language. These pressures provide explanations for the internal coherence of language structure, a perspective that is indeed necessary to explain how languages are not for the most part a heap of flotsam and jetsam accumulated during cultural evolution, but rather beautifully machined systems, with innovations constantly being adjusted to their functions.

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

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

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

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

Finally, it is salutary to recollect that it is only relatively recently that we have come
to recognize sign languages as fully developed languages with equal expressive power to
spoken languages. These languages with their easy access to iconicity and analog spatial
coding break out of the design space restricted by the strictly linear coding of the vocal-
auditory channel - the typology of these languages is still in development, and there are
plenty of surprises yet to come (see Meir et al. in press; Zeshan 2006a.b).
Section R4. Language variation and the future directions of cognitive science

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

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

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

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

A fundamental shift is that modern computational methods have revolutionized the possibility of studying change in systems using more and more complex and realistic simulations. Within the study of evolution, computational cladistics exploits this to the full, using for example Bayesian inference to run millions of simulations and Monte Carlo Markov chains to search for the optimum model that predicts back the data with the greatest likelihood. We can make history today, as Henry Ford thought we should.
In the coda of the paper we sketched a set of future directions for the language sciences based on evolutionary ideas, and these new methods put those directions within our grasp right now. Take the idea stated in thesis (3), that recurrent clustering of solutions will occur in grammars of non-closely related languages – such a claim can be tested by simulations. Equally tractable is the idea that changes cascade (thesis (4)), so that a few crucial ones may land a language in a gully of future developments. Thesis (5) about coevolution between brain, vocal organs and language has already begun being intensively explored by simulation (Christiansen & Chater 2008; Chater, Reali & Christiansen 2009). Thesis (7) suggests that we should investigate how the full range of attested language systems could have arisen – pie in the sky without computational simulation, but now thinkable. For example, we could follow Bickerton (1981) and start with a simple creole-like language, described by a set of formal features or characters, and use the rates and parameters of character change derived from recent work on the Bayesian phylogenetics of language families to simulate cultural evolution over 100,000+ years. Do we derive patterns of diversity like we now see, or would we need to model special historical circumstances like massive hybridization?

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

**R4.2 Learning and development**

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

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

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

McMurray & Wasserman correctly point out that our position radically moves the goal posts for language learning, suggesting that not only are a slew of specialized learning strategies employed to learn a language (and they provide very useful bibliographic leads here), but which of these strategies is deployed may depend on the language being learnt. We don’t necessarily learn Yéli Dnye with its 90 phonemes, flexible phrase order and widespread verb suppletion using the same strategies we use for English: “humans .. assemble language with a variety of learning mechanisms and sources of information, this assembly being guided by the particularities of the language they are learning”. Instead of talking about the passive acquisition of language, we should be talking about the active construction of many different skills. This perspective buries the idea of a single language acquisition device (LAD).
Christiansen & Chater, as well as Catania, emphasize that learning in development is the crucial filter through which languages have to pass. Languages have to be good to think with (to modify an adage of Levi-Strauss), otherwise they won’t make it. Christiansen & Chater have described (both in their 2008 BBS paper and Chater et al. 2009) interesting modelling that shows that the learning filter must be largely language-independent, and thus that properties of learning are unlikely to have evolved specifically for language. This is a new kind of evidence against the position taken by Pinker & Jackendoff that language-specific learning principles must be involved in the acquisition of language.

Finally, we would like to draw attention to one other crucial aspect of development, namely the way in which the environment is known to modulate developmental timing in the underlying biology of organisms, so that phenotypic variation can be achieved from the same genotype (‘phenotypic plasticity’), and conversely, phenotypic identity can be obtained from variant genotypes (‘developmental buffering’). In the conclusion to our paper we drew attention to the extraordinary achievement that is culture – generating phenotypic difference where there is no genetic difference, and phenotypic identity where there is genetic difference. These issues have been much explored in the biological literature on epigenesis and development (see West-Eberhard 2003 for a fascinating overview).

R4.3 The comparative perspective across species

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

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

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

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

R5 Situating language and cognition in the biology of variation

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

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

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

This research is going to revolutionize what we know about the mind and brain and how it works. By putting variation central, as the fuel of evolution, it will recast the language sciences. Some aspects of the language sciences are pre-adapted to the sea-change – sociolinguistics, dialectology, historical linguistics and typology – provided they can take the new mathematical methods on board. But we can look forward to the new psycholinguistics, centrally concerned with variation in human performance in the language domain both within and across language groups, and the new neurocognition of language which will explore both the varying demands that different languages put on the neural circuitry, and the way in
which superficial phenotypic standardization is achieved by distinct underlying processing strategies in different individuals.

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

R6. Appendix - Disputed data and generalizations


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

R6.2 Positionals and ideophones.

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

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

**R6.3 Coast Salish noun vs. verb distinction.**

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

**R6.4 Jemez / Kiowa number.**

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

**R6.5 Arrernte syllable structure.**
Nevins (and, briefly, Berent) take issue with us citing Arrernte as an example of a language that defies the ‘Universal CV preference’ by taking VC as the underlying syllable type. To contextualise their riposte, it is worth quoting Hyman (2008:13): ‘In each of the above cases, there is no “knock-out argument”. Anyone determined to maintain [these] universals can continue to do so, the worst consequence being an indeterminate or more awkward analysis… Architectural universals have this property: it all depends on your model and on what complications you are willing to live with.’

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

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

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

R6.6 Finite State Grammars and Cottontop Tamarins

Pullum & Scholz pull us up for propagating a misinterpretation of the findings in Fitch & Hauser (2004), by stating that Cottontop Tamarins have a general ability to learn Finite State languages. We stand corrected, and urge the reader to heed Pullum & Scholz’s clarification that Fitch & Hauser’s findings are restricted to the much smaller subset known as SL (Strictly Local) languages.
The investigation of recursive and augmentative structures in animal cognition is a current minor industry in cognitive science. If this is meant to shed light on the human language capacity, it is arguably quite misguided. Indefinite recursion, or discrete infinity as Chomsky prefers, is not an actual property of human language – no human is capable of indefinite centre-embedding for example. Only in the light of a radical distinction between competence and performance does this minor industry make any sense at all, and that little sense is undermined by the impossibility of testing animals directly for indefinite recursion.

**R6.7 Cinque's generalisation about Greenberg's Universal #20**, specifying strict ordering in noun phrases where the noun comes last, is raised by Rizzi as an example of how implicational universals can be made to follow from parameterized rules. However, Dryer (2009), drawing on a larger cross-linguistic sample, shows that you get better fit with the data if Cinque’s formal categories (like Adjective) are replaced by semantic categories (like ‘modifier denoting size, color etc.’). Cinque's parameterization just gives a discrete and approximate characterisation of statistical trends reflecting the interaction of many functional selectors.

**R6.8 Subjacency and ‘invisible Wh-movement’**

A number of commentators (Smolensky, Freidin, Rizzi) appealed to the Chomskyan notion of ‘subjacency’ as a convincing example of a highly abstract principle or rule-constraint which is manifested directly in languages like English. The idea in a nutshell is that movement of constituents is constrained so that they may not cross more than one ‘bounding node’ in the syntactic tree (in English, bounding nodes are a NP, i.e. noun phrase, or a complementizer phrase headed by *that*). Hence you can say "What does John believe that Mary saw ___?" but not "*What does John believe the rumour that Mary saw ___?".
Now consider Rizzi’s point that many languages including Chinese do not move their Wh-words (so called *in situ Wh*) - they would stay in the corresponding slots indicated in the just provided sentences – but appear to exhibit semantic interpretations that might constitute a parallel phenomenon. The apparent lack of wh-movement in Chinese, which at first seems an embarassment to the theory, is claimed however to mask *covert movement* at an underlying level, close to semantic interpretation: consequently the range of construals of a Chinese Wh-question is argued to be limited by the very same abstract constraint postulated for languages with overt movement (see examples in Rizzi's comment). For generativists this may seem like a double scoop: not only is the constraint of an abstract enough kind that children would find it hard to learn in English, but it even holds in Chinese where it is, in effect, invisible, so could not possibly be learnt! Moreover, it is a completely arbitrary and unmotivated constraint, so there is no apparent way for the child to infer its existence. Therefore it must be part of UG, a quirk of our innate language organ.

But this in fact is not at all a convincing example to the other camp.

First, to make it work in languages with and without overt ‘movement’, it has to be so particularized (‘parameterized’) for each language so that, as we noted in the paper, the child might as well learn the whole thing (Newmeyer 2005).

Second, there are perfectly good alternative models that do not use movement: wh-words are simply generated in the right place from the start, using other methods than movement to get the correct logical interpretations: within LFG, a combination of the FOCUS discourse function and prosodic structure can get *in-situ* Wh interpretation with no covert movement required (Mycock 2008). Through methods like these, LFG, HPSG, and Role and Reference Grammar have all developed ways of modelling both the English syntactic constraints and the Chinese interpretation constraints without any covert operations or unlearnable constraints.

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

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

R6.9 C-command.

Rizzi claims that ‘no language allows coreference between a pronoun and a NP when the pronoun c-commands the NP’ (*He said that John was sick; *each other saw the men). We pointed out that in languages (like Jiwarli) which lack constituency as the main organizing principle of sentence structure, notions like c-command cannot be defined (c-command is defined in terms of a particular kind of position higher in a syntactic constituency tree). But let us interpret this relation loosely and charitably, in terms of some general notion of domination or control. Then the observation would have very wide validity, but it would still be only a strong tendency. Counterexamples include Abaza reciprocals (Hewitt 1979) where the verbal affix corresponding to ‘each other’ occupies the subject rather than the object slot, and Guugu Yimidhirr pronominalisation, where it is possible to have a pronoun in the higher clause coreferential with a full NP in the lower clause (Levinson 1987).
Once again, then, we are dealing with a widespread but not universal pattern. The typological/functional paradigm explains it as emerging from a more general tendency in discourse (not just syntax): reference to entities proceeds with increasing generality, which is why ‘She came in. Barbara sat down’ is not a way of expressing ‘Barbara came in. She sat down.’ (see Levinson 2000 for a detailed Gricean account). Many languages have grammaticalised the results of this more general tendency, producing grammatical rules which can then be described by c-command (if you want to use that formalism) but also by other formalisms. Seeking the most general explanation for cross-linguistic patterning here directs us to more general pragmatic principles (‘use the least informative form compatible with ensuring successful reference given the current state of common ground’) rather than in terms of a specific syntactic constraint which only applies in a subset (even if a majority) of the world’s languages. Many strong tendencies across languages appear to have a pragmatic or functional base, undermining a presumption of innate syntax.

**R6.10 The ‘Verb object constraint’**.

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

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

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

(iii) There are twice as many VSO languages as VOS languages, 14% vs. 7% respectively in a worldwide sample by Dryer (2008), but only VOS languages seem likely to facilitate a ‘combination’ of verb and object.

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

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


University Press.


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

Endnotes

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

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

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

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

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

6 Abui, on František Kratochvíl’s (2007) analysis, comes rather close.

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