The Logical Structure of Linguistic Theory

Stephen R. Anderson, Yale University

Abstract

The object of inquiry in Linguistics is the human ability to acquire and use a natural language, and the goal of linguistic theory is an explicit characterization of that ability. Looking at the communicative abilities of other species, it becomes clear that our linguistic ability is specific to our species, undoubtedly a product of our biology. But how do we go about determining the specifics of this Language faculty? There are two primary ways in which we infer the nature of Language from the properties of individual languages: arguments from the Poverty of the Stimulus, and the search for universals that characterize every natural language. Arguments of the first sort are not easy to construct (though not as difficult as sometimes suggested), and apply only to a tiny part of Language as a whole. Arguments from universals or typological generalizations are also quite problematic. In phonology, morphology, and syntax, factors of historical development, functional underpinnings, limitations of the learning situation, among others conspire to compromise the explanatory value of arguments from observed cross-linguistic regularities. Confounding the situation is the likelihood that properties found across languages as a consequence of such external forces have been incorporated into the Language faculty evolutionarily through the 'Baldwin Effect.' The conflict between the biologically based specificity of the human Language faculty and the difficulty of establishing most of its properties in a secure way cannot, however, be avoided by ignoring or denying the reality of either of its poles.*

*This paper represents my Presidential Address to the Linguistic Society of America at its Annual Meeting in Chicago on 5 January, 2007. I am grateful to the audience there, especially Jerry Sadock, and also to many students in various courses at Yale touching on the matters dealt with here for comments and suggestions. Comments on the present paper by Andrew Carstairs-McCarthy, Brian Joseph, Fritz Newmeyer, Josh Tenenbaum, and referees for *Language* have also been quite useful and I acknowledge them with gratitude. The present paper is intended to be programmatic, rather than definitive. I hope to elaborate on many of the issues that are merely alluded to here in a book currently in preparation. Any resemblance between this and other works with the same or a similar title is purely coincidental.

The goal of this paper is to discuss what seems to me to be a false dichotomy in much theorizing about Language: to wit, the notion that attributing some property to the human cognitive faculty of Language[1] on the one hand, and providing an account of it that is external to this aspect of the structure of the mind are mutually exclusive. On that view, something is *either* a consequence of the structure of 'Universal Grammar'[2] *or else* it has a basis in function, or processing, or the workings of historical change, or something else external to the system of language itself and any specialized cognitive capacity that underlies it.

To the contrary, I want to suggest that these accounts of aspects of Language are not at all incompatible, and that in fact, we would expect a kind of duplication of foundation for much that is important about Language and its structure. But if this is correct, it actually leaves us in a somewhat worse state than before, with more basic epistemological questions than we thought we had and less clarity about how to go about answering them. I have no simple resolution to offer for this dilemma, but I argue that the straightforward answer that lies in denying the significance of one horn or the other simply will not do.

To establish the basis for my argument, let us ask what it is that we as linguists take as the goal of our work. The uncontroversial part comes straight from the dictionary: we can, presumably, agree that (as the Wikipedia says), 'Linguistics is the scientific study of language.' But when we try to be more precise, the sense of what we are studying is less obvious, in some ways more contentious, and subject to substantial change over the last century and a half or so of history.

For present purposes, I assume that — following the 'Cognitive Revolution' — what we want to study is not sets of sounds, or words, or sentences, or texts for their own sakes, but rather the system of knowledge or the cognitive capacity that underlies our ability to produce and understand these things. The central object of inquiry in Linguistics, that is, is the nature and structure of the cognitive faculty that supports Language. This is by no means all that linguists do, and I do not mean to denigrate the study of ways Language is used, the role of Language in society, and other pursuits. I do want to claim, though, that the central task for a 'scientific study of language' is to arrive at an understanding of an aspect of human cognitive organization. It is this that, like it or not, makes Cognitive Scientists of us all.[3]

---

[1] I capitalize the word *Language* here and below when referring to this general capacity, and leave *language* in lower case when referring to specific languages like English or French.

[2] I will generally avoid the use of this term, which tends to set off rioting in some quarters.

[3] I have been amazed at the number of linguists for whom this statement is controversial. And what right have I to decide for them what it is that constitutes 'linguistics'? Well, I'm the President, and if I don't want to eat broccoli I don't have to… But more seriously, so long as the reader is willing to grant

However we eventually characterize this faculty, it does seem clear that it is essentially a property of us as human beings, members of the species *Homo sapiens*. Absent severe and obvious deprivation or pathology, all humans acquire and use language spontaneously and without the need for explicit instruction. Furthermore, no member of any other species appears to have the capacity to do the same, with or without instruction. The 'Language faculty,' then, is both universal and unique to us within the animal kingdom. Before addressing the consequences of this fact for our work as linguists, I briefly justify its assertion by reviewing what we know about the communicative abilities of other animals.

1 The Uniqueness of Human Language

When we look at the communicative behavior of other species in nature, we find — as we tell our students in Linguistics 1 — that it is quite different in character from human language. All of the points below are discussed in Anderson 2004a, with some indication of the massive literature justifying these characterizations.

Most importantly, the communication systems of all other known animals are based on limited, fixed sets of discrete messages, all of which are essentially limited to the here and now. The inventory of these messages in each species constitutes a fixed list, and one that cannot be expanded by combining elements to form new and different complex messages. The same signal in different contexts can of course convey different information, but this flexibility of interpretation is entirely grounded in the lability of the situation, and not at all in the structure of the communication system itself.

In nearly all cases, the communication systems of non-human animals emerge without need for relevant experience: that is, they are not 'learned.' In some instances there may be a limited amount of 'fine tuning' possible concerning the precise conditions of use of some signal in the system, but this does not entail the capacity to acquire the use of genuinely novel signals, or to change the basic set in fundamental ways. Even in those cases where the system is learned, of which birdsong is by far the most robust example[4], the actual system acquired does not go beyond the char-

that the cognitive faculty underlying Language is *an* object of inquiry in Linguistics, the points below will still make sense.

[4]Some or all of the species in three out of twenty seven orders of birds display learned song in the sense developed in the literature. Apart from oscine songbirds, parrots and hummingbirds, vocal learning is also attested to at least a limited extent in some species of bats, cetaceans, and probably

acter of an essentially fixed list. Birds may (depending on species) learn a single song or many different songs (all nearly always conforming to a species-specific template), but where more than one song is available to a given bird, the difference between one and another serves the ends of variety in performance, rather than corresponding to distinct messages.

The most important properties of human natural language are quite different. On the one hand, human language seems to be distinctive to our species, and rooted in our biology, just as other animals' communication systems are part of their specific makeup.

Human language is learned, in the sense that experience affects which possibility from within a limited space will be realized in a given child. As just noted, a few other species — mostly birds — develop their communicative behavior on the basis of observed models, but in most animals, including all of the other primates, communication is entirely innate, and develops in a fixed way that is independent of experience.

Crucially, human language provides an infinite range of discrete, different messages. Where other species have fixed, limited sets of messages they can convey, humans have an unbounded range of things that can be expressed in language. In addition, human language use is voluntary, controlled mainly by cortical centers, while other animals produce communicative signals under various sorts of non-voluntary control.

The unboundedness of human natural language is grounded in some quite distinctive structural properties, properties that are apparently unique to it among communication systems. It is what Pinker (1994) describes as a 'Discrete Combinatorial System,' in which new messages are formed as novel combinations of existing elements, rather than as modulations of intensity or some other continuous variable in another message. This system is based on recursive, hierarchical combination. Recursion refers to the fact that structural units can include other instances of the same structural type as components. As a result, there is no limit to the number of different structures that can be accommodated by a small, fixed number of structural regularities.

Recursion, as illustrated by sentences like (1), is not found in any other communication system, and Hauser, Chomsky & Fitch (2002) propose it as the one unique characteristic of human language.[5]

---

also elephants. This does not change the conclusion here significantly, since the learning involved in all of these species is either of minor variations on signals from a fixed inventory, or else some sort of imitation of environmental sounds with little or no real communicative function.

[5]Everett 2005 has argued that Pirahã does not exhibit syntactic recursion, and that this fact shows

(1) [$_{S}$ [$_{NP}$ The claim that [$_{S}$ [$_{NP}$ the story about [$_{NP}$ his mother] that [$_{S}$ [$_{NP}$ Fred] told [$_{N}$ his psychiatrist]]] persuaded [$_{NP}$ the draft board] that [$_{S}$ he had a problem with authority]]] turned out to be false.]

Even if, along with Cheney & Seyfarth (2007), we accept the suggestion that the mental lives of non-human animals are based on thought processes that display recursion, the communication systems available to them do not have this property.

Another essential property of human language is what Hockett (1960) christened 'Duality of Patterning.' In every language, individually meaningless sounds combine to make meaningful words (or parts of words, 'morphemes') according to a specific system, and then these meaningful elements are combined according to a completely different system to make phrases, clauses, sentences, and so on. The distinctness of phonology from syntax in this way is not simply an ornament: it is what makes large vocabularies possible, by reducing the memorization of large numbers of linguistic *signifiants* to the task of remembering combinations of a small number of distinct elements of expression. Understanding the emergence and structural properties of phonology is thus just as important a task in capturing the nature of Language as understanding the properties of grammar.

Overall, the properties of Language make it unique in the animal world. Furthermore, with all due respect to the late Alex the parrot (Pepperberg 2000), Washoe and other chimpanzees (Gardner & Gardner 1969, Fouts 1997), and even Kanzi the bonobo (Savage-Rumbaugh, Shanker & Taylor 1998), efforts to teach such a system to other animals have not succeeded. This assertion is controversial in some quarters, but it would take me much too far afield to defend it here, and I will just refer to the discussion in Anderson 2004a for support. The bottom line is that there is no evidence that any other animal is *capable* of acquiring and using a system with the core properties of human language: a discrete combinatorial system, based on recursive, hierarchically organized syntax and displaying (at least) two independent levels of systematic structure, one for the composition of meaningful units and one for their combination into full messages.

---

that recursion cannot be part of the human Language faculty. Suppose we put aside the huge controversy of interpretation engendered by Everett's Pirahã material and accept for the sake of argument his claim that the language indeed lacks recursion. This still seems to tell us no more about whether recursion is part of the Language faculty than a language like Hawaiian, lacking a contrast between *t* and *k,* tells us about whether the Language faculty provides for distinctions between coronals and velars. My opinion here is like that of Fitch, Hauser & Chomsky 2005, p. 242, who take the position that a lack of recursion 'surely does not affect the argument that recursion is part of the human language faculty: … our language faculty provides us with a toolkit for building languages, but not all languages use all the tools.' I imagine, though, that this position will no more convince Everett than his convinces me.

Although the unique availability of language in this sense to humans, like the theory of evolution, remains controversial in some quarters, it should hardly be seen as surprising. In fact, there is no more reason to expect that our means of communication should be accessible to animals with a different biology than there is to expect ourselves to be able to catch bugs by emitting short pulses of high-frequency sound and listening for the echo, like a bat, or to tell whether the mouse that just scurried under the sofa is or is not a relative of the one we saw last week by sniffing the traces of urine it left behind.[6]

1.1 Rico the Border Collie

The popular press loves to resist this conclusion, highlighting cases in which animals supposedly acquire human language. Leaving aside the florid obituary for Alex and Washoe (Seibert 2007) that appeared in the *New York Times* not long ago, consider the case of Rico, the border collie (Kaminski, Call & Fischer 2004). Rico is a smart and interesting dog. He reliably fetches a particular item out of an array of several possibilities in response to a verbal command from his owner. His vocabulary is currently around 200 words — not actually earth shaking, in comparison to the number of commands trained dogs may respond to, but still impressive. He is particularly remarkable in that he can associate a new word with a novel item on the basis of a single trial — something that has been observed in human children learning words, and that psychologists call 'fast mapping.' No animal had ever been shown to be able to do this before.

But to put Rico's abilities into perspective, I invite the reader to watch the short film of him that is available in the online materials accompanying Kaminski et al. 2004. In it, we see Rico presented with an array of small toys, and hear his owner in the background urging him to find first the 'Tyrex' (a little blue dinaosaur) and then the 'Weinachtsmann' (a red Santa Claus doll), items on which he has already been trained. Finally, he is asked to find the 'sirikid,' an unknown word, and after some hesitation he picks out a small pinkish bunny, the only toy in the array which is unfamiliar to him.

Now it is important to note that, as far as we know, this is what Rico does: he

[6] A substantial research literature establishes the fact that mice deploy a specialized perceptual system, the vomeronasal organ (active in many other mammals for similar purposes), to recover information from urine traces that allows them to identify (possibly unfamiliar) relatives: cf. Sherbourne, Thom, Paterson, Jury, Ollier, Stockley, Beynon & Hurst 2007. Recent work (Cheetham, Thom, Jury, Ollier, Beynon & Hurst 2007) shows that a mouse can use this information to tell whether two other mice are related to one another, and not simply to herself.

doesn't talk, he doesn't read Shakespeare… This is certainly entertaining, and it is also scientifically interesting, because the size of his vocabulary is roughly similar to that at which the language-trained apes seem to top out, which is not bad. And 'fast mapping' is certainly a novel discovery in a dog. But there seems very little basis for confusing these skills with control of the essential facets of a human language.

There are a great many ways in which Rico's skills are a long way from Language. For one thing, he only learns words that refer to small, fetchable objects. Most words in a natural language are verbs, like *run* or *like*, adjectives like *small* and *unlikely*, grammatical words like *on* or *is*, etc. For another, Rico apparently only responds to his owner, while humans can treat one another as equivalent speakers and hearers.

Another, subtler point is due to my colleague Paul Bloom (2004). He points out that when a child learns a word like *sock*, she can use it to talk about the sock over there, to ask 'Where's my sock?' or to assert that 'That's MY sock,' etc. A word can be used in many ways to refer to a variety of situations. When Rico learns a word, however, it always refers to a single, very specific situation type: fetching an object.

So in comparison with any human language user, including pretty small children whose vocabulary may not be as big as his, Rico's accomplishments lose much of their interest. There is no reason to believe he controls an unbounded range of distinct messages: rather, it seems that the scope of his learning ability is well within the range of the 'signing' behavior of the language-trained apes. Rico is, perhaps, better at learning new items than they are, but he is, after all, a dog, and dogs have evolved so as to be much more sensitive to human communicative intentions than other animals (Hare, Brown, Williamson & Tomasello 2002).

There is also no evidence for syntax of any sort in Rico's ability, let alone anything displaying hierarchical, recursive structure. In fact, there is no need to imagine that Rico controls a discrete combinatory system of any kind. And since there is no evidence for phonological organization, either, there is no question of 'duality of patterning.'

Rico is undoubtedly quite a clever dog. But he is still a dog, and as far as we know only animals with human biology have a chance of mastering the essential properties of a human language. Rico provides us with no reason to doubt this conclusion.

1.2 Putty-nosed Monkeys

When we look at the facts, we must reject the claim that Rico displays anything like syntax (or phonology, for that matter) in what he does. But there have been other recent claims that some animals do in fact have a communication system that has a syntactic side to it. One such case was picked up eagerly in the popular science

press: when the *National Geographic News* of 17 May, 2006 headlined a story 'Monkeys use 'Sentences,' Study Suggests,' it was referring to the alarm calling behavior of an African primate species, the putty-nosed monkey (*Cercopitehcus nictitans*), as described in Arnold & Zuberbühler 2006.[7]

These animals have two basic alarm calls, although the difference between them is not as easy to interpret as a difference between two 'words' as in the case of, e.g., vervet monkeys (cf. Cheney & Seyfarth 1990), because the putty-nosed monkeys do not connect calls with specific predators as consistently as the vervets do. In any case, one of the calls is referred to as a 'pyow,' and is often (but not always) given when a leopard is sighted. The other call, a 'hack,' is given often (but not always) in reaction to spotting an eagle. The interesting point is that sometimes a monkey will produce a sequence of one to three 'pyows' followed shortly by some 'hacks,' and this seems to be a distinct signal. When one monkey does this, the others consistently move away with some urgency, a response which is different from the way they react to 'pyows' or 'hacks' alone.

Since the 'pyow-hack' call seems to consist of the same vocalizations as a simple 'pyow' followed by a simple 'hack,' it looks like what is going on is the combination of existing elements to form a new message. But is there any basis for the claim that the result is a 'sentence' or that what is involved is meaningfully described as 'syntax'? What seems evident is that 'pyow-hack' serves as a third call in the animals' repertoire. But there is no reason to believe that this involves some sort of free combination, since no other combinations occur. Furthermore, the 'meaning' of 'pyow-hack' does not appear to be a systematic combination of the 'meaning' of 'pyow' and the 'meaning' of 'hack.' We might possibly see this as a first, tentative step toward phonology — making a new symbol out of existing vocalizations — but there is no obvious reason to call this 'syntax.'[8]

[7]For an excellent analysis of the stories that circulated in the press around these matters, see Mark Liberman's *Language Log* posting at `http://itre.cis.upenn.edu/~myl/languagelog/archives/003192.html` and others referenced there.

[8]Observations of Campbell monkeys (*Cercopithecus campbelli*) in the Tai Forest reported in Zuberbühler 2002 come marginally closer to syntactic combination. These animals have two basic alarm calls, one for eagles and one for leopards, and a third 'Boom' call which can accompany either. A sequence of 'Boom' plus another alarm call is interpreted as indicating less certainty about the danger involved. This is the extent of the combinability involved, and even if one does not interpret 'Boom - Leopard' 'maybe a leopard' and 'Boom-Eagle' 'maybe an eagle' as two additional, phonologically complex calls — on the analogy of the putty-nosed monkey's 'pyow-hack,' the amount of 'syntax' entailed is completely trivial.

1.3 Starlings

Granting that the combinatorics displayed in these monkey alarm calls are quite trivial and far removed from the syntactic structures exhibited in human languages, a more serious challenge seems to be presented by work on European starlings (*Sturnus vulgaris*).[9] Recently, Gentner, Fenn, Margoliash & Nusbaum (2006) have reported that they were able to teach these birds to recognize a difference between song patterns that seems to mimic a rather complex syntactic structural property otherwise known to occur only in human languages.

In nature, starlings have rather complex songs which can be regarded as composed out of a number of component elements. A given bird's song will consist of a sequence of these elements, in various combinations. We can note that this structure already shows that these birds have something like phonology (known somewhat misleadingly as 'phonological syntax' in the literature on birdsong), because the individual pieces do not convey a message in themselves but only as components of a fully formed song.

Gentner and his colleagues divided the different song motifs they found in a group of starling songs into 'rattles' and 'warbles' (though without saying anything about the basis for this difference, or providing evidence that it has any reality for the birds in nature). They then prepared a set of new songs made up of various combinations of these elements. Some of these consisted of a number of 'rattle' motifs followed by same number of 'warbles,' while in others, the number of 'rattles' was different from the number of 'warbles.' They then showed that after a huge amount of training (10,000 to 50,000 trials), most (9 out of 11) of their birds could learn to distinguish songs of the first type (with equal numbers of motifs of the two classes) from those of the second, even in novel songs they had not previously heard.

If we think of these songs in formal terms, what was shown was that starlings could learn to distinguish patterns of the type A^nB^n from $(AB)^n$ or from A^mB^n where $m \neq n \leq 4$.[10] Similar patterns appear in human language, including cases of 'recursive center embedding,' and Gentner *et al* interpreted their results as showing that the birds were capable of learning a system with this property, and thus of being able to control recursive hierarchically organized structures of the type characteristic

[9] As in the case of the putty-nosed monkeys, Mark Liberman's analysis of the Gentner *et al* work on starlings, its reception in the press, and its real import is outstanding, and I have made use of his discussion here. See `http://itre.cis.upenn.edu/~myl/languagelog/archives/003076.html` among other *Language Log* postings for details.

[10] In formal language theory, A^nB^n describes a sequence AA…BB… of some number of As followed by the same number of Bs; $(AB)^n$ a sequence ABAB… of some number of As each immediately followed by a B; and A^mB^n a sequence AA…BB… of some number of As followed by a sequence of some (possibly different) number of Bs.

of human language syntax.

The first thing to note about this work is that center embedding (as in *the cheese the rat the cat chased ate was moldy*) is a rather odd kind of structure in itself, and one that most speakers have trouble interpreting. It is certainly not equivalent to recursive structure in general, and neither is it obvious that the birds must have learned a genuinely recursive rule in order to perform as they did.

There are in fact a variety of strategies the birds might be employing. For instance, if they could simply count the number of 'rattles' and then compare that with the number of 'warbles,' that would suffice. We know, in fact, that birds are capable of recognizing the cardinality of small sets — not by counting, in the specific way we do, but rather by 'subitizing' or recognizing quantities in a holistic way, so this strategy should be well within their capabilities without invoking the following of recursive structural rules. For further discussion of alternative accounts of the results of Gentner *et al.* see Corballis 2007.

There is no evidence that the birds learn to *produce* songs with the relevant structural property, as opposed to categorizing what they hear in order to obtain a food reward. There is also no sense in which the complex song (whatever its structure) has a 'meaning' that is a function of the 'meanings' of its parts. Overall, the most that could be said is that starlings are capable of learning to recognize an unusually complex 'phonological' pattern, a point that does not bear interestingly on discussions of possible syntax in non-human communication systems.

1.4 Language Abilities in Humans and Non-humans

What conclusions should we draw from this survey of cognitive and communicative capacities of humans as opposed to non-humans? The most basic result is that the communication systems of other species do not display the structural properties of human natural language. Contrary to the hopes and belief of Doctor Dolittle,[11] the ways other animals communicate in nature simply do not involve the most important structural attributes of human language, and in particular they do not involve anything like syntax, morphology or phonology in the relevant sense. More importantly, perhaps, it seems that members of other species are intrinsically lacking in the cognitive capacities underlying the learning and use of a system with those particular structural characteristics (even when offered M&Ms for successful performance). Human language is, apparently, uniquely accessible to members of the species *Homo sapiens*.

[11] Anyone who has never met the good doctor has many treats in store in consulting Lofting 1920 and its several sequels.

From a broader biological perspective, there is nothing remotely surprising about this result. The communication system of every species that has been seriously studied is deeply and essentially grounded in the biology of that species. Human language has its grounding in human biology, just as the communicative capacities of electric fish are grounded in their biology (Hughes 1999). This is not at all to denigrate the cognitive abilities of other species, or to imply that animal communication is uninteresting or inferior for the purposes to which its users put it. The point is simply that there are essential differences in this domain that are part and parcel of the biological differences among species.

2 Studying the Human Language Faculty

The previous section has argued that human language is a distinctive capacity of our species. It appears that a Language faculty that supports the rapid acquisition and use of natural language is a consequence of our biological nature. The facts seem consistent with the claim that such a faculty is determined by our biology, a faculty that also supports the development of competence in the language of the surrounding community in a largely effortless way during the first years of life by all normal humans who are not handicapped in their access to that language.

This is quite independent of whether the cognitive capacity in question is based wholly, partially, or not at all on capacities that are domain-specific to language. We can differ in our views of whether the bases of the Language faculty are limited in their applicability to language alone, or whether they are essentially connected in part or even entirely with more general abilities. The bottom line is that there is a package of capacities that underlie language learning and use, and this package as a whole is unique to our species.

As linguists, we wish to explicate precisely the nature and structure of that cognitive capacity, as argued above. But how are we to do that? In particular, how are we to identify the properties that we should attribute to the human Language faculty, and distinguish them from aspects of language structure and use that might have their origins elsewhere? That is the problem to which the remainder of this paper is devoted.

2.1 Sources of the Properties of Grammars

Our evidence for the Language faculty comes from the properties of the knowledge-structures (or 'grammars,' in one particular sense) that we find instantiated in individ-

ual speaker-hearers of natural languages. Accordingly, we can break down the problem of studying that faculty by asking what factors can contribute to the grammars we find, in the hope of identifying properties that bear its hallmark.

Grammars arise, of course, on the basis of the learner's experience with utterances in the surrounding community, the 'Primary Linguistic Data,' and the grammar attained will naturally reflect some properties of those data. Since the grammar that is acquired is not simply a registration or list of the utterances heard,[12] there must be some principles of inference that lead from the data to the grammar — a Learning Algorithm.

Another factor is the space of grammars that are cognitively possible, a function of the organization of the brain. In discussions of these matters, this is often conflated with the properties of the Learning Algorithm, but the two are logically quite distinct. It might well be that there are cognitively possible grammars which could never result from the available procedures of inference, and it is also possible that some possible outputs of those procedures would lie outside the space of possible grammars, and thus would need to be mapped in an imperfect way onto grammars that are possible. I generally disregard these possibilities here, though, and simply regard the combination of cognitive constraints on possible grammars and the principles by which grammars are related to data as collectively constituting the Language faculty we seek to understand.

In these terms, we can view the Learning Algorithm as a system of inference that maps a particular collection $\mathcal{D}$ of Primary Linguistic Data onto some specific (cognitively possible) grammar $\mathcal{G}$. If we now ask where the properties of $\mathcal{G}$ come from, we can identify at least three sources: (a) they might reflect regularities in the input data $\mathcal{D}$; or (b) they might be introduced as consequences of the way the Learning Algorithm manipulates the data; or (c) they might be things that are cognitively necessary, in the sense of being constitutive of Language in general, and thus definitional for the set of possible grammars. Obviously, if we want to understand the nature of the Language faculty, we need to separate properties of the first type from the other two. Regularities do not necessarily bear a clear indication of their origin, however, and the question thus arises of how we might distinguish the various possibilities. Some criteria for doing this have of course been proposed in the literature, and it is to these that we now turn.

[12] Or seen, in the case of a signed language, though I neglect such qualifications for the most part in what follows.

2.2 The 'Poverty of the Stimulus'

In identifying properties that are constitutive of the Language faculty, one form of argument appears to identify at least a subset of these with some precision. That is the argument from what is called the 'poverty of the stimulus.' In brief, the logic of this is that where speakers can be shown to have come to know something about the structure of their language, but where the data available to them as learners provided insufficient evidence for the property in question, it must be the case that the Language faculty itself is responsible.

There are actually a number of instances of this form of argument in the grammatical literature, in support of claims that one or another property of natural language must be derived from the structure of the Language faculty. The example that is most discussed, however, concerns the point that from very early on, children demonstrate an understanding of the fact that grammatical principles treat sentences in terms of their hierarchically organized constituents, and not just as strings of words. And the poster child here is the regularity involved in fronting Auxiliaries in English in the formation of *yes-no* questions.

If children indeed interpret sentences from a very early point in the learning process as if they are made up of parallel structural constituents (VPs, NPs, PPs, etc.) organized in a hierarchical fashion, and not simply strings of words, we need to provide an account of the basis for such an analysis. For a concrete example, consider the paradigm of sentences in

(2) a. All of the kids who <u>have</u> caught the mumps <u>must</u> stay home.
b. Must all of the kids who have caught the mumps [*e*] stay home?
c. *Have all of the kids who [*e*] caught the mumps must stay home?

Corresponding to a declarative sentence like (2a), English forms a 'yes-no' interrogative by inverting the subject with the auxiliary verb.[13] This requires us to identify the appropriate auxiliary verb to prepose to the beginning of the sentence in such a case, because (2a) contains more than one auxiliary, both underlined here.

The vast majority of sentences available to the learner on which to base such a generalization about the relation between declaratives and *yes-no* interrogatives will contain only a single auxiliary, and thus will be consistent with any one of a number of accounts. Among these are the following:

[13] I abstract away here from examples in which a 'dummy' verb *do* inverts with the subject in the absence of an overt auxiliary.

String-based: To form an interrogative, locate the leftmost auxiliary verb in the corresponding declarative and prepose it to the front of the sentence.

Structure-based: To form an interrogative, locate the nominal phrase that constitutes the subject of the corresponding declarative and the highest auxiliary verb within the predicate of that sentence, and invert them.

Of these, the String-based account would appear *a priori* to be the simpler one in formal terms. Nevertheless, as the difference in grammaticality between (2b) and (2c) demonstrates, it is the more elaborate Structure-based account which is correct. Now in fact, the evidence from child language suggests that children always produce correct questions of the form (2b), and do not make errors of the form (2c), which would be predicted if they entertained the String-based account (Crain & Nakayama 1987). Since this choice of the correct analysis seems to proceed despite the fact that clear evidence (sentences like (2b), for example) for the structure-based account are rare or non-existent in the actual Primary Language Data available to the early learner, we must account for this fact. The explanation that is standardly proposed is that the Language faculty itself specifies the Structure (rather than String) based nature of grammatical regularities, and thus this property does not have to learned anew by each child.

This argument has been the subject of considerable discussion in the recent literature. In particular, Pullum & Scholz (2002) argue that the claim that the structure-based nature of the regularity is under-determined by the available data is itself overstated. They argue that in a large corpus of English sentences (in particular, the *Wall Street Journal* corpus), there are in fact examples of the sort that would lead the child to posit a regularity based on structure, rather than on mere string sequence. While one might well question the extent to which *The Wall Street Journal* is representative of the input to the child, they show that some such examples also appear in the much more plausible CHILDES corpus. They suggest, then, that structure sensitivity as a property of grammatical principles does not need to be pre-specified by the Language faculty: it could be learned from the available data.

In a reply to Pullum & Scholz 2002, Legate & Yang (2002) argue that although there are some sentences of the relevant sort in the CHILDES corpus, their frequency is really much lower than what would appear to be required to drive learning, and so the conclusion that structure sensitivity need not be part of the Language faculty does not follow. They develop a precise account of the statistical prominence in the input data that seems to be necessary, and show that the level attained by sentences of the sort criterial for learning the structure sensitivity of the rule forming English *yes-no* questions is far below this threshold.

Data of the sort presumed by Legate and Yang may not have the importance they assume, however, if the conclusions of Perfors, Tenenbaum & Regier (2006) are correct. They show (see also Perfors, Tenenbaum & Regier 2008) that even in the absence of the putatively crucial sentence types, a sophisticated statistical learner can be constructed which will, after exposure to a corpus of suitably massaged data, emerge with a preference for structural accounts of grammatical structure over string based accounts. This might be taken to show that structure based grammar need not be built into the Language faculty, but could be deduced from the data.

But while this work does demonstrate that the attribution of structure sensitivity to the Language faculty is not logically necessary, we must still ask whether the conclusion that this property could be learned by a system like that developed by Perfors et al. is actually plausible as an account of human language learning. Does it really make sense, that is, to assume that children start out like this statistical learner, neutral with respect to the basis of grammatical regularities, and only come to the conclusion that these are structurally based after extensive exposure and contingent inference?

If this regularity is in fact learned, it is quite remarkable that children simply do not make errors of the sort we would expect if they were entertaining the string based hypothesis at some point, only to reject it later. In fact, we never find a language in which some generation of learners has, as it were, gone astray and made the wrong inference in this domain, such that their grammars were based on regularities of string sequence rather than structure. As far as we know, all grammatical rules in all languages, to the extent it matters, are based on the analysis of sentences in terms of their constituent structure and not just as strings of words.[14] And this in turn suggests that structure sensitivity is a property of Language, not just of particular rules in particular languages, rules and languages that could have been otherwise if learners had not made particular contingent statistical inferences.

It seems much more sensible to attribute structure sensitivity to the organization of the Language faculty than to statistical tendencies in the Primary Linguistic Data.

[14]Jerrold Sadock points out that the phenomenon in English known as 'Proximity Concord' (see Quirk, Greenbaum, Leech & Svartvik 1985), by which a sentence such as *No one except his own supporters agree(s) with him* or *A good knowledge of English, Russian and French are/is required for this position* may sometimes be produced with plural agreement where singular agreement would seem to be regular might constitute an exception to this generalization. There is an extensive literature on this phenomenon, both descriptive and experimental(see Bock, Cutler, Eberhard, Butterfield, Cutting & Humphreys 2006 for discussion and references in the context of a larger study of agreement phenomena in English), and it is reasonably clear that many factors are involved in such productions beyond simple string adjacency. A recent proposal of Marcus & Wagers (2007) for the representation of constituent structure in terms of linked partial trees (or 'treelets') explicitly addresses the kind of deviation from pure structure sensitivity represented by 'Proximity Concord' examples.

On this basis, learners never go astray for the simple reason that they never entertain the alternative hypothesis of string sensitivity.

That indeed seems like the most plausible conclusion, but it is a lot harder to go further and claim that it is a **necessary** one, because we cannot completely exclude the logical possibility that structure sensitivity is in fact learned.

As mentioned above, the literature contains other instances of arguments from the Poverty of the Stimulus,[15] although in general these have not been analyzed in nearly as much detail as the one considered above with respect to whether the stimulus really is impoverished enough to support the proposed conclusion, at least at the limits of logical possibility. This example suffices to illustrate my point, however. On the one hand, there are aspects of the structure of natural language whose emergence seems to suggest that they are substantively underdetermined by the Primary Linguistic Data available to the learner. On the other hand, it is extremely difficult to show, for a particular property of grammars, that it absolutely **must** be a consequence of the nature and structure of the Language faculty.

2.3 The Evidential Value of Universals

While some claims about the structure and content of the human Language faculty have been supported by 'Poverty of the Stimulus' arguments, most proposals about the properties that should be attributed to this cognitive capacity are grounded in discussions of cross-linguistic generalizations. When we find some property recurring in language after language, it is very tempting to say that this must be because it is determined by the Language faculty, either as a consequence of the way the learning algorithm works, or as a property of the class of grammars that are cognitively accessible.

Such a conclusion certainly does not follow as a matter of necessity, however, no matter how pervasive a phenomenon may be. It might always be the case that languages are all the same in this regard because the data available to language learners happen always to instantiate the regularity in question, for some external reason.

At one extreme, we might be dealing with an adventitious fact about the original language from which all modern languages are descended — 'proto-World,' assuming language was invented just once in one place — some fact which language change has never happened to alter. If that were true, there would not be any necessity to elevate it to the status of a cognitive necessity. Of course, for universals which we find

[15] For further discussion, see Anderson & Lightfoot 2002. Snyder 2007, pp. 174f provides another recent example of this argument type.

instantiated in signed languages (Sandler & Lillo-Martin 2006), this is not even a logical possibility, since as far as we know all existing signed languages are of comparatively recent origin, and cannot have inherited anything from 'proto-World.' None of this speculation is terribly serious, however, and I know of no actual proposals of this nature.

More seriously, something which we find to be true of all natural languages might well derive as a logical consequence of the structure of the Language faculty, and this is a common assumption. Alternatively, however, there might be some external force shaping the data in question in a way that imposes the observed regularity. To the extent that is true, we could evade the conclusion that the Language faculty is responsible by arguing that the regularity is simply learned from the data in every case. I elaborate on this apparent dichotomy below.

These two accounts of linguistic universals — as deriving either from the structure of the Language faculty or from external forces shaping the Primary Linguistic Data – are commonly seen as mutually exclusive. Kiparsky (2008), for example, attempts to lay out a set of criteria that distinguish categorically between 'true universals' (regularities due to the nature of language, the Language faculty) and 'typological generalizations, which are simply the results of typical paths of change.'

In fact, though, it is at least logically possible that **both** characterizations might be true for any particular observed regularity. That is, there might be an external force (e.g., change) which influences the data to display a particular pattern, while in addition the Language faculty is so constituted that learning and/or the space of possible languages is limited to, or at least strongly favors, languages that conform to that pattern. To the extent the two effects coincide, it will be difficult to disentangle the two modes of explanation. In the remainder of this section, I survey briefly a range of potential instances of this state of affairs.

2.3.1 Phonological Universals

Much discussion of linguistic universals from the 1960s through the 1990s assumed the validity of inferences from pervasive regularity across languages to the structure of the cognitive capacity for Language. Around the turn of the millennium, however, the balance began to shift somewhat. In phonology, Juliette Blevins (2004) developed the argument, based on work by a variety of scholars on the factors contributing to linguistic change, that virtually all of the standardly cited regularities of phonological structure are actually products of the mechanisms of change.

For example, we find in many languages that obstruent consonants are neutralized as voiceless and unaspirated in word-final position. This might be attributed to

a principle of grammar to the effect that these are precisely the least 'Marked' among the obstruents; final position is where various oppositions are commonly neutralized, and neutralization is always to the least 'Marked' value for the feature in question. Blevins, however, identifies a number of factors in articulation which cumulatively have the effect that voicing is particularly difficult to maintain in this position, and observes that even when an obstruent is actually pronounced with some voicing, the perceptual cues necessary for the listener to recover that fact may be missing or obscured. As a result of these effects, we predict that subsequent generations of learners will hear many final obstruents as voiceless even when speakers intended them to be voiced, and thus that a rule of final devoicing is a natural outcome of linguistic change. Where subsequent stages of the language have such a rule, then, its presence is due to phonetic effects and their interpretation in the process of transmission of language across generations (a major category of linguistic change[16]), and not to the structure of the cognitive faculty for Language.

Explanations of this sort do not depend on properties of the Language faculty in any essential way, and to the extent they can be generalized, deprive us of a basis for inferring properties of that faculty from phonological universals. On this view, the locus of explanation in phonology shifts from synchronic structure to diachrony, more or less as our Neogrammarian ancestors told us. The regularities we find are regularities of the input data, as shaped by factors of phonetic production and perception in the operation of linguistic change, and tell us nothing about cognitive organization. On Blevins' account, to the extent we can ground phonological regularities in the properties of change they need not be informative about the structure of the Language faculty.

2.3.2 Syntactic Universals

A very similar line is pursued by Newmeyer (2006) with respect to cross-linguistic regularities of syntactic structure, based on proposals in work by Hawkins (1994, 2004). In this account, allowable variation in syntactic form that is perfectly within the bounds of cognitive possibility — and thus, within the limits of the Language faculty — tends to be reduced to favor structures with specific advantages in functional or processing terms. The functional considerations lead speakers to choose the struc-

[16] But see Labov 2007, p. 346, fn.4, who writes: 'Halle (1962) argued that linguistic change is the result of children's imperfect learning in another sense: that late additions to adults' grammar are reorganized by children as a simpler model, which does not exactly match the parents' original grammar. Although Lightfoot (1997, 1999) argues for this model as a means of explaining completed changes, such a process has not yet been directly observed in the study of changes in progress.'

tures in question whenever possible. Simplifying somewhat, later learners assume that the only structures they hear are the only ones that are possible in the language, and so deduce a corresponding rule. Such a rule enforces, within the grammar of subsequent generations, the functional or processing considerations that drove the original biases in the input data.

For instance, much of the literature assumes that the Language faculty provides a 'Head Parameter' (originating in Chomsky 1981, with many subsequent reformulations), according to which constituents within a given language either have heads consistently preceding their complements, or consistently following them. Hawkins suggests that in languages where verbs precede their complements, there is a general preference for lighter constituents to precede heavier ones, a processing effect. But in fact, across categories heads tend to be lighter than their complements, so to the extent this preference affects sentence productions, heads of all types will tend to precede their complements. If this becomes enough of a regularity to become a grammatical rule, it will give the appearance of a 'heads first' setting for the 'Head Parameter,' but the basis of this generalization lies in the consequences of a processing preference, not in a restriction imposed by the Language faculty.

When regularities of this sort become entrenched and recur across languages, they may be apprehended as linguistic universals. The important point, though, is that they owe their appearance to external functional or processing effects which have conspired to make them properties of the data available to the learner. These universals do not derive from the structure of the Language faculty, however, but rather from the other factors that have shaped the Primary Linguistic Data in particular ways.

2.3.3 Morphological Universals

Results parallel to those proposed for phonology and syntax can also be cited for morphology. Some have argued, for example, that some universal principle must prevent the appearance in languages of rules of metathesis that mark morphological categories,[17] as opposed perhaps to rules of vowel or consonant change (called, in that case, Ablaut, mutation, metaphony, etc.). It is true that morphological metathesis rules are vanishingly rare — though not, in fact, totally non-existent. But as detailed elsewhere (Anderson 2004b and literature cited there going back to Janda 1984), the explanation for this fact is to be sought not in a restriction imposed by the Language faculty, but rather in the specific diachronic origins of non-concatenative morphology, and the nature of phonological rules of metathesis. To summarize, morphological

[17] McCarthy 1981 for example.

metathesis rules are rare because there is virtually no sequence of diachronic changes that could give rise to them.

If all of these lines of argument are correct, we might presume that the search for linguistic universals is hopelessly compromised as a source of evidence for the structure of the Language faculty. But sometimes that search does lead to useful results, even if not the ones we had in mind.

For instance, in a recent paper, Andrew Spencer (2006) has suggested that there really are not any universals in morphology. As a basis for that, he shows a number of ways in which words turn out to have an organization that does not correspond to a structured string of classical morphemes: Ablaut, multiple and zero exponence, etc. These are all rather familiar — see Anderson 1992 for a survey, among much other literature. But Spencer's conclusion that this leaves us without evidence for the substantive content of the Language faculty is at least premature. In particular, we do learn the valuable lesson that our cognitive capacity for language is such that we can assign significance not only to discrete signs, but to other aspects of word form as well.

Similarly, Andrew Garrett (2008) argues convincingly that one proposed universal of morphological structure — a drive toward paradigm regularity — is illusory. Paradigm regularity effects are instances of the shift of words from one paradigm to another: assuming the regular paradigm exists for new words to shift into, the shift from other paradigms to this one is based on its status in the language, where it may — but need not — constitute the default.

In fact, when we look at more elaborate paradigm shifts, such as those studied by Martin Maiden (1992, 2005) in Romance, following proposals of Aronoff 1994, we see good evidence that these paradigms really do have some systematic status in the cognitive organization of language, something that might not be evident if we paid attention only to the synchronic facts of specific languages.

Even investigations that do not turn up universals of the standard sort may thus still be quite informative about the nature of the Language faculty. What we have to avoid is the notion that all properties of Language in general will have the character of statements to the effect that 'all languages have property $\mathcal{P}$' or 'if a language has property $\mathcal{P}$, it also has property $\mathcal{P}'$'. Some interesting and important properties of Language concern the form of linguistic knowledge, not just its substantive content.

3 The Evolutionary Origins of the Language Faculty

In the end, we are left with a puzzle. On the one hand, we know from comparative considerations that human language must be a richly structured capacity, one that is unique to our species. But on the other hand, when we try to determine the properties of that capacity, we are faced with a general absence of necessity arguments: what we find could, logically, just be a result of external influences shaping the data, with learnable consequences that are of little value for determining the substantive content of the hypothesized Language faculty.

Discussions of these matters generally assume that the two modes of analysis are incompatible: to the extent one can show that factors such as articulation and perception, processing and historical change have (or at least could have) shaped the data in the way that we find them, this is taken to demonstrate that the patterns found should not be attributed to the Language faculty. For example, Blevins (2006a, 2006b) and Kiparsky (2006) disagree on the correctness of Blevins' account of Final Devoicing, but each assumes that what is at stake is an essentially binary choice between an explanation based in the mechanisms of diachronic change and one based in the content of the Language faculty.

I suggest, though, that it is not necessary to see these lines of argument as mutually exclusive. It might be, that is, that some — perhaps many — properties with an external basis (in recurrent patterns of diachronic change, for example) are also characteristic of the Language faculty.[18]

To see how that could be the case, let us ask how this Language faculty, construed as part of human biology, arose. It seems quite implausible to suggest that the basic Darwinian mechanisms of random genetic variation and natural selection could have resulted by themselves in the very specific properties that seem to be candidates for part of the Language faculty. Try to imagine, for instance, how adherence to the principle of Subjacency[19] might provide a competitive advantage that would increase a speaker's likely reproductive success. In fact, though, there is an evolutionary mechanism that is widely assumed to have been capable of shaping the Language faculty to the form in which we find it: so-called 'Baldwin Effects.'

[18] The position I maintain here is not in fact an original one: it was formulated quite explicitly (though without reference to the 'Baldwin Effect' which I invoke below) in Newmeyer 1990, and Fritz Newmeyer has pointed out to me that there are various other antecedents. Nonetheless, most discourse on matters of explanation in Linguistics continues to presume that external accounts and ones based on the human Language faculty are mutually exclusive, so it seems worth pursuing the matter further.

[19] Chomsky 1973 and much subsequent literature. Lightfoot 1991 offers the same example, as I learned after writing this.

3.1 Baldwin Effects

This is a notion that goes back to several independent but simultaneous proposals by evolutionary theorists at the end of the 19th century. It remains somewhat controversial, in part because at least some discussions have treated the 'Baldwin Effect' as a special case of genetic assimilation, although there are various reasons to think that that is not at all the right way to view it. A number of other formulations exist that have the right consequences for our purposes, and the notion is accepted by many scholars. A recent collection of papers edited by Weber & Depew (2003) surveys the history of the Baldwinian idea and the issues associated with it, and reaches generally optimistic conclusions.

With specific reference to Language, Newmeyer (2006, p. 213) rejects a Baldwinian account of the origin of constraints in syntax. His basis for this, however, is a single paper (Yamauchi 2001) reporting an unsuccessful computer simulation. Interestingly, the same author's subsequent dissertation (Yamauchi 2004) reports considerably more work on the topic, and concludes that when combined with ideas from the literature on 'niche construction' in evolution, Baldwinian ideas show great promise in accounting for the origin of properties of human language.

The core notion appealed to in a Baldwinian account of the origin of some behavior or cognitive ability of a species is the claim that when a behavior provides an advantage within the specific context of the organism, it is advantageous to be able to acquire that behavior quickly and efficiently.

The use of language is surely a candidate for such an analysis: once language emerged in a human population, it quickly became essential to social organization and other aspects of life that provided humans with some competitive advantages. That had the consequence that in order to be a functioning member of a human society where language was in use (and to have any chance of reproducing), an individual had to be able to acquire the language of that society. Where any specific property of that language might have come from is essentially irrelevant: community members have to learn the language as they find it.

As a result, the learning of languages is going to be selectionally highly favored, once the ecological niche in which humans found themselves had been shaped to make extensive use of language. And consequently, to the extent certain properties recur in language across societies, it will be efficient for the learning process to incorporate those into the Language faculty as predispositions. This follows from the Baldwin Effect mechanism.

What this means is that once Language had emerged in a human population, properties of languages that were commonly encountered were quite likely to be in-

corporated into the biological foundations for language learning and use, even if the origins of those properties were to be found in external influences and not in the original structure of the Language faculty. Much of the specific shaping of that faculty, indeed, must on this account have taken place after its original emergence. If this analysis is correct, it vitiates the argument that the Language faculty could not be as specific as much current writing supposes, because there would be no reason for that structure to be in place prior to the emergence of the full-fledged capacity which the Language faculty underlies:

> [T]here is an evolution-based argument argument against typological generalizations being encoded in [the Language faculty]. If [this] is related to a specific genetic endowment of our species and goes back to some sort of protolanguage capacity humans had before spreading out from Africa (see Wunderlich 2004), then it seems implausible that the typological variation which originated afterwards could be a part of this universal language capacity. [Newmeyer 2006, p. 119]

Once we allow for the subsequent shaping of the Language faculty by Baldwinian mechanisms to enhance the ease of learning and use of Language, it is entirely reasonable to assume that just those typological generalizations that recur in many (or all) of the languages to be learned could be incorporated into the faculty underlying their learning.

3.2 The Emergence of a Dilemma

Recalling what has been said above, when we look at the properties that seem plausibly to be part of the human Language faculty, it is hard to show that their presence in particular languages must be a consequence of the structure of that faculty on the basis of their internal structure or distribution alone.

For example, simulations suggest that basic properties like structure sensitivity might plausibly have arisen spontaneously in communicative interactions over time. Studies like those of Perfors et al. suggest that this property might be acquired by modern learners through generalized learning on the basis of overall statistical properties of the data, at least as a possibility. Substantive properties, both in phonology and in syntax, might in turn be driven by the external forces operative in language use and language change, forces that provide an account of the observed regularities that does need to appeal to structural properties of the Language faculty.

On the other hand, to the very extent such forces exert pervasive shaping effects on the languages humans speak, it is not implausible to suggest that evolution, through

Baldwin Effects, is likely to incorporate those same regularities into our genetically determined disposition as human language learners. And the result of that incorporation will be a duplication: the regularities for which we find external grounding in forces of usage, performance, and change will tend to recur as characteristics that the Language faculty expects to find in every language, since that expectation will increase the efficiency of learning the language of the surrounding community.

The importance of this line of argument is that functional or external explanations of cross-linguistic regularities are not, in principle, incompatible with the attribution of those same regularities to the human cognitive capacity for language. But with this comes a serious problem of evidence: for any given regularity that has an external basis, we still need some way to argue for whether or not it also forms part of the Language faculty.

It is quite possible that the external forces working on languages will be reflected only imperfectly as cognitive properties of Language, and such differences as we can find may possibly provide us with ways of teasing the two apart. It is not clear at present, though, how this might be of help. The bottom line seems to be that we have no secure way of identifying a discovered regularity as specifically deriving from the structure of the mind, and not just a product of other influences. From one point of view, distinguishing the two may not matter much if (as seems likely) the structure of the mind is itself ultimately a product of Baldwinian evolution. But since there is no reason to be believe that all such widely distributed regularities have in fact been incorporated into the Language faculty, the issue remains of how to distinguish those that have from those that have not.

4 Conclusions

I return to our central goal as linguists: the devising of ways to infer the properties of the human Language faculty from the available data. In trying to provide a scientific account of the nature and structure of human language, we need to find ways to build substantive theories of the cognitive capacity for Language on the basis of observable evidence. The major problem is that although the observed facts we have to work with clearly flow from instantiations of that capacity, there are very few aspects of those facts whose signature is such that we can construct an argument that they *necessarily* flow from the nature of that faculty and nothing else. Some scholars will rejoice in this conclusion, and see it as a validation of the position that there really is not much content to the notion of a species-specific Language faculty after all. That seems inconsistent, however, with the abundant evidence seemingly showing that the

human capacity to acquire and use language as we do is a quite specific part of our nature as humans, inaccessible in principle to all other known organisms. So I, at least, reject that alternative.

The moral of all this, I think, concerns the nature of scientific inquiry. As linguists, we have developed a rich set of tools for studying Language and languages, and a rich collection of results based on those tools. But it is still possible that although we can ask the question 'What is the structure of the human Language faculty?' the tools available to us are not adequate to provide a real answer of the sort we seek. In the context of broader inquiry into the nature of cognition and the mind, and their relation to the brain, this result would be depressing, but not very surprising.

On the other hand, it would be a mistake to conclude that the absence of solid evidence for the nature of the Language faculty should be taken as solid evidence of the absence of such a faculty. So I must end not with a bang, but a whimper: we need to work toward better tools of investigation. That effort will not be aided, though, by attempts to deny the reality and significance either of a complex and organized human cognitive capacity for Language, or of important forces external to that capacity that have profound effects in shaping the properties of languages — and thereby, of Language.

References

Anderson, Stephen R. 1992. A–Morphous morphology. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2004a. Doctor Dolittle's delusion: Animals and the uniqueness of human language. New Haven: Yale University Press.

Anderson, Stephen R. 2004b. Morphological universals and diachrony. Yearbook of Morphology **2004**. 1–17.

Anderson, Stephen R. & David W. Lightfoot. 2002. The language organ: Linguistics as cognitive physiology. Cambridge: Cambridge University press.

Arnold, Kate & Klaus Zuberbühler. 2006. Semantic combinations in primate calls. Nature **441**. 303.

Aronoff, Mark. 1994. Morphology by itself. Cambridge, MA: MIT Press.

Blevins, Juliette. 2004. Evolutionary phonology: The emergence of sound patterns. Cambridge: Cambridge University Press.

Blevins, Juliette. 2006a. A theoretical synopsis of Evolutionary Phonology. Theoretical Linguistics **32**. 117–166.

Blevins, Juliette. 2006b. Reply to commentaries. Theoretical Linguistics **32**. 245–256.

Bloom, Paul. 2004. Can a dog learn a word? Science **304**. 1605–1606.

Bock, Kathryn, Anne Cutler, Kathleen M. Eberhard, Sally Butterfield, J. Cooper Cutting & Karin R. Humphreys. 2006. Number agreement in British and American English: Disagreeing to agree collectively. Language **82**. 64–113.

Cheetham, Sarah A., Michael D. Thom, Francine Jury, William E. R. Ollier, Robert J. Beynon & Jane L. Hurst. 2007. The genetic basis of individual-recognition signals in the mouse. Current Biology **17**. 1771–1777.

Cheney, Dorothy L. & Robert M. Seyfarth. 1990. How monkeys see the world. Chicago: University of Chicago Press.

Cheney, Dorothy L. & Robert M. Seyfarth. 2007. Baboon metaphysics: The evolution of a social mind. Chicago: University of Chicago Press.

Chomsky, Noam. 1973. Conditions on transformations. A Festschrift for Morris Halle, ed. by S. R. Anderson & P. Kiparsky. pp. 232–286. New York: Holt, Rinehart & Winston.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Corballis, MIchael C. 2007. Recursion, language and starlings. Cognitive Science **31**. 697–704.

Crain, Stephen & Mineharu Nakayama. 1987. Structure dependence in grammar formation. Language **24**. 139–186.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. Current Anthropology **36**. 621–646.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. Cognition **97**. 179–210.

Fouts, Roger. 1997. Next of kin: What chimpanzees have taught me about who we are. New York: William Morrow.

Gardner, R. Allen & Beatrix T. Gardner. 1969. Teaching sign language to a chimpanzee. Science **165**. 664–672.

Garrett, Andrew. 2008. Linguistic universals and language change. *in* Good (2008). pp. 125–143.

Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. Nature **440**. 1204–1207.

Good, Jeff, ed. 2008. Linguistic universals and language change. Oxford: Oxford University Press.

Hare, Brian, Michelle Brown, Christina Williamson & Michael Tomasello. 2002. The domestication of social cognition in dogs. Science **298**. 1634–1636.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? Science **298**. 1569–1579.

Hawkins, John A. 1994. A performance theory of order and constituency. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. Efficiency and complexity in grammars. Oxford: Oxford University Press.

Hockett, Charles F. 1960. Logical considerations in the study of animal communication. Animal sounds and animal communication, ed. by W. E. Lanyon & W. N. Tavolga. pp. 392–430. Washington, DC: American Institute of Biological Sciences.

Hughes, Howard C. 1999. Sensory exotica. Cambridge, MA: MIT Press.

Janda, Richard D. 1984. Why morphological metathesis rules are rare: On the possibility of historical explanation in linguistics. Proceedings of the Annual Meeting of the Berkeley Linguistics Society **10**. 87–103.

Kaminski, Juliane, Josep Call & Julia Fischer. 2004. Word learning in a domestic dog: Evidence for "fast mapping". Science **304**. 1682–1683.

Kiparsky, Paul. 2006. Amphichronic linguistics *vs.* Evolutionary Phonology. Theoretical Linguistics **32**. 217–236.

Kiparsky, Paul. 2008. Universals constrain change; change results in typological generalizations. *in* Good (2008). pp. 23–53.

Labov, William. 2007. Transmission and diffusion. Language **83**. 344–387.

Legate, Julie Anne & Charles D. Yang. 2002. Empirical re-assessment of stimulus poverty arguments. Linguistic Review **19**. 151–162.

Lightfoot, David. 1991. Subjacency and sex. Language and Communication **11**. 67–69.

Lofting, Hugh. 1920. The story of Doctor Dolittle. New York: Frederick A. Stokes.

Maiden, Martin. 1992. Irregularity as a determinant of mophological change. Journal of Linguistics **28**. 285–312.

Maiden, Martin. 2005. Morphological autonomy and diachrony. Yearbook of Morphology **2004**. 137–175.

Marcus, Gary & Matthew Wagers. 2007. Tree structure and the representation of sentences: A reappraisal. Manuscript submitted to *Trends in Cognitive Science*.

McCarthy, John J. 1981. A prosodic theory of non-concatenative morphology. Linguistic Inquiry **12**. 373–418.

Newmeyer, Frederick J. 1990. Speaker-hearer asymmetry as a factor in language evolution: A function explanation for formal principles of grammar. Proceedings of the Berkeley Linguistics Society. Vol. 16.

Newmeyer, Frederick J. 2006. Possible and probable languages. Oxford: Oxford University Press.

Pepperberg, Irene M. 2000. The Alex studies: Cognitive and communicative abilities of grey parrots. Cambridge, MA: Harvard University Press.

Perfors, Amy, Joshua B. Tenenbaum & Terry Regier. 2006. Poverty of the stimulus? a rational approach. Proceedings of the twenty-eighth annual conference of the Cognitive Science Society.

Perfors, Amy, Joshua B. Tenenbaum & Terry Regier. 2008. The learnability of abstract syntactic principles. Manuscript submitted to *Cognition*.

Pinker, Steven. 1994. The language instinct. New York: William Morrow.

Pullum, Geoffrey K. & Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. The Linguistic Review **19**. 9–50.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. A comprehensive grammar of the English language. London and New York: Longman.

Sandler, Wendy & Diane Lillo-Martin. 2006. Sign language and linguistic universals. Cambridge: Cambridge University Press.

Savage-Rumbaugh, Sue, Stuart G. Shanker & Talbot J. Taylor. 1998. Apes, language, and the human mind. New York: Oxford University Press.

Seibert, Charles. 2007. The communicators. New York Times. December 30.

Sherbourne, Amy L., Micheal D. Thom, Steve Paterson, Francine Jury, William E. R. Ollier, Paula Stockley, Robert J. Beynon & Jane L. Hurst. 2007. The genetic basis of inbreeding avoidance in house mice. Current Biology **17**. 2061–2066.

Snyder, William. 2007. Child language: The parametric approach. Oxford: Oxford University Press.

Spencer, Andrew. 2006. Morphological universals. Linguistic universals, ed. by R. Mairal & J. Gil. pp. 101–129. Cambridge: Cambridge University Press.

Weber, Bruce H. & David J. Depew, eds. 2003. Evolution and learning: The Baldwin effect reconsidered. Cambridge, MA: MIT Press.

Wunderlich, Dieter. 2004. Why assume UG? Studies in Language **28**. 615–641.

Yamauchi, Hajime. 2001. The difficulty of the Baldwinian account of linguistic innateness. Advances in artificial life: Proceedings of the 6th European conference on artificial life, ed. by J. Kelemen & P. Sosík. pp. 391–400. Berlin: Springer Verlag.

Yamauchi, Hajime. 2004. Baldwinian Accounts of Language Evolution. PhD thesis. University of Edinburgh.

Zuberbühler, Klaus. 2002. A syntactic rule in forest monkey communication. Animal Behaviour **63**. 293–299.