PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2008) **363**, 3591–3603
doi:10.1098/rstb.2008.0145
Published online 19 September 2008

Cultural evolution: implications for understanding the human language faculty and its evolution

Kenny Smith[1,*] **and Simon Kirby**[2]

[1]*Cognition and Communication Research Centre, Division of Psychology, Northumbria University, Northumberland Building, Northumberland Road, Newcastle NE1 8ST, UK*
[2]*Language Evolution and Computation Research Unit, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK*

Human language is unique among the communication systems of the natural world: it is socially learned and, as a consequence of its recursively compositional structure, offers open-ended communicative potential. The structure of this communication system can be explained as a consequence of the evolution of the human biological capacity for language or the cultural evolution of language itself. We argue, supported by a formal model, that an explanatory account that involves some role for cultural evolution has profound implications for our understanding of the biological evolution of the language faculty: under a number of reasonable scenarios, cultural evolution can shield the language faculty from selection, such that strongly constraining language-specific learning biases are unlikely to evolve. We therefore argue that language is best seen as a consequence of cultural evolution in populations with a weak and/or domain-general language faculty.

Keywords: language; communication; language faculty; cultural evolution; biological evolution

1. INTRODUCTION

When compared with other animals, humans strike us as special. First, we are highly cultural—while culture appears not to be unique to humans (Whiten 2005), its ubiquity in human society and human cognition is highly distinctive. Second, humans have a unique communication system, *language*. Language differs from the communication systems of non-human animals along a number of fairly well-defined dimensions, to be elucidated below.

The first main contention of this article is that the co-occurrence of these two unusual properties is not coincidental: they are causally related. The view that language facilitates human culture is not a new one: it is for this reason that Maynard Smith & Szathmáry (1995) described language as the most recent major evolutionary transition in the history of life on Earth. However, we will argue that the relationship works in the other direction as well—at least some of the distinctive features of human language are adaptations to its cultural transmission, and cultural evolution potentially plays a major role in explaining why the human communication system has the particular features that it does. The most extreme form of this argument, to be expounded in §5, is that a communication system looking a lot like human language is a natural and inevitable consequence of cultural evolution in populations of a particular sort of social learner.

Our second main contention is that a serious consideration of cultural evolution radically changes the kinds of evolutionary stories we can tell about the human capacity for language. Specifically, as we will argue in §4, cultural evolution potentially shields the human language faculty from selection, ruling out the evolution of a strongly constraining and domain-specific language faculty in our species.

2. LANGUAGE DESIGN

What is special about language? In an early attempt to answer this question, Hockett (1960) identified 13 *design features* of language. Of particular relevance are the following three features, whose conjunction (to a first approximation) distinguish language from the communication systems of all other animals.

—*Semanticity*: 'there are relatively fixed associations between elements of messages (e.g. words) and recurrent features or situations of the world around us' (Hockett 1960, p. 6).
—*Productivity*: '[language provides] the capacity to say things that have never been said or heard before and yet to be understood by other speakers of the language' (Hockett 1960, p. 6).
—*Traditional (i.e. cultural) transmission*: 'Human genes carry the capacity to acquire a language, and probably also a strong drive towards such acquisition, but the detailed conventions of any one language are transmitted extragenetically by learning and teaching' (Hockett 1960, p. 6).

In isolation, each of these features is not particularly rare. Limited non-productive semantic communication systems are common in the natural world, notably in the alarm-calling behaviours of various species (e.g. diverse species of bird and monkey; Marler 1955; Cheney & Seyfarth 1990; Evans *et al.* 1993; Zuberbuhler 2001).

* Author for correspondence (kenny.smith@northumbria.ac.uk).

One contribution of 11 to a Theme Issue 'Cultural transmission and the evolution of human behaviour'.

This journal is © 2008 The Royal Society

However, such semantic communication systems are not traditionally transmitted: the consensus view is that there is no vocal learning of the form of alarm calls in such systems, although there may well be a role for learning in establishing the precise situations under which calls must be produced, or how calls must be responded to (usage and comprehension learning, respectively: Slater 2005).

Traditionally transmitted communication systems also exist in non-humans—in mammals (seals, bats, whales and dolphins) and in birds (most notably the oscine passerines, the songbirds). Aspects of the production of communicative signals in all of these groups show sensitivity to input and (particularly in the case of songbirds and whales) patterns of local dialects characteristic of cultural transmission. Furthermore, some bird song (e.g. that of the starling, willow warbler and Bengalese finch; Eens 1997; Gil & Slater 2000; Okanoya 2004) is productive, to the extent that these songs are constructed according to rules that generate several possible songs with the same underlying structure. However, none of these systems rise above more than a superficial degree of semanticity: this is clearest in the songbirds, where song seems to serve a dual function as a means of attracting mates and repelling rivals (Catchpole & Slater 1995). Indeed, the same song being sung by a particular individual can be differentially interpreted depending on the identity of the listener, with female listeners interpreting it as sexual advertisement and males interpreting it as territory defence.

Finally, traditionally transmitted and semantic communication systems seem to exist, to a limited extent, in the gestural communication systems of our nearest extant relatives, the apes. Pairs of apes develop, through repeated interactions, meaningful gestures that can be subsequently used to communicate about various situations (feeding, play, sex, etc.; Call & Tomasello 2004). However, such systems are not productive: each sign is underpinned by a rather laborious history of interaction (through a process known as ontogenetic ritualization; Tomasello 1996) and consequently such systems cannot be expanded to include meaningful novel signals.

While there are, in principle, several ways of designing a productive semantic communication system, in human language these features are underpinned by four subsidiary design features (two from Hockett 1960, the others implicit in his account).

—*Arbitrariness*: 'the ties between the meaningful message elements and their meanings can be arbitrary' (Hockett 1960, p. 6). Arbitrary signals of this sort can be contrasted with, for example, signals that have meaning by resemblance (icons) or signals that are causally related to their meanings (indexes, such as the gestural signals established by ontogenetic ritualization).

—*Duality of patterning*: 'The meaningful elements in any language (e.g. words)…constitute an enormous stock. Yet they are represented by small arrangements of a relatively very small stock of distinguishable sounds which are in themselves wholly meaningless' (Hockett 1960, p. 6). Languages have a few tens of phonemes that are combined to form tens of thousands of words.

—*Recursion*: a signal of a given category can contain component parts that are of the same category, and this embedding can be repeated without limit. For example, 'I think he saw her' is a sentence of English, which contains an embedded sentence, 'he saw her'. This embedding can be continued indefinitely ('I think she said he saw her', 'You know I think she said he saw her', and so on), yielding an infinite number of sentences.

—*Compositionality*: the meaning of a complex signal is a function of the meaning of its parts and the way in which they are combined (Krifka 2001). For example, the sentence 'She slapped him' consists of four component parts—'she', 'him', 'slap' and '-ed' (the past tense marker). The meaning of the sentence is determined by the meaning of these parts and the order in which they are combined—'She slapped him' differs predictably in meaning from 'They slapped him' and 'She kicked him', but also from 'He slapped her', where the order of the male and female pronouns has been reversed.

The combination of these four subsidiary features results in a system that is productive and semantic. Duality of patterning allows for the generation of an extremely large set of basic communicative units from a small inventory of discriminable sounds.[1] Arbitrariness allows those basic units to be mapped onto the world in a flexible fashion, without additional constraints of iconicity or indexicality. Recursion allows that large inventory of basic units to be combined to form a truly open-ended system. Finally, compositional structure makes the interpretation of novel utterances possible—in a recursive compositional system, if you know the meaning of the basic elements and the effects associated with combining them, you can deduce the meaning of any utterance in the system, including infinitely many entirely novel utterances.

Again, we can ask to what extent these subsidiary design features are realized in the communication systems of non-human animals. Perhaps, as we might expect, the productive systems highlighted above (e.g. the song of certain species of bird) appear to adopt rudiments of duality of patterning, in that they consist of recombinable subunits (notes or syllables). Semantic systems show limited amounts of arbitrariness (in alarm-calling systems) and compositionality (in the boom alarm combination call in Campbell's monkeys, where the preceding boom serves to change the meaning of the subsequent alarm call, or reduces its immediacy; Zuberbuhler 2002). Finally, there is little evidence for recursion in the communicative behaviour of non-humans—indeed, Hauser *et al.* (2002) argue that it is entirely absent from the cognitive repertoires of all non-human species, although this point is still a matter for debate (Kinsella in press).

The foregoing discussion suggests that language is unusual owing to the particular bundle of features it possesses, and the pervasiveness of those features in language, rather than possession of any individual unique element. Why is language designed like this? This at first seems like a fairly straightforward question to answer: language is designed like this because these design features make for a useful communication system, specifically a system with open-ended expressivity. A population of individuals sharing such a system

can in principle communicate with each other about anything they chose, including survival relevant issues such as where to find food and shelter, how to deal with predators and prey, how social relationships are to be managed, and so on. Furthermore, each individual only requires a finite (and fairly small) set of cognitive resources to achieve this expressive range—a few tens of speech sounds, a few tens of thousands of meaningful words created from those speech sounds and a few hundred grammatical rules constraining the combination of those words into meaningful sequences.

However, the fact that human language makes good design sense in various ways does not explain how language came to have these properties. To truly answer the question 'why is language designed like this?' we need to establish the *mechanisms* that explain this fit between function and form: how did the manifest advantages of such a linguistic system become realized in language as a system of human behaviour?

We will review two potential mechanisms below: one that explains the fit between function and form as arising from the biological evolution of the human language faculty (in §3*a*), and the other that views it as a consequence of the cultural evolution of language itself (in §3*b*). In both cases we will see how these contrasting explanations have been applied to the specific question of the evolution of compositionality, one of the language design features subserving productivity. The evolution of compositionality has received a great deal of attention in the literature, which is why we focus on it here. Ultimately, all these design features will require such explanation, and a similar research effort aimed at understanding the evolution of duality of patterning (in both biological and cultural terms) is underway (see, e.g. Nowak & Krakauer 1999; Nowak *et al.* 1999; Oudeyer 2005; Zuidema & de Boer in press). In §4, we will turn to the issue of interactions between biological and cultural evolutionary accounts.

3. TWO EXPLANATIONS FOR LANGUAGE DESIGN

(a) *An explanation from evolutionary biology*

The uniqueness of human language must have some biological basis—there must be *some* feature of human biology that results in this unusually rich, expressive system of communication in our species alone. One obvious biological adaptation for vocal communication is the unusual structure of the human vocal tract, which provides a wide range of highly discriminable sounds (see Fitch 2000, for review). However, language is more than just speech: the design features picked out above are ambivalent as to the modality of the system in question, and while adaptations for speech are important in that they provide a good clue as to the age, modality and selective importance of language in human evolution history, they are peripheral to what we see as the fundamental design features of language.

Moving beyond the productive apparatus, then, what is our biological endowment for language? Linguists have approached this question via a consideration of the problem of language acquisition: how do children acquire a complex and richly structured language with apparent ease at a relatively early stage in their lives? The difficulty of reconciling the precocity of language acquisition with the complexity of language leads to the fairly prominent view that humans must be endowed with a highly structured and constraining language-specific mental faculty that, to a large degree, prefigures much of the structure of language. One of the main cornerstones of this argument is known as the *argument from the poverty of the stimulus*, a term introduced by Chomsky (1980): the data language must be learned from lack direct evidence for a number of features of the grammars that children must learn. If children end up with grammars that contain features for which they received no evidence in their input, then (the argument goes) those features of language must be prefigured in the acquisition device. Rather than a flexible and open-ended process of social learning, language acquisition is 'better understood as the growth of cognitive structures along an internally directed course under the triggering and partially shaping effect of the environment' (Chomsky 1980, p. 34). Design features of language are then naturally explained as features of the internally directed course of acquisition.

This account has been so successful that it now constitutes a scientific orthodoxy, at least in some form (and despite a lack of empirical support for stimulus poverty arguments; Pullum & Scholz 2002). However, even if we accept that language is the way it is because the language faculty forces it to be that way, this simply pushes the question back one step: why is the language faculty designed the way it is?

A well-established solution to such questions in biological systems is that of evolution by natural selection. In a highly influential article, Pinker & Bloom (1990) argued that this solution can be applied to language, assuming (following the argument above) that language is to a non-trivial extent a biological capacity: 'It would be natural, then, to expect everyone to agree that human language is the product of Darwinian natural selection. The only successful account of the origin of complex biological structure is the theory of natural selection' (Pinker & Bloom 1990, p. 707). To support this claim, Pinker and Bloom provided a number of arguments that language, in general, offers reproductive pay-offs. Some of these arguments have been rehearsed in brief above—for example, the open-ended communicative possibilities that language affords might plausibly have reproductive consequences. The compositionality of language should then be viewed in the context of the functional advantages that compositionality affords: language must be compositional because the language faculty forces it to be that way, and the language faculty must be designed like that because a compositional language offers reproductive advantages (what could be more useful than the ability to produce entirely novel utterances about a wide range of situations, and have them understood?).

Nowak *et al.* (2000) develop this argument in a mathematical model of the evolution of compositionality. They assume two types of language learner: those who learn a *holistic* (non-compositional) mapping between meanings and signals, and those who learn a simple compositional system. They consider the case of populations of such learners converged on stable

languages and find that, as expected, populations of compositional learners have higher within-population communicative accuracy than learners who learn in a holistic fashion, assuming that the number of events that individuals are required to communicate about is not small. Under conditions where there are a large number of fitness relevant situations to communicate about, the productivity advantage of compositional language pays off in evolutionary terms.

(b) *An explanatory role for cultural evolution*

The biological account of the evolution of language sketched by Pinker & Bloom (1990) is an attractive one. However, their assertion that it is the *only* possible explanation for the evident adaptive design of language is based on a false premise: while we would accept that '[t]he only successful account of the origin of complex biological structure is the theory of natural selection', we would dispute that language is solely a biological structure. As already discussed in §2, language is manifestly a socially learned, culturally transmitted system. Individuals acquire their knowledge of language by observing the linguistic behaviour of others, and go on to use this knowledge to produce further examples of linguistic behaviour, which others can learn from in turn (see, e.g. Andersen 1973; Hurford 1990; Kirby 1999). We have previously termed this process of cultural transmission *iterated learning*: learning from the behaviour of another, where that behaviour was itself acquired through the same process of learning. The fact that language is socially learned and culturally transmitted opens up a second possible explanation for the design features of language: those features arose through *cultural*, rather than biological, evolution. Rather than traditional transmission being another design feature that a biological account must explain, traditional transmission is the feature from which the other structural properties of language spring.

One of the primary objections to this account has already been stated: the argument from the poverty of the stimulus suggests that language learning should be impossible, and therefore the apparent social learning of language must be illusory. However, cultural transmission offers a potential solution to this conundrum that does not require an assumption of innateness—while the poverty of the stimulus poses a challenge for individual learners, language adapts over cultural time so as to minimize this problem, because its survival depends upon it. We have dubbed this process *cultural selection for learnability*:

> in order for linguistic forms to persist from one generation to the next, they must repeatedly survive the processes of expression [production] and induction [learning]. That is, the output of one generation must be successfully learned by the next if these linguistic forms are to survive.
>
> (Brighton *et al.* 2005, p. 303).

Cultural selection for learnability offers a solution to the conundrum posed by the argument from the poverty of the stimulus:

> Human children appear preadapted to guess the rules of syntax correctly, precisely because languages evolve so as to embody in their syntax the most frequently guessed patterns. The brain has coevolved with respect to language, but languages have done most of the adapting
>
> (Deacon 1997, p. 122).

As a consequence of cultural selection for learnability, the poverty of the stimulus problem induces a pressure for languages that are learnable from the kinds of data learners can expect to see: 'the poverty of the stimulus solves the poverty of the stimulus' (Zuidema 2003), but only if we look at language in its proper context of cultural transmission.

How does cultural selection for learnability explain compositionality? As discussed above, languages are infinitely expressive (due to the combination of recursion and compositionality). However, such languages must be transmitted through a finite set of learning data. We call this mismatch between the size of the system to be transmitted and its medium of transmission the *learning bottleneck* (Kirby 2002*a*; Smith *et al.* 2003). Compositionality provides an elegant solution to this problem: to learn a compositional system, a learner must master a finite set of words and rules for their combination, which can be learned from a finite set of data but can generate a far larger system. This fit between the form of language (it is compositional) and a property of the transmission medium (it is finite but the system passing through it is infinite) is suggestive, and computational models[2] of the iterated learning process (known as *iterated learning models*) have repeatedly demonstrated that cultural evolution driven by cultural selection for learnability can account for this goodness of fit.

In their simplest form, iterated learning models consist of a chain of simulated language learner/users (known as *agents*). Each agent in this chain learns their language by observing a set of utterances produced by the preceding agent in the chain, and in turn produces example utterances for the next agent to learn from. The treatment of language learning and language production varies from model to model, with similar results having been shown for a fairly wide range of models (see Kirby 2002*b*, for review). In all cases, the crucial feature is that these agents are capable of learning *both* holistic and compositional languages: they can memorize a holistic meaning–signal mapping, but they can also generalize to a partially or wholly compositional language when the data they are learning from merit such generalization. In other words, compositionality is not hard-wired into the language learners.

These models show that the presence of a learning bottleneck is a key factor in determining the evolution of compositionality. In conditions where there is no learning bottleneck, the set of utterances produced by one agent for another to learn from covers (or is highly likely to cover) the full space of expressions that any agent will ever be called upon to produce. This is of course impossible for human language, where the set of expressions is extremely large or infinite. On the other hand, where there is a learning bottleneck, there is some (typically high) probability that learner/users will subsequently be called upon to produce a novel expression and will therefore be required to generalize.

It is this pressure to generalize that leads to the evolution of compositional languages: whereas compositional languages or compositional parts of language are generalizable and can therefore be successfully transmitted through a bottleneck, holistic (sub)systems are by definition not generalizable and will be subject to change. This adaptive dynamic has been used to explain the putative transition from a hypothesized holistic protolanguage (see, e.g. Wray 1998) to a modern compositional system.

Compositionality can therefore be explained as a cultural adaptation *by language* to the problem of transmission through a learning bottleneck. Note that throughout this explanation we are appealing to a different notion of function to that discussed in §3*a*: while Pinker & Bloom (1990) and Nowak *et al.* (2000) appealed to compositionality as an adaptation for communication, it is also a potential adaptation for cultural transmission.

(c) *Explanations for language design: conclusions*

Based on the preceding discussion, we believe there is at least as good a case for explaining one design feature of language as a cultural, rather than biological, adaptation. Of course, cultural and biological evolution can happily work in the same direction, and in some cases (such as the evolution of duality of patterning), the distinction between cultural and biological mechanisms seems rather minor: in both cases, the linguistic system is being optimized for its ability to produce numerous maximally discriminable signals (see Zuidema & de Boer in press). For the case of compositionality, there is a difference in the function being optimized by the two alternative adaptive mechanisms: communicative usage under the biological account and learnability under the cultural account. However, given that the conclusions in all cases are the same—compositionality and duality of patterning are good ideas—the temptation might be to gloss over the differences in mechanism leading to these adaptations.

However, the two competing explanations make rather different predictions about the structure of the human language faculty, which, as linguists, is our ultimate object of study. Biological accounts suggest a fairly direct mapping between properties of the language faculty and properties of language, whereas the cultural accounts propose a rather more opaque relationship. This opaque relationship between the language faculty and language design significantly complicates evolutionary accounts of the language faculty, as we will see in §4.

4. IMPLICATIONS FOR UNDERSTANDING THE EVOLUTION OF THE LANGUAGE FACULTY

Acknowledging a role for cultural processes in the evolution of language might not actually have any consequences for our understanding of the human language faculty in its present form. For example, it would be perfectly consistent to accept that compositionality was initially a cultural adaptation by language to maximize its own transmissibility, but to argue that this feature has subsequently been assimilated into the language faculty. Arguments of this sort, appealing to the mechanism known as the Baldwin effect or genetic assimilation (Baldwin 1896; Waddington 1975) are in fact reasonably common in evolutionary linguistics. Pinker & Bloom (1990) appeal directly to the notion that initially learned aspects of language will gradually be assimilated into the language faculty, with any remaining residue of learning being a result of diminishing returns from assimilation. More detailed arguments and formal models of this effect are presented by Briscoe (2000, 2003). The prevalence of assimilational accounts in the evolutionary linguistics literature can be neatly characterized in a parenthetical remark from Ray Jackendoff: 'I agree with practically everyone that the 'Baldwin effect' had something to do with it' (Jackendoff 2002, p. 237).

We believe this picture is fundamentally wrong, for at least two reasons. First, we have previously argued that the coevolution of culturally transmitted systems and biological predispositions for learning those system is somewhat problematic, due to time lags introduced by the cumulative and frequency-dependent nature of cultural evolution (Smith 2004). Our focus here will be on a second problem: social transmission and cultural evolution can shield the language-learning machinery of individuals from selection, such that two rather different language faculties end up being behaviourally equivalent and selectively neutral. This neutrality rules out evolution of more strongly constraining language faculties via assimilational processes. Furthermore, selection itself can drive evolution precisely into these conditions—a plausible evolutionary scenario sees natural selection acting on the language faculty selecting for conditions where the language faculty is shielded from selection.

(a) *The link between language structure and biological predispositions*

A useful vehicle to develop this argument is provided by a series of recent papers that seek to explicitly address the link between language structure, biological predispositions and constraints on cultural transmission (Griffiths & Kalish 2005, 2007; Kirby *et al.* 2007; Griffiths *et al.* 2008). These papers are based around iterated learning models where learners apply the principles of Bayesian inference to language acquisition. A learner's confidence that a particular grammar h accounts for the linguistic data d that they have encountered is given by

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h'} P(d|h')P(h')}, \tag{4.1}$$

where $P(h)$ is the prior belief in each grammar and $P(d|h)$ gives the probability that grammar h could have generated the observed data. Based on the posterior probability of the various grammars, $P(h|d)$, the learner then selects a grammar and produces utterances that will form the basis, through social learning, of language acquisition in others. This learning model provides a transparent division between the contribution of the learning bias of individuals prior to encountering data (the prior) and the observed data in shaping behaviour. We will equate prior bias with the innate language

faculty of individuals—while Griffiths and Kalish rightly point out that the prior need not necessarily take the form of an innate bias at all (e.g. it might be derived from non-linguistic data), ours is a possible and (we believe) natural interpretation.

Within this framework, Griffiths & Kalish (2005, 2007) showed that cultural transmission factors (such as noise or a learning bottleneck imposed by partial data) have *no* effect on the distribution of languages delivered by cultural evolution: the outcome of cultural evolution is solely determined by the prior biases of learners, given by $P(h)$. In other words, only the structure of the language faculty matters in determining the outcomes of linguistic evolution. Griffiths & Kalish (2007) and Kirby *et al.* (2007) demonstrated that this result is a consequence of the assumption that learners select a grammar with probability proportional to $P(h|d)$—if learners instead select the grammar that maximizes the posterior probability (known as MAP learners), then cultural transmission factors play an important role in determining the distribution of languages delivered by cultural evolution: while the distribution of languages produced by cultural evolution will be approximately centred on the language most favoured by the prior, different transmission bottlenecks (for example) lead to different distributions. Furthermore, and crucially for our arguments here, for MAP learners the strength of the prior bias is irrelevant over a wide range of the parameter space (Kirby *et al.* 2007)—it matters which language is most favoured in the prior, but not *how much* it is favoured over alternatives.

These models suggest two candidate components of the innate language faculty: first, the prior bias, $P(h)$, and second, the strategy for selecting a grammar based on $P(h|d)$—sampling proportional to $P(h|d)$, or selecting the grammar that maximizes $P(h|d)$. Furthermore, they allow us to explore, using a single framework, the evolution of the language faculty under two rather different conceptions of what cultural evolution does: either it ensures that the distribution of languages ultimately delivered by cultural evolution reflects exactly the biases of language learners, regardless of transmission factors such as the learning bottleneck (if learners sample from the posterior), or it allows such transmission factors to play a significant role in determining the languages we see, and obscures differences in the nature of individual language faculties (if learners maximize). We can therefore straightforwardly extend models of this sort to ask whether cultural evolution alters the plausibility of scenarios regarding the biological evolution of the language faculty and, if so, under what conditions.

(b) *The model of learning and cultural transmission*

We adopt Kirby *et al.*'s (2007) model of language and language learning. Full details are given below, but to briefly summarize: languages are mappings between meanings and signals, where learners have a parametrizable preference for regular languages. Implicit in this model of learning is a model of cultural transmission, which can be used to calculate the stable outcomes of cultural evolution.

A language consists of a system for expressing m meanings, where each meaning can be expressed using one of k means of expression, called *signal classes*. In a perfectly regular (or systematic) language the same signal class will be used to express each meaning—for example, the same inflectional paradigm will be used for each verb, or the same compositional rules will be used to construct an utterance for each meaning. By contrast, in a perfectly irregular system each meaning will be associated with a distinct signal class—each verb an irregular, each complex utterance an idiom.

We will assume two types of prior bias. For *unbiased* learners, all grammars have the same prior probability: $P(h) = 1/k^m$. *Biased* learners have a preference for languages that use a consistent means of expression, such that each meaning is expressed using the same signal class. Following Kirby *et al.* (2007), this prior is given by the expression

$$P(h) = \frac{\Gamma(k\alpha)}{\Gamma(\alpha)^k \Gamma(m + k\alpha)} \prod_{j=1}^{k} \Gamma(n_j + \alpha), \tag{4.2}$$

where $\Gamma(x) = (x-1)!$ when x is an integer; n_j is the number of meanings expressed using class j; and $\alpha \geq 1$ determines the strength of the preference for consistency: low α gives a strong preference for consistent languages and higher α leads to a weaker preference for such languages.[3] Kirby *et al.* (2007) justify the use of this particular prior distribution on the basis that Bayesian inference with this prior can be viewed as hypothesis selection based on minimum description length principles, which has been argued (convincingly, to our minds) to be relevant to cognition in general and language acquisition in particular (see, e.g. Brighton 2003; Chater & Vitányi 2003). However, the precise details of this prior are unimportant for our purposes here: any prior that provides a (partial) ordering over hypotheses supports our conclusions.

The probability of a particular dataset d (consisting of b observed meaning–form pairs: b gives the bottleneck on language transmission) being produced by an individual with grammar h is

$$P(d|h) = \prod_{\langle xy \rangle \in d} P(y|x,h) \frac{1}{m}, \tag{4.3}$$

where all meanings are equiprobable; $\langle xy \rangle$ is a meaning–signal pair consisting of a meaning x and a signal class y; and $P(y|x,h)$ gives the probability of y being produced to convey x given grammar h and noise ϵ:

$$P(y|x,h) = \begin{cases} 1-\epsilon & \text{if } y \text{ is the class corresponding to } x \text{ in } h \\ \dfrac{\epsilon}{k-1} & \text{otherwise} \end{cases}. \tag{4.4}$$

Bayes' rule can then be applied to give a posterior distribution over hypotheses, given a particular set of utterances. This posterior distribution is used by a learner to select a grammar, according to one of two strategies. Sampling learners simply select a grammar proportional to its posterior probability: $P_{\mathrm{L}}(h|d) = P(h|d)$. MAP learners select randomly from among those

Table 1. $P(h)$ for three grammars given various types of bias (unbiased, weak bias ($\alpha=40$) and strong bias ($\alpha=1$), denoted by *u*, *bw* and *bs*, respectively), and the frequency of those grammars in the stationary distribution for sampling and MAP learners. (Grammars are given as strings of characters, with the first character giving the class used to express the first meaning and so on: *aaa* is a perfectly regular language and *abc* is perfectly irregular. All results here and throughout the paper are for $m=3$, $k=3$, $b=3$ and $\epsilon=0.1$. For MAP learners, qualitatively similar results are obtainable for a wide range of the parameter space. For sampling learners, as shown by Griffiths & Kalish (2007), qualitatively similar results are obtainable for any region of the parameter space where $\epsilon>0$.)

	$P(h)$			Q^*, sampler			Q^*, maximizer		
h	*u*	*bw*	*bs*	*u*	*bw*	*bs*	*u*	*bw*	*bs*
aaa	0.0370	0.0389	0.1	0.0370	0.0389	0.1	0.0370	0.2499	0.2499
aab	0.0370	0.0370	0.0333	0.0370	0.0370	0.0333	0.0370	0.0135	0.0135
abc	0.0370	0.0361	0.0167	0.0370	0.0361	0.0167	0.0370	0.0014	0.0014

grammars with the highest posterior probability:

$$P_{\mathrm{L}}(h|d) = \begin{cases} 1/|\mathcal{H}| & \text{if } h \in \mathcal{H} \\ 0 & \text{otherwise} \end{cases}, \tag{4.5}$$

where $\mathcal{H}$ is the set of hypotheses for which $P(h|d)$ is at a maximum.

A model of cultural transmission follows straightforwardly from this model of learning: the probability of a learner at generation n arriving at grammar h_n given exposure to data produced by grammar h_{n-1} is simply

$$P(h_n = i|h_{n-1} = j) = \sum_d P_{\mathrm{L}}(h_n = i|d)P(d|h_{n-1} = j). \tag{4.6}$$

The matrix of all such transition probabilities is known as the Q matrix (Nowak *et al.* 2001): entry Q_{ij} gives the transition probability from grammar j to i. As discussed in Griffiths & Kalish (2005, 2007), the stable outcome of cultural evolution (the *stationary distribution* of languages) can be calculated given this Q matrix and is proportional to its first eigenvector.[4] We will denote the probability of grammar i in the stationary distribution as Q_i^*.

Table 1 gives some example prior probabilities and stationary distributions, for various strengths of prior and both selection strategies. As shown in table 1, strength of prior determines the outcome of cultural evolution for sampling learners, but is unimportant for MAP learners as long as some bias exists.

(c) *Evaluating within-population communicative accuracy*

In order to calculate which selection strategies and priors are favoured by biological evolution, we need to define a measure that determines reproductive success and therefore predicts the evolutionary trajectory of the language faculty. One possibility (following Nowak *et al.* 2000) is that the relevant quality is the extent to which members of a genetically homogeneous population can communicate with one another: the probability that any two randomly selected individuals drawn from a population at equilibrium (at the stationary distribution provided by cultural evolution) will share the same language.[5] This quantity, C, is simply

$$C = \sum_h (Q_h^*)^2, \tag{4.7}$$

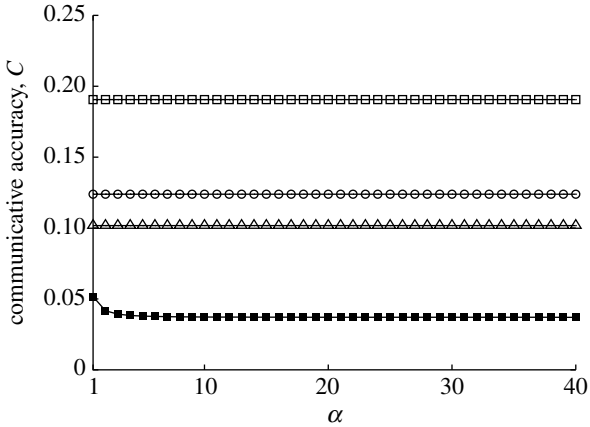

Figure 1. Fitness of sampler (filled squares) and MAP (open squares, $b=3$; circles, $b=6$; triangles, $b=9$) learners. Results for MAP learners are for $\epsilon=0.1$ and for various bottlenecks (values of b). C for sampler populations is the same regardless of noise level and amount of data, as these factors have no impact on the stationary distribution.

where the sum is over all possible grammars. The within-population communicative accuracies of various combinations of strength of prior and hypothesis selection strategy are given in figure 1. Three results are apparent from this figure. First, in sampling populations, stronger priors (lower values of α) yield higher communicative accuracy. Stronger priors make for a less uniform stationary distribution, with more regular languages being over-represented. This skewing away from the uniform distribution results in greater within-population coherence. By contrast, strength of prior is irrelevant in populations of MAP learners: it has no impact on the stationary distribution, and as a result there is no communicative advantage associated with stronger priors.

Finally, MAP populations always have higher within-population communicative accuracy than sampling learners. As shown by Kirby *et al.* (2007), and illustrated in table 1, in MAP populations the differences between languages are amplified by cultural evolution, and the extent of the amplification is inversely proportional to the size of the transmission bottleneck: tight bottlenecks give greater amplification, such that the *a priori* most likely language is greatly overrepresented in the stationary distribution. By contrast, cultural evolution in sampling populations

returns the unmodified prior distribution. Amplifying the differences between languages provided by the prior increases the likelihood that two individuals from a population will share the same language, and therefore yields higher C. This analysis therefore suggests that evolution should favour MAP learning over sampling, leading to selective neutrality with respect to the strength of prior biases: strength of prior bias is shielded in MAP populations.

(d) *Evaluating evolutionary stability*

While the preceding analysis tells us which priors and hypothesis selection strategies are objectively best from a communicative point of view, it tells us nothing about the evolvability of those features. A more satisfying solution is to evaluate the evolutionary stability of hypothesis selection strategies and priors. In order to do this, we require a slightly more detailed means of evaluating how communication influences reproductive success. We make the following initial assumptions (see below for a slightly different set of assumptions): (i) a population consists of several subpopulations, (ii) each subpopulation has converged on a single grammar through social learning, with the probability of each grammar being used by a subpopulation given by that grammar's probability in the stationary distribution (as suggested by the analysis provided in Griffiths *et al.* 2008, §5*a*), and (iii) natural selection favours learners who arrive at the same grammar[6] as their peers in a particular subpopulation, where peers are other learners exposed to the language of the subpopulation. Given these assumptions, the communicative accuracy between two individuals A and B is given by

$$\mathrm{ca}(A,B) = \sum_{h}\sum_{h'} Q^{A}_{hh'} \cdot Q^{B}_{hh'} \cdot Q^{*}_{h'}, \tag{4.8}$$

where the superscripts on Q indicate that learners A and B may have different selection strategies and priors. The *relative communicative accuracy* of a single learner A with respect to a large and homogeneous population of individuals of type B is therefore given by $\mathrm{rca}(A,B) = \mathrm{ca}(A,B)/\mathrm{ca}(B,B)$. Where this quantity is greater than 1, the combination of selection strategy and prior (the *learning behaviour*) of individual A offers some reproductive advantage relative to the population learning behaviour, and may come to dominate the population. Where relative communicative accuracy is less than 1, learning behaviour A will tend to be selected against; and when relative communicative accuracy is 1, both learning behaviours are equivalent and genetic drift will ensue. Following Maynard Smith & Price (1973), the conditions for evolutionary stability for a behaviour of interest, I, are therefore: (i) $\mathrm{rca}(\mathcal{J},I) < 1$ for all $\mathcal{J} \neq I$ (populations of type I resist invasion by all other learning behaviours) or (ii) $\mathrm{rca}(\mathcal{J},I) = 1$ for some $\mathcal{J} \neq I$, but in each such case $\mathrm{rca}(I,\mathcal{J}) > 1$. The second condition covers situations where the minority behaviour $\mathcal{J}$ can increase by drift to the point where encounters between type $\mathcal{J}$ individuals become common, at which point type I individuals are positively selected for and the dominance of behaviour I is re-established.

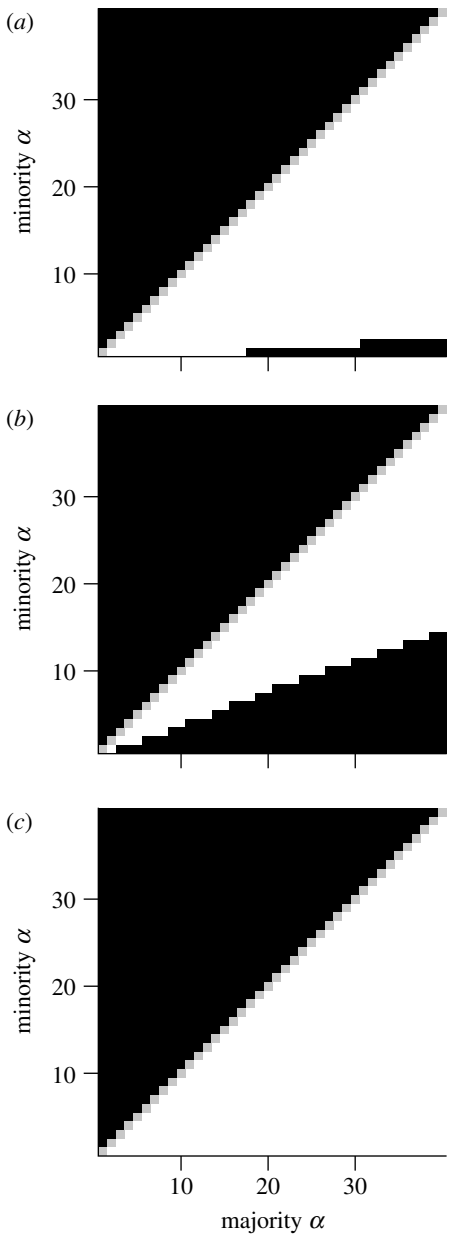

Figure 2. (*a*,*b*) Relative communicative accuracy of various combinations of strengths of prior in sampling populations ((*a*) $b=3$; (*b*) $b=9$). Black cells indicate the minority α-value will be selected against ($\mathrm{rca}<1$), white cells indicate the minority α-value will be selected for ($\mathrm{rca}>1$) and grey cells (seen on the diagonal) indicate selective neutrality ($\mathrm{rca}=1$). (*c*) An example result given the global evaluation of communicative accuracy ($b=9$); see text for details.

Again, we can use this model to ask several evolutionary questions. First, what does the evolution of the strength of prior preference look like in sampling and MAP populations? Do different learning models (and their associated differences in the outcomes of cultural evolution) change the evolutionary stability of different strengths of prior? Figure 2*a*,*b* shows the results of numerical calculations of evolutionary stability in sampling populations.

As can be seen in figure 2, in sampling populations there is selection against weaker priors and selection for a stronger prior bias than the population norm. The stronger the strength of the prior bias of the population, the more dominant languages favoured by that prior bias (i.e. the regular languages) will be, and the greater

Table 2. Relative communicative accuracy of various strengths of prior in (*a*) sampling and (*b*) maximizing populations.

		majority		
		u	*bs*	*bw*
(*a*) minority	*u*	—	0.81	0.9997
	bs	0.98	—	0.82
	bw	0.99998	0.99	—
(*b*) minority	*u*	—	0.45	0.45
	bs	1	—	1
	bw	1	1	—

Table 3. Relative communicative accuracy of the minority hypothesis selection strategy for strong ($\alpha=1$), weak ($\alpha=40$) and flat priors.

		majority	
	minority	sampling	maximizing
strong	sampling	—	0.6
	maximizing	1.39	—
weak	sampling	—	0.38
	maximizing	1.14	—
flat	sampling	—	0.88
	maximizing	1.12	—

the advantage in being biased in favour of acquiring those languages. In maximizing populations, strength of prior is *always* selectively neutral—regardless of the α of the majority and minority, rca = 1. These results are therefore not plotted. In accordance with the analysis given in §4*c*, the evolutionary stability of different strengths of prior bias differs markedly across sampling and maximizing populations.

The degree to which stronger priors are favoured in sampling populations is somewhat sensitive to the strength of the population prior, however. For example, populations with a very weak prior bias in favour of regularity ($\alpha=40$) resist invasion by mutants with much stronger prior preferences for regularity ($\alpha\approx1$)—given the relatively uniform distribution of languages in these populations, the added ease of acquiring highly regular languages provided by a very strong preference for such languages is counteracted by the reduced likelihood of being born into populations speaking such languages. This phenomenon becomes more marked given larger bottlenecks (higher *b*, meaning that learners observe more data during learning), as this reduces the likelihood of misconvergence for individuals with weaker priors.

As shown in table 2, this tendency for weaker majority priors to reduce the extent to which strengthened priors can invade also pertains for populations where the majority have a flat, unbiased prior: in such populations every language is equally probable, and any bias to acquire a particular language is penalized due to the consequent decreased ability to acquire the other languages. Consequently, there are two evolutionarily stable strategies (ESSs) in sampling populations: the strongest possible prior ($\alpha=1$) or a completely unbiased prior.

In contrast with the situation in sampler populations, selection in MAP populations is neutral with respect to bias strength. This follows naturally from the insensitivity of MAP learners to the strength of their prior biases—all that matters is the ranking of those languages under the prior. Consequently, strength of the prior makes no difference to the stationary distribution provided by cultural evolution (as shown in table 1), nor to the ability of learners to acquire those languages. Consequently, strength of prior is selectively neutral. However, as shown in table 2, this neutrality with respect to strength of prior bias only applies when there is *some* prior bias—a completely flat prior is not evolutionarily stable (through the second clause of the definition given above). Consequently, the set of all non-flat priors constitutes the evolutionarily stable set (Thomas 1985) for MAP learners.

These two models of learning, which make different predictions about the outcomes of cultural evolution, also make different predictions about the coevolution of the language faculty and language. In particular, the learning model that predicts an opaque relationship between bias and outcomes of cultural evolution (MAP learning) also predicts extensive shielding of the language faculty from selection, and certainly no positive selection for increasingly constraining language faculties.

We can ask a final evolutionary question: is it better to be a sampling learner or a maximizing learner? As shown in table 3, maximizing is the ESS, regardless of bias strength. MAP learners boost the probability that the most likely grammar will be learned, and are consequently more likely to arrive at the same grammar as some other learner exposed to the same data-generating source. Given the fitness function that emphasizes convergence on the same language as other members of the population, maximizing is obviously the best thing to do.

It is worth noting that this pattern of results is not dependent on the assumption that learners are rewarded for arriving at the same grammar as other individuals reared in the same linguistically homogeneous subpopulation. While this strikes us as a reasonable fitness function, it is not the only possible one. For example, we could make the following rather different set of assumptions: (i) rather than consisting of several subpopulations, there is a single well-mixed (meta-)population, (ii) the probability of each grammar being used by any individual is given by that grammar's probability in the stationary distribution (again, as suggested by the analysis provided in Griffiths *et al.* 2008, §5*a*) and (iii) natural selection favours learners who arrive at the same grammar as any randomly selected peer. This *global* fitness function evaluates learners on their ability to communicate with any randomly selected individual from a linguistically heterogeneous metapopulation, as opposed to rewarding communication within a local subpopulation.

Under this alternative global fitness regime, the same pattern of results emerges: stronger priors are favoured in sampling populations; prior strength is selectively neutral in maximizing populations; and maximizing is favoured over sampling. There are three relatively minor qualitative differences.

Table 4. Some results for the alternative global evaluation of fitness. (Relative communicative accuracy of various strengths of prior in sampling populations. Unbiased priors are no longer evolutionarily stable.)

		majority		
		u	*bs*	*bw*
minority	*u*	—	0.79	0.9997
	bs	1	—	1.008
	bw	1	0.79	—

Table 5. Some results for the alternative global evaluation of fitness. (Relative communicative accuracy of sampling versus maximizing for flat priors. Under these circumstances, maximizing is no longer preferred over sampling.)

		majority	
		sampling	maximizing
minority	sampling	—	1
	maximizing	1	—

First, in sampling populations, the flat prior no longer constitutes an ESS: the sole ESS is the strongest possible prior, $\alpha = 1$. This is shown in table 4 (cf. table 2*a*). Populations converged on a flat prior are invasible by drift by learners with a non-flat prior: in a well-mixed unbiased population, any randomly selected unbiased individual will be equally likely to have each grammar; therefore, no matter what grammar a biased learner converges on, they will successfully communicate with an identical proportion of the population. Note the difference with the local population case, where biased learners are penalized whenever they are born into a subpopulation using a language disfavoured by the biased learner's prior.

Second, and following the same reasoning, in sampling populations, the zone of selection where much stronger priors than the population norm are selected against disappears (figure 2*c*, cf. figure 2*b*)—again, global mixing removes the risk associated with strong priors of converging on the wrong language within a homogeneous subpopulation.

Finally, and again by the same reasoning, in populations with a flat prior bias, sampling and maximizing are equivalent: under the global evaluation, the selective advantage to maximizing disappears in such populations (but not in biased populations, where maximizing is still preferred). This is illustrated in table 5 (cf. table 3, subtable for flat priors). Given the completely uniform distribution over languages and global mixing, samplers (who would normally be disproportionately penalized in situations where they fail to converge on their subpopulation's language) will successfully communicate with an equal proportion of the population regardless of which grammar they (mis)learn.

Notwithstanding these minor differences, the results of the evolutionary analysis seem to be fairly robust under different conceptions of how communicative success within a population should be defined. It is an open question whether the same conclusion pertains given more complex fitness functions—for example, one that favours the acquisition of a specific language rather than the (locally or globally) most frequent, or favours the acquisition of a frequent but not too frequent language.

To summarize: this model therefore suggests that selection for communication acting on the language faculty should select learners who maximize over the posterior distribution rather than sampling from it. Given this initial selective choice, selection should consequently be neutral with respect to the strength of the prior preference in favour of particular linguistic structures: strongly constraining language faculties are no more likely than extremely weak biases. This is the first main conclusion we would like to draw from this model: the prediction that selection will favour less flexible rather than more flexible learning is not always warranted, because learning and cultural evolution can shield the language faculty from selection.

Furthermore, there are conditions under which selection will favour the *weakest possible* prior biases. Cost can be integrated into the model by assuming that natural selection favours learners who arrive at the same grammar as their peers *and* minimize cost as a function of α: for example, individuals might reproduce proportionally to their weighted communicative accuracy, where the weighted communicative accuracy between two individuals A and B is given by

$$\mathrm{ca}'(A,B) = \frac{\mathrm{ca}(A,B)}{c(\alpha_A) + c(\alpha_B)}, \tag{4.9}$$

where $\mathrm{ca}(A, B)$ is as given in equation (4.8); $c(\alpha)$ is a cost function, with higher values corresponding to greater cost; and α_A gives the strength of prior of individual A. If we assume that strong prior biases have some cost (i.e. $c(\alpha)$ is inversely proportional to α: perhaps stronger priors require additional, restrictive but costly cognitive machinery), there are conditions under which only weak bias would be evolutionarily stable. There will be some high value of α, which we will call α^*, for which: (i) the prior is sufficiently weak that its costs relative to the unbiased strategy are low enough to allow α^* individuals to invade unbiased populations and (ii) the prior remains sufficiently strong that α^* populations resist invasion by unbiased individuals. Under such a scenario, the extremely weak α^* prior becomes the sole ESS: evolution will favour maximization and the weakest possible (but not flat) prior.

The evolutionary argument sketched above only applies if we assume that the only selective advantage to a particular prior bias arises from its communicative function: were particular priors to offer some non-linguistic advantage, less prone to being shielded by learning and cultural evolution, then it could be selected for (or against) based on those more selectively obvious functions. This means that language-learning strategies that are strongly constraining but domain general are more likely to evolve than constraining domain-specific strategies. In other words, the traditional transmission of language means that the most likely strongly constraining language faculty will not be a language faculty in the strict sense at all, but a more general cognitive faculty applied to the acquisition of language.

(e) *Conclusions on the evolution of the language faculty*

Cultural transmission alters our expectations about how the language faculty should evolve. Different language faculties may lead to identical outcomes for cultural evolution (as shown by Kirby *et al.* 2007), resulting in selective neutrality over those language faculties. If we believe that stronger prior biases (more restrictive innate machinery) cost, then selection can favour the *weakest possible* prior bias in favour of a particular type of linguistic structure. This result, in conjunction with those provided in Smith (2004), leads us to suspect that a domain-specific language faculty, selected for its linguistic consequences, is unlikely to be strongly constraining—such a language faculty is the least likely to evolve. Any strong constraints from the language faculty are likely to be domain general, not domain specific, and owe their strength to selection for alinguistic, non-cultural functions.

5. CONCLUSIONS

How do the predictions of this evolutionary modelling relate back to the uniqueness question posed in §2: what is it about humans that gives us our unusual communication system? Based on the argument outlined in §4, we think that a strongly constraining domain-specific language faculty is unlikely to have evolved in humans. It strikes us as more likely that humans have a collection of domain-general biases that they bring to the language-learning task (see also Christiansen & Chater in press)—while these biases might be rather weak, their cumulative application during language learning will lead to strong universal patterns of linguistic structure, such as the design features identified in §2. This prediction is consonant with a recent body of work in developmental linguistics that seeks to identify how domain-general statistical learning techniques can be used to acquire language from data (e.g. Saffran *et al.* 1996; Gómez 2002), what the biases of those statistical processes are (e.g. Newport & Aslin 2004; Endress *et al.* 2005; Hudson Kam & Newport 2005) and whether non-linguistic species share the same capacities and biases (Hauser *et al.* 2001; Fitch & Hauser 2004; Newport *et al.* 2004; Gentner *et al.* 2006). The comparative aspect of this work strikes us as particularly important, in that it sheds light on the extent to which humans actually have *any* specializations for the acquisition of language, rather than specializations for sequence learning or for social learning more generally.

We would not be surprised if species-specific specializations for the acquisition of linguistic structure turn out to be rare or even non-existent—this is what the evolutionary argument in §4 suggests. Rather, we suspect that the uniqueness of language arises from the co-occurrence of a number of relatively unusual cognitive capacities that constitute preconditions for the cultural evolution of these linguistic features.

The modelling literature on the cultural evolution of communication is a useful tool in identifying what these preconditions might look like. The components required for cultural evolution to produce a simple, traditionally transmitted, semantic and productive system seem to be fairly minimal:

(i) ability to modify own produced signal forms based on observed usage;
(ii) ability to learn to associate meanings with signals;
(iii) ability to infer communicative intentions (meanings) in others;
(iv) these meanings are drawn from a reasonably large and structured meaning space.

At a first approximation, cultural evolution in a population of social learners meeting these four preconditions should yield a productive and semantic communication system. The first three conditions are simply the component parts required for a traditionally transmitted semantic communication system. The fourth stipulation relates to the requirement for a learning bottleneck (a bottleneck is more likely if the system contains lots of meanings to be communicated), but also the possibility of generalizing from meaning to meaning, which requires some similarity structure among meanings—when this similarity structure is reduced, the learnability advantage of compositional language is reduced (Brighton 2002; Smith *et al.* 2003).

As discussed in §2, some of these abilities are present in non-human species. The first is present in all vocal learners, and also in non-human primates capable of sustaining ritualized gestural communication systems. While the second is not present to any meaningful extent in songbirds and other vocal learners (at least with reference to their vocally learned communication systems), it is again present in gesturally communicating apes. However, the third ability seems to be rare, including among other primates. The ability to infer the communicative intentions of others underpins the learning of large sets of arbitrary meaning–form associations: in order to learn such associations, you must be able, somewhat reliably, to infer the meaning associated with each utterance you encounter. While some other primates are able to learn to infer the meaning of a communicative signal, the laborious process of ontogenetic ritualization by which meaning for signals become established stands in stark contrast to the human ability to rapidly and accurately infer communicative intentions (see, e.g. Bloom 1997, 2000). There is no good evidence that apes are able to acquire communicative signs by more streamlined observational processes (and see Tomasello *et al.* 1997, for a negative result).

This difference in efficiency in inferring the meaning associated with a signal may be important for the following reasons. First, the requirement for a shared history of interaction reduces the potential spread of any signal, limiting it to the individuals able to devote time to its establishment. Second, it restricts the ultimate size of the repertoires of signals that can be learned, which directly links to the fourth requirement listed above, for a large and structured meaning space to associate signals with. The (limited) success of apes trained on artificial communication systems (see, e.g. Savage-Rumbaugh *et al.* 1986), with modest inventories of communicative tokens and meaningful combinations of those tokens, suggests that there may

be no fundamental cognitive barrier to the acquisition of a small productive communication system in apes, if only sufficient scaffolding is provided for its acquisition. Such communication systems may not exist in the wild (and we offer this as no more than a tentative suggestion) due to the limits on repertoire size imposed by the ritualization process, which prevents the establishment of large and systematically related sets of conventional meaning–signal pairings and therefore removes any cultural evolutionary pressure for productive structure.

Of course the question then becomes: why are these preconditions for the cultural evolution of language only met in humans? Why have humans evolved the unusual abilities to vocally learn, infer meaning efficiently, and so on? We have at present no answer to these questions, but we suspect that exploring the evolution of the preconditions for the cultural evolution of language is likely to be more profitable, and more amenable to the comparative approach, than seeking to establish the evolutionary history of a strongly constraining, domain-specific and species-unique faculty of language.

K.S. was funded by a British Academy Postdoctoral Research Fellowship. The development of the model described in §4 was facilitated by our attendance at the NWO-funded Masterclass on Language Evolution, organized by Bart de Boer and Paul Vogt.

ENDNOTES

[1]Or some other discriminable unit, e.g. handshape, orientation and location in sign language (Stokoe 1960).

[2]See Kirby *et al.* (2008) for an experimental treatment of the same phenomenon.

[3]We will only consider the case where α takes integer values.

[4]Provided the Markov process described by the Q matrix is ergodic (see Griffiths & Kalish 2007, p. 446).

[5]This assumes that identical grammars are required for communication. We could instead evaluate communicative accuracy based on the degree of similarity between grammars or the degree of similarity between datasets produced by those grammars. These more graded notions of communication produce qualitatively similar results to those presented here, the quantitative change being a reduction in the magnitude of the selection differentials in some cases.

[6]Again, more graded notions of match between grammars produce qualitatively similar results.

REFERENCES

Andersen, H. 1973 Abductive and deductive change. *Language* **40**, 765–793. (doi:10.2307/412063)

Baldwin, J. M. 1896 A new factor in evolution. *Am. Nat.* **30**, 441–451. (doi:10.1086/276408)

Bloom, P. 1997 Intentionality and word learning. *Trends Cogn. Sci.* **1**, 9–12. (doi:10.1016/S1364-6613(97)01006-1)

Bloom, P. 2000 *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Brighton, H. 2002 Compositional syntax from cultural transmission. *Artif. Life* **8**, 25–54. (doi:10.1162/106454602753694756)

Brighton, H. 2003 Simplicity as a driving force in linguistic evolution, PhD thesis, The University of Edinburgh.

Brighton, H., Kirby, S. & Smith, K. 2005 Cultural selection for learnability: three principles underlying the view that language adapts to be learnable. In *Language origins: perspectives on evolution* (ed. M. Tallerman), pp. 291–309. Oxford, UK: Oxford University Press.

Briscoe, E. 2000 Grammatical acquisition: inductive bias and coevolution of language and the language acquisition device. *Language* **76**, 245–296. (doi:10.2307/417657)

Briscoe, E. 2003 Grammatical assimilation. In *Language evolution* (eds M. H. Christiansen & S. Kirby), pp. 295–316. Oxford, UK: Oxford University Press.

Call, J. & Tomasello, M. (eds) 2004 *The gestural communication of apes and monkeys*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Catchpole, C. K. & Slater, P. J. B. 1995 *Bird song: biological themes and variations*. Cambridge, UK: Cambridge University Press.

Chater, N. & Vitányi, P. M. B. 2003 Simplicity: a unifying principle in cognitive science? *Trends Cogn. Sci.* **7**, 19–22. (doi:10.1016/S1364-6613(02)00005-0)

Cheney, D. & Seyfarth, R. 1990 *How monkeys see the world: inside the mind of another species*. Chicago, IL: University of Chicago Press.

Chomsky, N. 1980 *Rules and representations*. London, UK: Basil Blackwell.

Christiansen, M. & Chater, N. In press. Language as shaped by the brain. *Behav. Brain Sci*.

Deacon, T. 1997 *The symbolic species*. London, UK: Penguin.

Eens, M. 1997 Understanding the complex song of European starling: an integrated ethological approach. *Adv. Study Behav.* **26**, 355–434.

Endress, A. D., Scholl, B. J. & Mehler, J. 2005 The role of salience in the extraction of algebraic rules. *J. Exp. Psychol. Gen.* **134**, 406–419. (doi:10.1037/0096-3445.134.3.406)

Evans, C. S., Evans, L. & Marler, P. 1993 On the meaning of alarm calls: functional reference in an avian vocal system. *Anim. Behav.* **46**, 23–38. (doi:10.1006/anbe.1993.1158)

Fitch, W. T. 2000 The evolution of speech: a comparative review. *Trends Cogn. Sci.* **4**, 258–267. (doi:10.1016/S1364-6613(00)01494-7)

Fitch, W. T. & Hauser, M. D. 2004 Computational constraints on syntactic processing in a nonhuman primate. *Science* **303**, 377–380. (doi:10.1126/science.1089401)

Gentner, T. Q., Fenn, K. M., Margoliash, D. & Nusbaum, H. C. 2006 Recursive syntactic pattern learning by songbirds. *Nature* **440**, 1204–1207. (doi:10.1038/nature04675)

Gil, D. & Slater, P. J. B. 2000 Song organisation and singing patterns of the willow warbler, *Phylloscopus trochilus*. *Behaviour* **137**, 759–782. (doi:10.1163/156853900502330)

Gómez, R. L. 2002 Variability and detection of invariant structure. *Psychol. Sci.* **13**, 431–436. (doi:10.1111/1467-9280.00476)

Griffiths, T. L. & Kalish, M. L. 2005 A Bayesian view of language evolution by iterated learning. In *Proc. 27th Annual Conf. of the Cognitive Science Society* (eds B. G. Bara, L. Barsalou & M. Bucciarelli), pp. 827–832. Mahwah, NJ: Erlbaum.

Griffiths, T. L. & Kalish, M. L. 2007 Language evolution by iterated learning with Bayesian agents. *Cogn. Sci.* **31**, 441–480.

Griffiths, T. L., Kalish, M. L. & Lewandowsky, S. 2008 Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Phil. Trans. R. Soc. B* **363**, 3591–3603. (doi:10.1098/rstb.2008.0146)

Hauser, M. D., Newport, E. L. & Aslin, R. N. 2001 Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* **78**, B53–B64. (doi:10.1016/S0010-0277(00)00132-3)

Hauser, M. D., Chomsky, N. & Fitch, W. T. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579. (doi:10.1126/science.298.5598.1569)

Hockett, C. F. 1960 The origin of speech. *Sci. Am.* **203**, 88–96.

Hudson Kam, C. L. & Newport, E. L. 2005 Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Lang. Learn. Dev.* **1**, 151–195. (doi:10.1207/s15473341lld0102_3)

Hurford, J. R. 1990 Nativist and functional explanations in language acquisition. In *Logical issues in language acquisition* (ed. I. M. Roca), pp. 85–136. Dordrecht, The Netherlands: Foris.

Jackendoff, R. 2002 *Foundations of language: brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.

Kinsella, A. R. In press. *Language evolution and syntactic theory*. Cambridge, UK: Cambridge University Press.

Kirby, S. 1999 *Function, selection and innateness: the emergence of language universals*. Oxford, UK: Oxford University Press.

Kirby, S. 2002*a* Learning, bottlenecks and the evolution of recursive syntax. In *Linguistic evolution through language acquisition: formal and computational models* (ed. E. Briscoe), pp. 173–203. Cambridge, UK: Cambridge University Press.

Kirby, S. 2002*b* Natural language from artificial life. *Artif. Life* **8**, 185–215. (doi:10.1162/106454602320184248)

Kirby, S., Dowman, M. & Griffiths, T. L. 2007 Innateness and culture in the evolution of language. *Proc. Natl Acad. Sci. USA* **104**, 5241–5245. (doi:10.1073/pnas.0608222104)

Kirby, S., Cornish, H. & Smith, K. 2008 Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl Acad. Sci. USA* **105**, 10 681–10 686. (doi:10.1073/pnas.0707835105)

Krifka, M. 2001 Compositionality. In *The MIT encyclopaedia of the cognitive sciences* (eds R. A. Wilson & F. Keil), pp. 152–153. Cambridge, MA: MIT Press.

Marler, P. 1955 Characteristics of some animal calls. *Nature* **176**, 6–8. (doi:10.1038/176006a0)

Maynard Smith, J. & Price, G. R. 1973 The logic of animal conflict. *Nature* **146**, 15–18. (doi:10.1038/246015a0)

Maynard Smith, J. & Szathmáry, E. 1995 *The major transitions in evolution*. Oxford, UK: Oxford University Press.

Newport, E. L. & Aslin, R. N. 2004 Learning at a distance I: statistical learning of non-adjacent dependencies. *Cogn. Psychol.* **48**, 127–162. (doi:10.1016/S0010-0285(03)00128-2)

Newport, E. L., Hauser, M. D., Spaepan, G. & Aslin, R. N. 2004 Learning at a distance II: statistical learning of non-adjacent dependencies in a non-human primate. *Cogn. Psychol.* **49**, 85–117. (doi:10.1016/j.cogpsych.2003.12.002)

Nowak, M. A. & Krakauer, D. C. 1999 The evolution of language. *Proc. Natl Acad. Sci. USA* **96**, 8028–8033. (doi:10.1073/pnas.96.14.8028)

Nowak, M. A., Krakauer, D. C. & Dress, A. 1999 An error limit for the evolution of language. *Proc. R. Soc. B* **266**, 2131–2136. (doi:10.1098/rspb.1999.0898)

Nowak, M. A., Plotkin, J. B. & Jansen, V. A. A. 2000 The evolution of syntactic communication. *Nature* **404**, 495–498. (doi:10.1038/35006635)

Nowak, M. A., Komarova, N. L. & Niyogi, P. 2001 Evolution of universal grammar. *Science* **291**, 114–117. (doi:10.1126/science.291.5501.114)

Okanoya, K. 2004 The bengalese finch: a window on the behavioral neurobiology of birdsong syntax. *Ann. N. Y. Acad. Sci.* **1016**, 724–735. (doi:10.1196/annals.1298.026)

Oudeyer, P.-Y. 2005 The self-organization of speech sounds. *J. Theor. Biol.* **233**, 435–449. (doi:10.1016/j.jtbi.2004.10.025)

Pinker, S. & Bloom, P. 1990 Natural language and natural selection. *Behav. Brain Sci.* **13**, 707–784.

Pullum, G. K. & Scholz, B. C. 2002 Empirical assessment of stimulus poverty arguments. *Linguist. Rev.* **19**, 9–50. (doi:10.1515/tlir.19.1-2.9)

Saffran, J. R., Aslin, R. N. & Newport, E. L. 1996 Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928. (doi:10.1126/science.274.5294.1926)

Savage-Rumbaugh, S., McDonald, K., Sevcik, R. A., Hopkins, W. D. & Rubert, E. 1986 Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *J. Exp. Psychol. Gen.* **115**, 211–235. (doi:10.1037/0096-3445.115.3.211)

Slater, P. J. B. 2005 Animal communication: vocal learning. In *The encyclopedia of language and linguistics* (ed. K. Brown), pp. 291–294, 2nd edn. Amsterdam, The Netherlands: Elsevier

Smith, K. 2004 The evolution of vocabulary. *J. Theor. Biol.* **228**, 127–142. (doi:10.1016/j.jtbi.2003.12.016)

Smith, K., Brighton, H. & Kirby, S. 2003 Complex systems in language evolution: the cultural emergence of compositional structure. *Adv. Complex Syst.* **6**, 537–558. (doi:10.1142/S0219525903001055)

Stokoe, W. C. 1960 *Sign language structure*. Silver Spring, MD: Linstok Press.

Thomas, B. 1985 On evolutionarily stable sets. *J. Math. Biol.* **22**, 105–115. (doi:10.1007/BF00276549)

Tomasello, M. 1996 Do apes ape? In *Social learning in animals: the roots of culture* (eds C. Heyes & B. Galef), pp. 319–436. San Diego, CA: Academic Press.

Tomasello, M., Call, J., Warren, J., Frost, G., Carpenter, M. & Nagell, K. 1997 The ontogeny of chimpanzee gestural signals: a comparison across groups and generations. *Evol. Commun.* **1**, 223–259.

Waddington, C. H. 1975 *The evolution of an evolutionist*. Ithaca, NY: Cornell University Press.

Whiten, A. 2005 The second inheritance system of chimpanzees and humans. *Nature* **437**, 52–55. (doi:10.1038/nature04023)

Wray, A. 1998 Protolanguage as a holistic system for social interaction. *Lang. Commun.* **18**, 47–67. (doi:10.1016/S0271-5309(97)00033-5)

Zuberbuhler, K. 2001 Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behav. Ecol. Sociobiol.* **50**, 414–422. (doi:10.1007/s002650100383)

Zuberbuhler, K. 2002 A syntactic rule in forest monkey communication. *Anim. Behav.* **63**, 293–299. (doi:10.1006/anbe.2001.1914)

Zuidema, W. H. 2003 How the poverty of the stimulus solves the poverty of the stimulus. In *Advances in neural information processing systems 15* (*Proceedings of NIPS '02*) (eds S. Becker, S. Thrun & K. Obermayer), pp. 51–58. Cambridge, MA: MIT Press.

Zuidema, W. H. & de Boer, B. In press. The evolution of combinatorial phonology. *J. Phon.*