15

Syntax As an Adaptation to the Learner

Simon Kirby, Morten H. Christiansen, and Nick Chater

Abstract

Please add.

Introduction

One of the most obviously striking features of human language, especially in comparison with all other communication systems in nature, is syntax. More precisely, language is unique in providing an open-ended system for relating signals and meanings, one which has its own internal structure. The particular structure of the mapping between meanings and signals varies from language to language, and for many researchers, the central challenge for linguistic theory is an explanation of the constraints on this variation. In other words, linguistics seeks an explanatory account of the universals of syntactic structure.

A hugely influential approach to this explanatory challenge has involved a direct appeal to biology. In this view, syntax arises from our species-specific biological endowment which is specific to language. We have the languages we do because an innately given "language faculty" has a particular structure that constrains the possible types of language (e.g., Hoekstra and Kooij 1988). In particular, Chomsky (1975) suggests that it is a set of innate constraints on language acquisition that determines the nature of syntactic universals.

This view directly relates universal properties of syntax on the one hand, with a universally shared biological trait on the other[1]. One issue with this

[1] Note that there is a presumption here that the language faculty is uniform across members of our species, or at least it is uniform with respect to the constraints on cross-linguistic variation. This is a reasonable assumption to make in that there is no obvious evidence that some individuals find particular types of language harder to acquire than other individuals. However, it has recently been challenged as a result of large-scale statistical analysis of genetic and linguistic variation (Dediu and Ladd 2007).

attempt at explanation (for a discussion, see, e.g., Hurford 1990) is that it simply replaces one explanatory challenge with another. Although it appears to answer the question of why we have the particular language universals we do, it immediately poses another: Why is our language faculty constrained in the way it is? In a landmark paper, Pinker and Bloom (1990) directly address this question in an attempt to support the nativist approach to explanation. They set out what might be called the orthodox evolutionary approach to language (see Figure 15.1). In this approach, our innate language faculty shapes the structure of language and is in turn shaped by biological evolution driven by natural selection for communicative function. This is motivated by the observation of the apparent adaptive nature of syntactic structure:

> Grammar is a complex mechanism tailored to the transmission of propositional structures through a serial interface... Evolutionary theory offers clear criteria for when a trait should be attributed to natural selection: complex design for some function, and the absence of alternative processes capable of explaining such complexity. Human language meets this criterion (Pinker and Bloom 1990, p. 707).

Pinker and Bloom (1990) provide an influential recasting of Chomskyan nativism in evolutionary terms, one that takes us from observed universals of syntactic structure, through an inferred innate Universal Grammar, grounded firmly in standard mechanisms of evolutionary biology. To critically assess the foundations of this view, it is worth unpacking some of the motivations for assuming this kind of evolutionary nativism. In this paper we will consider three in the light of recent research on the adaptive mechanisms underlying human language:

1. *Universals*. Languages vary, but that variation is constrained. The nativist approach provides a simple and compelling account of this: the constraints on cross-linguistic variation *directly* reflect the languages we can acquire.
2. *The appearance of design.* This is the point made in the quotation from Pinker and Bloom (1990) above. Language structure is adaptive – natural selection of innate constraints appears to be the only available explanation.
3. *Poverty of the stimulus.* For many linguists, this is the most familiar reason for assuming innate constraints. Children have access to only limited and degraded evidence that underdetermines the language they are attempting to acquire. Nevertheless children robustly converge on the correct language. Language acquisition therefore appears to be impossible without significant innate knowledge about the languages children may face.

These motivations seem well-founded and reasonable, and appear to provide solid ground on which to build an evolutionary account of the origins

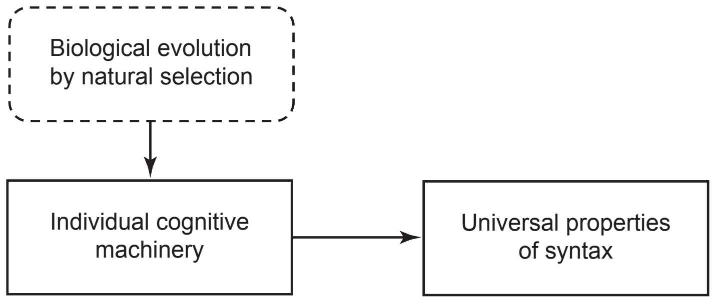

Figure 15.1 The orthodox evolutionary approach to explaining syntactic structure. Biological evolution by natural selection shapes our innate cognitive mechanisms for acquiring language which directly determine the universal properties of syntax.

of syntactic structure. However, in this paper we wish to argue that there is something missing from orthodox evolutionary approach sketched in Figure 15.1 that undermines each of these motivations. A key unstated assumption underlying Pinker and Bloom's framework (and indeed the standard nativist position more generally) is that there is a straightforward link between our innate language faculty and universal properties of language structure. This assumption seems reasonable on the face of it, but on closer inspection it is problematic. After all, these are two very different kinds of entities: a genetically determined universally shared part of our cognitive machinery; and constraints on the variation of internalized patterns of linguistic behavior shared within speech communities.

What is the mechanism that bridges the gap between an individual-level phenomenon (the structure of a language-learner's cognitive machinery) and a population-level phenomenon (the distribution of possible languages)? As Kirby et al. (2004) argue, the solution to this problem is to explicitly model the way in which individual behavior leads to population effects over time. Language emerges out of a repeated cycle of language learning and language use, and it is by studying this socio-cultural process directly that we will see how properties of the individual leave their mark on the universal structure of language (Figure 15.2).

Of course, it is not a priori obvious that the extra box in Figure 15.2 will add anything substantial to the picture – that considering the role of *cultural* as well as biological evolution will change anything. The goal of this chapter is to argue the contrary. By ignoring or downplaying the importance of cultural evolution, evolutionary linguistics risks coming to the wrong conclusions. It is crucial that researchers interested in language evolution do not make the mistake of assuming that evolution is a purely biological process that can be studied in isolation from the dynamics operating at shorter timescales. Although it would be convenient if we could say that the study of social transmission and cultural evolution is purely the realm of historical linguistics and therefore

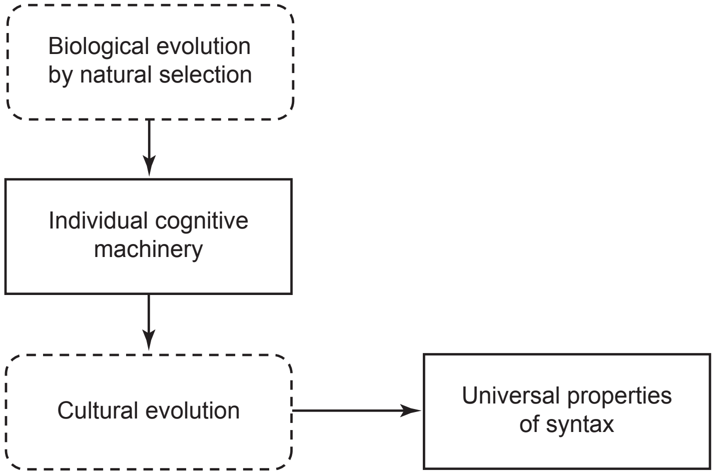

Figure 15.2 The place of cultural evolution in determining the universal properties of syntax. Biological evolution shapes our individual cognitive machinery, but this is only indirectly connected to the object of explanation. Individuals influence a process of social transmission and cultural evolution that eventually leads to emergent universals.

evolutionary linguistics can essentially ignore these mechanisms to focus purely on natural selection, this is not how evolutionary systems work.

By taking the role of cultural evolution seriously, we will show that the motivations for a *strongly-constraining domain specific linguistic nativism*[2] are undermined. Throughout we will stress the importance of taking an empirical approach to language evolution. Three complex systems are involved in the emergence of syntax: individual learning, cultural transmission, and biological evolution. We cannot reasonably expect our intuitions about the interactions of these to be sound. One response is to build models, both in the computer and in the laboratory, which allow us to explore in miniature how the processes underlying language evolution work, and then apply what we learn from the models to better understand the real object of enquiry: human language.

The Logical Problem of Language Evolution

Before turning to models of cultural evolution, we first wish to explore some of the issues underlying Motivation 2, above – that the existence of language structure implies an explanation in terms of natural selection of innate constraints. We argue that advocates of a richly structured, domain-specific, innate UG

[2] It is important to stress that our arguments in this chapter apply to a particular nativist stance: one which infers innate constraints on language that are both specific to language and map directly onto language universals. It is a common misunderstanding that this position is synonymous with generative approaches to language, but this is not necessarily the case (for extensive discussion, see Kirby 1999).

confront a "logical problem of language evolution" (Christiansen and Chater 2008). To see this, we begin by noting that, as for any other putative biological structure, an evolutionary story for UG can take one of two routes. One route is to assume that brain mechanisms specific to language acquisition have evolved over long periods of natural selection (e.g., Pinker and Bloom 1990). The other rejects the idea that UG has arisen through adaptation and proposes that UG has emerged by nonadaptationist means, (e.g., Lightfoot 2000).

The nonadaptationist account can rapidly be put aside as an explanation for a domain-specific, richly structured UG. The nonadaptationist account boils down to the idea that some process of *chance variation* leads to the creation of UG. Yet the probability of randomly building a fully functioning, and complete novel, biological system by chance is infinitesimally small (Christiansen and Chater 2008). To be sure, so-called "evo-devo" research in biology has shown how a single mutation can lead, via a cascade of genetic ramifications, to dramatic phylogenetic consequences (e.g., additional pairs of legs instead of antennae, Carroll 2001). But such mechanisms cannot explain how an intricate and functional system can arise, de novo.

What of the adaptationist account? UG is intended to characterize a set of universal grammatical principles that holds across all languages; it is a central assumption that these principles are arbitrary. This implies that many combinations of arbitrary principles will be equally adaptive—as long as speakers adopt the *same* arbitrary principles. Pinker and Bloom (1990) draw an analogy with protocols for communication between computers: it does not matter what specific settings are adopted, as long as everyone adopts the same settings. Yet the claim that a particular "protocol" can become genetically embedded through adaptation faces three fundamental difficulties (Christiansen and Chater 2008).

The first problem stems from the spatial dispersion of human, which occurred within Africa, and ultimately beyond Africa, before and during the period (100–200 K years) within which most scholars assume language emerged. Each sub-population would be expected to create highly divergent linguistic systems. But, if so, each population will develop a UG as an adaptation to a *different* linguistic environment—hence, UGs should, like other adaptations, diverge to fit their local environment. Yet modern human populations do not seem to be selectively adapted to learn languages from their own language groups. Instead, every human appears, to a first approximation, equally ready to learn any of the world's languages.

The second problem is that natural selection produces adaptations designed to fit the *specific* environment in which selection occurs, i.e., a language with a specific syntax and phonology. It is thus puzzling that an adaptation for UG would have resulted in the genetic encoding of highly abstract grammatical properties, rather than fixing the superficial properties of one specific language.

The third, and perhaps most fundamental, problem is that linguistic conventions change much more rapidly than genes do, thus creating a "moving target" for natural selection. Computational simulations have shown that even under conditions of relatively slow linguistic change, arbitrary principles do not become genetically fixed. Chater et al. (2009) illustrate this problem in a series of computer simulations. They model the specific evolutionary mechanism to which Pinker and Bloom appeal to explain the evolution of innate knowledge of language. This mechanism is the *Baldwin effect*: that information which is initially acquired during development can become gradually encoded in the genome (see also review and discussion in Briscoe 2003, this volume, and Deacon 1997). Chater et al. assume the simplest possible set-up: that (binary) linguistic principles and language "genes" stand in one-to-one correspondence. Each gene has three alleles—a neutral allele, and two alleles, each encoding a bias for a version of the linguistic principle. Agents learn the language by trial-and-error, where their guesses are biased according to which alleles they have[3]. The fittest agents (i.e., the fastest learners) are allowed to reproduce, and a new generation of agents is produced by sexual recombination and mutation. When the language is fixed, there is a selection pressure in favor of the "correctly" biased genes, and these rapidly dominate the population (Figure 15.3).

However, when language is allowed to change gradually (e.g., due to grammaticalization-like processes or exogenous forces such as language contact), the effect reverses—biased genes are severely selected against when they are inconsistent with the linguistic environment, and neutral genes come to dominate the population. The selection in favor of neutral genes occurs even for low levels of language change (i.e., the effect occurs, to some degree, even if language change equals the rate of genetic mutation). But, of course, linguistic change (prior to any genetic encoding) is likely to have been much faster than genetic change[4].

It remains possible, though, that the origin of language did have a substantial impact on human genetic evolution. The above arguments only preclude biological adaptations for *arbitrary* features of language. There might have been features that are universally stable across linguistic environments that might lead to biological adaptation (such as the means of producing speech, the need for enhanced memory capacity, or complex pragmatic inferences – see Kirby and Hurford 1997, and Christiansen et al. 2006, for computational models that look at nonarbitrary adaptation). In addition, the situation becomes

[3] Because the learning biases are probabilistic, learners are always able to learn the language eventually, even if their genetic biases are in the wrong direction (which will make them slow learners).

[4] It may be tempting to object that UG principles do not change, and hence provide a stable environment over which adaptation can operate. However, such an objection would be circular because it presupposes what an evolutionary theory of UG is meant to explain. That is, an innate UG is supposed to explain language universals and thus it cannot be assumed that language universals predate the emergence of UG.

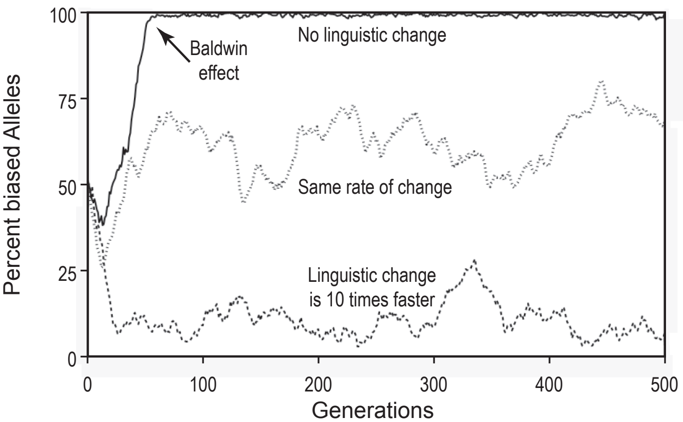

Figure 15.3 The effect of linguistic change on the genetic encoding of arbitrary linguistic principles. The simulation has a population of 100 agents, a genome size of 20, survival of the top 50% of the population, and starts with 50% neutral alleles. With no linguistic change, a Baldwin effect occurs—i.e., alleles encoding specific aspects of language emerge rapidly. But when the language change, biased alleles are no longer advantageous and are selected against. The results are typical of those obtained using a wide range of parameters (Chater et al. 2009).

more complex when we look in more detail at interactions between cultural evolution and biological evolution of *weak* constraints. We will return to this issue later.

Language and Cultural Evolution

The problem with the straightforward application of the arguments from biological adaptation to theories of UG lie principally in our poor understanding of exactly how the process of cultural evolution works for language. Specifically, we need to move towards a general theory of how particular kinds of UG constraints or biases lead to language structure when mediated by a process of cultural transmission (Figure 15.2). Only once we have this can we hope to disentangle the precise roles of the different adaptive processes involved.

The *Iterated Learning Model* (e.g., Kirby et al. 2004; Brighton et al. 2005) aims to provide a general solution to this problem. The idea is simple: to build idealized models of the process of cultural transmission that show how global effects emerge from the repeated process of individuals learning and producing linguistic behavior. The simplest iterated learning models consist of a chain of *agents* (individuals modeled in simulation, or in an experimental setting) each of which observes the linguistic behavior of the previous agent in the chain, attempts to learn the underlying linguistic system, and then goes on to produce observable behavior for the next agent down the chain. Like the parlor game, *Telephone*, this produces a cultural dynamic whereby the behavior produced by agents may change over time purely by virtue of being passed-on by an iterated cycle of learning and production. In general, we define iterated learning to be a cultural process whereby an individual learns a behavior by observing another individual's behavior, *who acquired it in the same way*.[5]

This general approach—modeling the way in which linguistic behavior is repeatedly transmitted between individuals—has been studied extensively in the literature, using everything from dramatically idealized simple models (e.g., Kirby et al. 2007) to extremely sophisticated models involving realistic populations of agents interacting socially and grounded in a real environment (e.g., Steels 2003). A thorough review of the result of this modeling work is well beyond the scope of this article, but one of the recurrent observations relates to the importance of what have been called *transmission bottlenecks*. Specifically, if a learner is given imperfect information about the language they are trying to acquire (i.e., where there is some kind of bottleneck on the transmission of language from one individual to another, be it in terms of noise, processing constraints, or simply not hearing all the relevant data) then cultural transmission becomes an *adaptive system*. What this means is that language will adapt so that it appears to be designed to fit through whatever bottleneck the experimenter imposes.

A classic example of this kind of result is provided by several studies into the emergence of compositional syntax (for a review see, Brighton et al. 2005). The existence of compositional structure in the mapping between meanings and strings is an apparently adaptive feature of human language syntax—it is a crucial part of what enables us to have open-ended expressivity, an assuredly adaptive trait[6]. However, computational models of iterated learning which start from random noncompositional initial languages, or with no language at all, show that this property emerges from the repeated cycle of production and learning without any biological evolution of the agents. The reason is straightforward: compositional structure improves the stability of languages transmitted through a bottleneck. To put it another way, compositionality is an adaptation by language to improve its own survival. There is nothing mysterious or

5 Note that this does not limit iterated learning to purely vertical transmission. Indeed, one of the earliest models of this process (Batali 1998) employed purely horizontal transmission (i.e. with individuals learning, producing and then learning again in a completely mutually interacting population). Batali's results bear striking similarities to the quite different models with only vertical transmission. It is the similarity of results across a range of models that has lead researchers to attempt to understand the dynamics of iterated learning in as general terms as possible (e.g., Griffiths and Kalish 2007).

6 Of course, compositionality is only one aspect of the uniquely human structure of syntax (and a very basic one at that). A language with just compositionality and none of the other features of human syntax might arguably be described as "protolanguage," so more work is needed to see if similar processes as those described here can take us further. Nevertheless, it is important to understand that these results do have significant implications as they stand for the particular kind of nativism we are discussing. Furthermore, the Bayesian model we turn to later is completely general and not reliant on a particular view of what constitutes syntactic structure.

teleological about this. Rather, it is the inevitable consequence of the process of cultural transmission. As Hurford (2000) puts it succinctly, "social transmission favors linguistic generalization."

To check the generality of the conclusions from the computational models, we developed an experimental framework for iterated learning (Kirby et al. 2008). In our experiments, human participants were faced with an artificial language learning task in which they were required to learn to associate strings of written syllables with pictures of colored moving shapes. Each picture was either a square, triangle or circle, was colored either red, blue or black, and was depicted as bouncing, spiraling, or moving horizontally. Although, in the testing phase, participants were asked to produce strings for all 27 different possible pictures, they were only actually trained on a random subset of 14 of these.

The crucial aspect of these experiments that makes them relevant here is that the language a participant is trained on is actually a random sample of the output of the previous participant in the experiment at test, with the very first participant being trained on a randomly constructed language (i.e., one which exhibits no compositional structure). With this experimental set up we are able to observe in the laboratory exactly how a simple language like this one evolves culturally. Two questions present themselves: Will languages adapt to be increasingly learnable? Will structure emerge?

The answer turns out to be "yes" to both questions, but the exact kind of structure that emerges depends in an interesting way on the nature of the bottleneck. Figure 15.4 shows quantitative results for the experiment outlined above (with the lines marked "unfiltered"). Clearly, the languages become more learnable and more structured over time, purely as a result of being transmitted repeatedly from individual to individual. We start with a language that is impossible to learn in the sense that there is no way of accurately guessing what an unseen meaning might be called, and end with a language where participants do extremely well in generalizing accurately to unseen examples.

What does the emergent structure that makes this possible look like? It turns out that, in this version of the experiment, what emerges is a kind of structured lexical underspecification. The number of distinct strings in the language plummets from 27 at the start to a handful after 10 "generations" (the exact number varies from replication to replication). These remaining strings thus refer to a set of meanings, rather than a single one. What is fascinating is that these sets show distinctive structure (picked up by the quantitative measure in Figure 15.4). For example, in one run of the experiment, a single word emerged to refer to all the horizontally moving objects. This kind of nonrandom underspecification of meanings in the language allows learners to generalize accurately to unseen meanings.

Why weren't we seeing the kind of compositional structure that was apparent in the computational models? One difference is that the simulations typically built-in some motivation for the agents to maintain expressivity and avoid

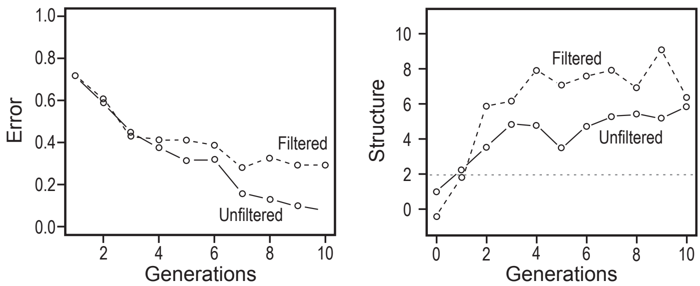

Figure 15.4 The average of four cultural transmission chains in each of two experimental conditions showing languages evolve to be more learnable and more structured over time. Each generation here is an individual experimental participant who learns an artificial language produced by the previous participant in the experiment. The graph on the left shows the average error in learning the language (a score of 1 means that the strings produced were completely dissimilar to the target, a score of 0 means they were exactly correct). The graph on the right shows a measure of structure in the mapping between strings and meanings (for more information, see Kirby et al. 2008). Wherever this structure measure is above the dotted line, the language is nonrandom at the 95% confidence interval. The "unfiltered" condition indicates that the training data is passed directly to the next participant, whereas in the "filtered" condition ambiguous strings are removed (see text).

collapsing all meaning distinctions to a single string, for example. We wanted to show that a similar result could be achieved in the human experiment just by a minimal change to the transmission bottleneck. Accordingly, we reran the experiments with a single alteration: before giving the training data to a participant, we scanned it for any underspecification. If the same string was used for more than one meaning, we simply filtered all but one of those instances out of the training data. This filtering step corresponds to the pressure in real language use to maintain expressivity, such that distinct meanings tend to be assigned distinct signals.[7] Note that participants were not aware we were doing this (indeed, in neither version of the experiment did any participant ever guess that the experiment involved cultural transmission in any case). However, the difference in outcome was dramatic. Figure 15.4 shows that the quantitative results showed the same trend, albeit revealing this was a more difficult task. The big difference was in the particular structure of the language.

7 Filtering underspecification from input is not necessarily a particularly realistic way of achieving this, although it is likely that something like filtering based on communicative utility is a real mechanism in language transmission. In the experimental model it should be seen as a stand-in for a more complex suite of communicatively motivated pressures. An alternative might have been to set up the experiment within the frame of an overtly communicative task. However, it was crucial for our purposes to demonstrate that participants were not intentionally and intelligently *designing* a communication system (for example on analogy with their own language), as they may well have done if this became the overt goal of the experimental setup. Rather, we wanted to show that the cultural transmission process alone is all that is required to generate adaptive structure.

With filtering in place, the kind of structured underspecification we saw previously no longer could take hold. Nevertheless, a different adaptation emerged which lead to both an increase in learnability *and* the maintenance of expressivity. This adaptation was exactly the same as the one that emerged in the computational models: compositionality. The examples below (from data presented in Kirby et al. 2008) show how three "morphemes" emerged in one chain encoding color shape and motion respectively (note the hyphens are included for clarity only, they are not present in the participants' output or input):

(1) (a) n-eke-ki
"black-triangle-horizontal"
(b) n-eki-pilu
"black-triangle-spiral"
(c) l-aho-ki
"blue-circle-horizontal"
(d) l-aho-plo
"blue-circle-bounce"
(e) r-e-plo
"red-square-bounce"
(f) r-e-pilu
"red-square-spiral"

This result was not invented by one particularly smart individual in the experiment, but rather appeared cumulatively and without deliberate design on behalf of the participants. The participants were not trying to construct a perfect language to fit through the bottleneck (which would have been impossible given that they could not know the constraints we were placing on the bottleneck). They were simply trying their best to give us back what we gave them. Many participants did not even realize that we were asking them to generalize to unseen meanings. Nevertheless the language underwent cumulative cultural adaptation, just as predicted by the computational models. Clearly, these are adult participants that already have a native language and as such we need to be aware that the biases they bring to bear on the learning task are a combination of biologically basic ones and those that arise from their existing specific cultural inheritance (i.e., their native language). Of course, the close fit of the experimental results with those predicted by the simulation models speaks against the idea that acquired biases are the primary driver. More importantly, however, the primary purpose of these models is not as a discovery procedure for our biological biases but rather as a way of determining how a culturally transmitted language responds to whatever biases and transmission pressures are placed upon it.

This result, and others like it, cast doubt on all the motivations for strongly constraining domain-specific innateness listed in the introduction. Firstly, and most importantly, it demonstrates that there is more than one mechanism capable of delivering the appearance of design. Natural selection (in the biological sense) is no longer the only possible explanation for adaptive structure in language—the mere fact that language is transmitted culturally induces adaptation by language itself. Secondly, it recasts the so-called "poverty of the stimulus" problem in a new light. As others have argued (e.g., Zuidema 2003) these results show language structure does not exist in spite of the impoverished stimulus available to the child, but rather precisely because of the stimulus poverty. When more, and better, data is provided to learners in the models the languages that emerge exhibit less structure.[8]

What of the remaining motivation: universals? In some sense, the experiments with compositionality already address this issue. If we were to look only at the end result of the simulations, compositional language, we might be lead to the wrong conclusion that the learners were equipped with a mechanism that constrained them only to learn compositional languages. However, this would be a mistake. In these models, even when learners do not reliably acquire compositional structure an exceptionless universal outcome can still be expected (Hurford 2000).

To understand the relationship between universals and UG better, we implemented a mathematical model of iterated learning using Bayesian agents (Kirby et al. 2007). This allows us to control very precisely the contribution of innateness and see what language universals emerges for a given transmission bottleneck. The innate contribution is represented in the model in terms of a prior bias over possible languages. That is, we are able to provide a probability distribution over languages that reflects the innate preference for one language over another. These innate biases can therefore be arbitrarily strong or weak, covering the spectrum of possibilities from hard constraints to slight tendencies.

By treating iterated learning as a Markov process in which the transition between languages is determined by the Bayesian model of learning, we are able to predict exactly what universals should emerge for a given model of innateness. The most striking result from this work is that, given reasonable assumptions about how learners select hypotheses, innate biases are *not* reflected directly in language universals. Specifically, the strength of the language universals that emerge is independent of the strength of the innate bias and is instead determined by the nature of the transmission bottleneck. Simplifying somewhat, in conditions of data poverty, arbitrarily weak innate predispositions are amplified by cultural transmission. Indeed, the strength of innate bias makes absolutely no difference to the final distribution of language types in the model.

What this means is that we cannot infer strongly constraining innateness simply by looking at language universals. This observation actually has some empirical support. For example, there are cases where culturally transmitted birdsong for a particular species has an exceptionless universal, but where a bird of that species can nevertheless acquire an atypical song (e.g., Hultsch 1991). Similarly, Dediu and Ladd (2007) present evidence that there is genetic

8 Of course, this in itself does not provide a solution to learnability arguments, but note that iterated learning ensures that the training data provided for a learner will be the best possible data for the particular learning problem learners typically face (because cultural evolution will tend to maximize the learnability of the language).

variation in prior disposition to acquire tone languages which result in a clear and strong skewing of language types in different populations. Nevertheless, it is clear that any normal individual can acquire any existing language whatever their genetic makeup. Both these cases suggest that whatever biases lead to the population-level effects, their effect at an individual level can be tiny.

Finally, this result has implications for the biological evolution of innateness discussed in the previous section. Smith and Kirby (2008) look at the coevolution of Bayesian learners and the languages they transmit culturally. They argue that cultural transmission shields bias strength from the view of natural selection (see also Deacon 2003a), leading to the possibility that strong biases may be impossible to maintain against mutation pressure.

So, where does this leave the biological evolution of innate bias? The results of the models discussed in the previous two sections do not rule out the evolution of innate bias, but they narrow down the possible ways this evolution could take place. Two particularly plausible alternatives remain: either that bias is not domain-specific (and therefore could be subject to selection pressures not solely determined by the emerging cultural system); or it could be a bias weak enough not to have a strong impact on a single individual, but nevertheless be amplified by cultural transmission.

Biases that Shape Syntax

We have proposed that language has adapted to biases or constraints deriving from language learners and users: biases which may not be specific to language. But how far can these constraints be identified? To what extent can linguistic structure previously ascribed to an innate UG be identified as having a nonlinguistic basis? Clearly, establishing a complete answer to this question would require a vast program of research. Here we divide the constraints into four groups relating to thought, pragmatics, perceptuo-motor factors, and cognition (Christiansen and Chater 2008).

Constraints from thought. The structure of mental representation and reasoning must, we suggest, have a fundamental impact on the nature of language. The structure of human concepts and categorization must strongly influence lexical semantics; the infinite range of possible thoughts must drive the compositionality of natural language (as discussed above); the mental representation of time is likely to have influenced the linguistic systems of tense and aspect; and so on. While the Whorfian hypothesis that language influences thought remains controversial, there can be little doubt that thought profoundly influences language.

Pragmatic constraints. Similarly, language is likely to be substantially shaped by the pragmatic constraints involved in linguistic communication. Pragmatic processes may, indeed, be crucial in understanding many aspects of linguistic structure, as well as the processes of language change.

Levinson (2000) notes that "discourse" and syntactic anaphora have interesting parallels, which provide the starting point for a detailed theory of anaphora and binding. Levinson argues how initially pragmatic constraints may, over time, become "fossilized" in syntax, leading to some of the complex syntactic patterns described by binding theory. Thus, one of the paradigm cases for arbitrary UG constraints may derive, at least in part, from pragmatics.

Perceptuo-motor factors. The motor and perceptual machinery underpinning language seems, moreover, inevitably to influence language structure. The seriality of vocal output, most obviously, forces a sequential construction of messages. A perceptual system with a limited capacity for storing sensory input forces a code which can interpreted incrementally (rather than the many practical codes in communication engineering, where information is stored in large blocks). The noisiness and variability (across contexts and speakers) of vocal or signed signals may, moreover, force a "digital" communication system, with a small number of basic units: i.e., phonetic features or phonemes. These discrete units in turn appear closely related to the vocal apparatus and to "natural" perceptual boundaries.

Cognitive mechanisms of learning and processing. Another source of constraints derives from the nature of cognitive architecture, including learning, processing and memory. In particular, language processing involves extracting regularities from highly complex sequential input, pointing to an obvious connection between sequential learning and language: both involve the extraction and further processing of discrete elements occurring in complex temporal sequences. It is therefore not surprising that sequential learning tasks have become an important experimental paradigm for studying language acquisition and processing (sometimes under the guise of "artificial grammar/language learning" or "statistical learning"; for reviews, see Gómez and Gerken 2000).

Syntax Shaped by Sequential Learning

If language has evolved to fit human sequential learning mechanisms, then constraints on the learning and processing of sequential structure should be reflected in the universal properties of human language. Importantly, many of the cognitive constraints that have shaped the evolution of language would still be at play in our current language ability. Thus, the study of how artificial sequential material is learned may reveal selectional pressures operating on the evolution of natural languages. We summarize a series of modeling and experimental results that indicate how constraints on sequential learning may have given rise to certain word-order universals relating to head-ordering, as well as interactions between case and word-order flexibility.

Assuming that language acquisition and processing share mechanisms with sequential learning in other domains, then breakdown of language would be expected to be associated with impaired sequential learning. This prediction is particularly interesting, because breakdown in sequential learning does generally not co-occur with cognitive impairments. This prediction has been tested using an artificial grammar learning task involving agrammatic aphasics who typically have damage in or around Broca's area and have severe problems with the hierarchical structure of sentences (Christiansen and Ellefson 2002). Although both aphasics and normal controls, matched for age, socio-economic status, and abstract reasoning abilities, were able to successfully complete a training task in involve same-different judgments on symbol strings, only the control group could correctly determine which of a set of novel strings was generated by the same rules as the training strings.

We would predict that basic word-order universals might arise from constraints on sequential learning, if sequential learning and language share common mechanisms. To pursue this hypothesis, let us begin with the *heads* of phrases: the word that determines the properties and meaning of the phrase as a whole (such as the noun *boy* in the noun phrase "*the boy with the bicycle*"). Across the world's languages, there is a statistical tendency toward a basic format in which the head of a phrase consistently is placed in the same position — either first or last — across different types of phrase. English is considered to be a head-first language, meaning that the head is most frequently placed first in a phrase, as when the verb is placed before the object noun-phrase in a transitive verb-phrase such as "*eat curry*." A head-last language, such as Hindi, typically uses the opposite order, and hence the equivalent of "*curry eat*." Likewise, head-first languages tend to have prepositions before the noun-phrase in prepositional phrases (such as "*with a fork*"), whereas head-last languages tend to have postpositions following the noun-phrase in postpositional phrases (such as "*a fork with*"). In the traditional UG framework, head-order consistency has been explained by innate language-specific constraints on the phrase structure of languages.

A very different picture emerges if we hypothesize that word order has evolved to fit human sequential learning mechanisms. Christiansen and Devlin (1997) trained simple recurrent networks[9] (Elman 1990; SRN) on corpora generated by 32 different grammars that differed in head-order consistency (i.e., inconsistent grammars would mix head-first and head-last phrases). The networks were trained to predict the next lexical category in a sentence. Although these networks had no built-in linguistic biases, their predictions were sensitive to the amount of head-order consistency found in the grammars, such that there was a strong correlation between the degree of head-order consistency in a grammar and how successfully the networks learned the language: the more inconsistent the grammar, the harder it is to learn (Figure 15.5). Christiansen and Devlin further analyzed frequency data on the world's natural languages

[9] It is sometimes objected that these kinds of networks lack biological plausibility because they are typically trained using back-propagation. However, recent advances in neural computation undermine this objection by demonstrating that the kind of networks employed in the simulations reported here can be trained with similar results using reinforcement learning (Grüning 2007), which is a neurobiologically plausible learning algorithm.

concerning the specific syntactic constructions used in the simulations. They found that languages incorporating patterns that the networks found hard learn tended to be less frequent.

Incorporating systems of case marking, Lupyan and Christiansen (2002) were able to relate learnability in the networks with attested frequency of different orders of subjects (S), verbs (V) and objects (O), across the world's languages. Subject-first languages, which make up the majority of language types (SOV: 51% and SVO: 23%), were easily learned by the networks. Object-first languages, on the other hand, were not well learned, and have very low frequency in the world's languages (OVS: 0.75% and OSV: 0.25%). Using rule-based language induction, Kirby (1999) arrived at a similar account of typological universals.

Lupyan and Christiansen (2002) also modeled data from a study by Slobin and Bever (1982) showing differences in performance across English, Italian, Turkish, and Serbo-Croatian when children acted out reversible transitive sentences, such as "*the horse kicked the cow*," using familiar toy animals. Like the children, the networks initially showed the best performance in Turkish, with English and Italian quickly catching up, and with Serbo-Croatian lagging behind. The close match between network performance across training and that of children across age is illustrated in Figure 15.6. Because of their consistent use of case and word order, respectively, Turkish and English were more easily learned than Italian and, in particular, the highly inconsistent Serbo-Croatian language. With repeated exposure, the networks learning Serbo-Croatian eventually caught up, as do the children learning this language.

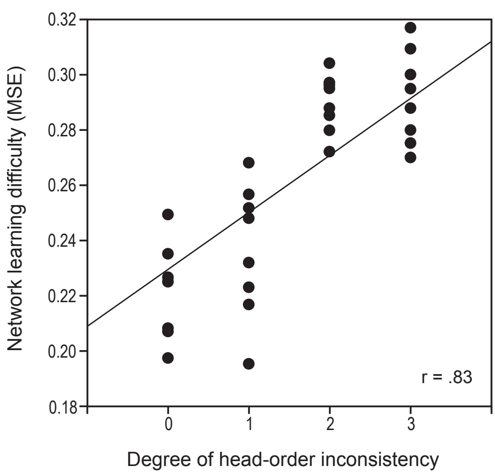

Figure 15.5 Relating the degree of head-order inconsistency and ease-of-learning in a connectionist network. Higher degrees of head-order inconsistency result in increased learning difficulty. (Adapted from Christiansen and Devlin 1997).

To determine whether these sequential learning biases would result in the emergence of consistent head ordering across successive generations of learners, Reali and Christiansen (2009) trained SRNs to map words onto grammatical roles. Prior to the introduction of language, the SRNs were first allowed to evolve "biologically" to improve their ability to perform a sequential learning task. Specifically, the initial weights from the best learner at each generation were chosen as the basis for the next, with copies of the parent's weights mutated slightly. After 500 generations, the SRNs had evolved a considerably better ability to deal with sequential structure. A language with no word-order constraints was introduced into the simulation. Crucially, both language and networks were allowed to change while the networks at the same time also had to maintain the same level of performance on the sequential learning task as obtained after initial evolution of sequential learning biases (on the assumption that this skill would still have been crucial for hominid survival after the emergence of language). Over generations, a consistent head-ordering emerged due to linguistic adaptation rather than biological adaptations (of initial weights). Indeed, the pressure toward maintaining a high level of sequential learning performance prevented the SRNs from adapting biologically to language.

If sequential learning is a fundamental human skill, as explored in these simulations, it should be possible to uncover the source of some of the universal constraints on language by studying human performance on sequential learning tasks. In a sequential learning experiment (Christiansen and Ellefson 2002), human participants learned sequences generated by either a consistent or inconsistent grammar from Christiansen and Devlin (1997). When tested

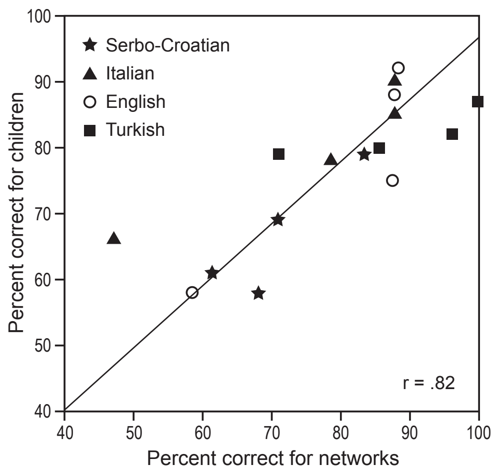

Figure 15.6 Using network performance as a function of training to predict the improvements in children's performance with increasing age in Turkish, English, Italian, and Serbo-Croatian. (Network results from Lupyan and Christiansen 2002; child data from Slobin and Bever 1982).

on novel sequences, the participants trained on the grammar with a consistent head-ordering were significantly better at distinguishing grammatical from ungrammatical items compared to participants trained on the inconsistently head-ordered grammar. Together, these simulations and human experiments suggest that sequential learning constraints may provide an alternative explanation of head-order consistency without UG. Specifically, constraints on basic word order may derive from nonlinguistic constraints on the learning and processing of complex sequential structure. Grammatical constructions with highly inconsistent head-ordering may simply be too hard to learn and therefore tend to disappear.

It is possible, moreover, that human sequential learning abilities are a crucial pre-adaptation to language. Conway and Christiansen (2001) reviewed evidence on sequential learning in nonhuman primates and concluded that although the performance of nonhuman primates on learning fixed sequences and certain types of statistical structure is similar to that of humans, the former has problems dealing with the kind of hierarchical sequential structure characteristic of human languages. This sequential learning may help explain why only humans have complex linguistic abilities.

Summary and Conclusion

The fundamental explanatory goal of linguistics is to answer the question: Why are languages the way they are and not some other way? We firmly believe that the only viable approach to this question is an evolutionary one: to answer why languages have the structure that they do, we need to ask how they came to be that way. However, language is not the result of a single adaptive system, and approaches that look to biological adaptation as their sole explanatory mechanism lead us to the wrong conclusions. In particular, we have highlighted the importance of taking into account the interactions between individual learning biases and cultural evolution in order to understand the sources of linguistic structure.

The problem is that the interactions between culture, biology and individual learning are very complex, perhaps uniquely so when it comes to human language. The solution is to explore theories about their interactions by building, in miniature, models that take seriously the notion that population-level phenomena like languages must emerge from the lower-level interactions between individual learners. We have briefly summarized here a spectrum of different modeling approaches, from mathematical models, through computational simulations, to novel experimental frameworks. All the results we have so far come together to demonstrate that languages adapt culturally under influence from limitations on human learning and processing. Furthermore, this kind of cultural adaptation may reduce the influence of biological adaptation. We are left with the conclusion that many key features of language, such as syntactic structure, may be adaptation by language to the problem of being passed-on through generations of language learners.

Acknowledgments

The authors are listed in reverse alphabetical order and have contributed equally to this paper. SK was supported by ESRC Grant Number RES-062-23-1537 and by AHRC Grant Number AH/F017677/1. MHC was supported by a Charles A. Ryskamp Fellowship from the American Council of Learned Societies and by the Santa Fe Institute; NC was supported by a Major Research Fellowship from the Leverhulme Trust and by ESRC Grant Number RES-000-22-2768.