Published in *Erkenntnis*, 80, 895-922 (2015).
Please refer to the published version: http://link.springer.com/article/10.1007/s10670-014-9688-8

Nativism, Empiricism, and Ockham's Razor

Simon Fitzpatrick

Department of Philosophy

John Carroll University

sfitzpatrick@jcu.edu

Abstract

This paper discusses the role that appeals to theoretical simplicity (or parsimony) have played in the debate between nativists and empiricists in cognitive science. Both sides have been keen to make use of such appeals in defence of their respective positions about the structure and ontogeny of the human mind. Focusing on the standard simplicity argument employed by empiricist-minded philosophers and cognitive scientists—what I call "the argument for minimal innateness"—I indentify various problems with such arguments—in particular, the apparent arbitrariness of the relevant notions of simplicity at work. I then argue that simplicity ought not be seen as a theoretical desideratum in its own right, but rather as a stand-in for desirable features of theories. In this deflationary vein, I argue that the best way of interpreting the argument for minimal innateness is to view it as an indirect appeal to various potential biological constraints on the amount of innate structure that can wired into the human mind. I then consider how nativists may respond to this biologized version of the argument, and discuss the role that similar biological concerns have played in recent theorizing in the Minimalist Programme in generative linguistics.

1. Introduction

In the course of his attack on the Cartesian doctrine of innate ideas, John Locke claimed that:

> [M]en, barely by the use of their natural faculties, may attain to all the knowledge they have, without the help of any innate impressions, and may arrive at certainty without any such original notions or principles… It would be impertinent to suppose, the ideas of colours innate in a creature, to whom God hath given sight… and no less unreasonable would it be to attribute [innate ideas] when we may observe in our selves faculties, fit to attain as easy and certain knowledge of them, as if they were originally imprinted on the Mind. (Locke, 1690, Book 1, Chapter 2, Section 1).

Locke's argument can be understood as an appeal to *Ockham's Razor*, often expressed in the slogan, "entities are not to be multiplied beyond necessity" or, more generally, as the maxim that simpler theories or explanations are to be preferred to less simple ones, other things being equal. Given that we possess the ability to acquire all our ideas and concepts from experience, it is, according to Locke, unnecessary to appeal to any innate ideas or concepts in explaining the development of individual human understanding. It is simpler to presume empiricism: that all of our ideas and concepts are derived from experience.[1]

Many subsequent philosophers and psychologists have followed Locke's lead in appealing to Ockham's Razor when defending broadly empiricist views of the structure and ontogeny of the human mind. It is a common view in modern cognitive science, for example, that we should attribute to infants an innate language faculty, an innate folk physics, or an innate folk psychology only as an absolute last resort, if it seems impossible to account for the development of their linguistic, physical, or psychological understanding through general learning processes alone. The philosopher Jesse Prinz (2012) has recently defended this attitude, using an argument

that mirrors Locke's. Here, for instance, is Prinz's argument against Chomskian nativist theories of language:

> The statistical learning approach has a great advantage over the innateness hypothesis: simplicity. All sides to the debate must admit that humans automatically engage in statistical learning. Our ability to notice patterns, recognize familiar objects and make decisions based on prior decisions depends on this… Nativist approaches clearly can explain language acquisition (they can build in as much innate language as the evidence requires), but there is no decisive evidence that these approaches are on the right track. Statistical approaches have not yet proven that they can explain language acquisition, but there are no decisive arguments for the conclusion that these approaches will fail. So which should we pick? One answer is that we should actively pursue both possibilities. It's good for science to explore all serious options. But, there is also a sense in which the statistical approach wins out in a stalemate. As the simpler theory, it should be the default explanation. We should not posit an innate language faculty unless we are forced to by decisive arguments, and we are not in that position yet. So even if we grant that both approaches should be actively pursued, we should also feel confident enough to regard the statistical approach as the more likely. (Prinz, 2012, p168-169; see also, p145-146)[2]

However, empiricists have not had Ockham's Razor all to themselves. Nativists have also often appealed to considerations of theoretical "simplicity" or "parsimony".[3] For instance, Robert Matthews endorses the following argument for linguistic nativism, which he finds in Chomsky's work:

> [Given that t]here are at present no substantive proposals as to how such domain specific [linguistic] knowledge might be acquired… it is reasonable on grounds of theoretical parsimony to conclude that this knowledge, which is apparently species-general, is, like other such traits, innate, i.e., endogenously determined, since to conclude otherwise (in the absence of any such proposals) would be to presume that the development of cognitive structures in the brain is to be treated differently than the development of physical structures elsewhere in the body. (Matthews, 2001, p221)

Matthews, following Chomsky, seeks to establish the opposite presumption to that defended by Locke and Prinz: nativism, not empiricism, ought to be the "default" assumption in cognitive science. Ockham's razor tells us to assume that species-typical psychological traits are, like species-typical bodily traits, innate, unless we have good reason to believe otherwise. Similarly, Steven Pinker (1984, p33-37) has argued that empiricist theories of cognitive development, while being more parsimonious with respect to infants' innate endowment, are often *less* parsimonious overall. This is because they often imply that the organization of the infant's cognitive architecture changes radically over course of development, and this violates what Pinker calls the *continuity assumption*: "the null hypothesis in developmental psychology is that the cognitive mechanisms of children and adults are identical" (1984, p7). In so far as nativist theories better satisfy this assumption, they should be preferred, other things being equal.

This enthusiasm for appeals to notions of theoretical simplicity in debates about the structure of the mind is perhaps not surprising, given that a long tradition in both science and philosophy has held that simplicity—broadly construed to include various kinds of theoretical and ontological economy, unity, and so forth—is a general *theoretical virtue,* and constitutes one of the core criteria that should be used for evaluating and choosing between rival theories. There

are some well known cases in the history of science where considerations of simplicity have been thought to have played an indispensible role in facilitating theoretical progress—the Copernican appeal to the "simplicity" and "harmony" of heliocentric astronomy (see Martens, 2009) and Newton's use of his first two "rules of reasoning in natural philosophy" (both of which embodied an explicit preference for simple theories) in the argument for universal gravitation (Newton, 1999) being two of the most prominent examples. For many, such examples are enough to show that relative simplicity ought to be taken very seriously in the evaluation of scientific and philosophical theories,[4] including those considered by cognitive science.

Nonetheless, one does wonder about the cogency of the sorts of simplicity arguments that have featured in the nativist-empiricist dispute. As we have just seen, such arguments are sometimes intended to establish particular theories or research programmes as "default explanations" or presumptive "null hypotheses", shifting the burden of proof onto the opposition; and sometimes they are invoked as tiebreakers when the empirical data is seen as failing to decide between competing theories—though theory T1 and T2 can both account for the empirical data, we should regard T1 as "more likely" or otherwise preferable to T2 because it is "simpler" or "more parsimonious". Yet, the accounting involved in these evaluations of comparative simplicity often appears dubious. At the very least, the notions of simplicity invoked seem sufficiently vague to arouse doubts about how principled such judgements really are—it is instructive that *both* sides have claimed Ockham's Razor to be on their side. Moreover, given the known complexity of evolved biological systems, it is far from clear why we should care about the relative simplicity of psychological theories, and even less clear why any type of theory should, as a result of such considerations, be regarded as the "default explanation" or presumptive "null hypothesis" in cognitive science. Merely pointing to other instances in the

history or current practice of science where simplicity considerations have been seen as important is not sufficient to allay such concerns.

My aim in this paper is to look at the role that arguments appealing to simplicity have played in the debate between nativist and empiricist views of the mind, and attempt to disentangle sense from mere rhetorical point-scoring. My goal is not to defend a particular nativist or empiricist position about the structure of the mind, but rather to frame this particular aspect of the debate in a more productive way. My focus will be on the standard argument employed by Locke, Prinz, and others to defend empiricist views in cognitive science—what will I call "the argument for minimal innateness"—but my conclusions will also have bearing on the arguments employed by nativists. In Section 2, I look closer at the argument for minimal innateness, and show that (somewhat surprisingly) even some nativists have endorsed it. I look at one tempting justification for the argument, which appeals to the idea that nativist accounts often posit explanatorily redundant entities. I show why this justification fails, and highlight the fact that the argument for minimal innateness must therefore rest on substantive assumptions about the role of a particular kind of parsimony in cognitive science. In Section 3, I turn to the issue of assessing the comparative simplicity of rival psychological theories, and highlight grounds for being sceptical that the relative simplicity of such theories can in fact be assessed in a principled way. In Section 4, I articulate my own preferred view of the role of simplicity considerations in choosing between scientific theories: what I call a "deflationary" account of simplicity, which builds, most notably, on the work of Elliott Sober. According to this view, it is misleading to regard simplicity as a theoretical desideratum in its own right. Instead, appeals to simplicity ought to be seen as indirect appeals to other, less problematic virtues of theories. In this vein, in Section 5, I argue that the best way of interpreting the argument for minimal innateness is to view it as an indirect appeal to various potential biological constraints on the amount of innate

structure that can be wired into the human mind. In Section 6, I consider how nativists may try to respond to this biologized version of the argument, and in Section 7 I look at the role that similar biological concerns have played in recent nativist theorizing about the structure and evolution of language within the Minimalist Programme in generative linguistics. The Minimalist Programme itself has often been seen as motivated by a particularly strong presumption of theoretical simplicity in its conception of language, which critics have found both obscure and ill motivated (e.g., Lappin et al., 2000; Pinker and Jackendoff, 2005; Kinsella and Marcus, 2009). However, I argue that recent articulations of the programme clearly fall in line with a deflationary conception: it is best understood not as a drive towards simplicity for simplicity's sake, but rather as an attempt to construct an evolutionarily plausible nativist account of language.

2. The argument for minimal innateness

The argument for minimal innateness, exhibited in the words of Locke, Prinz, and many others, asserts that we should, on grounds of theoretical simplicity, posit as little innate[5] structure as possible in explaining cognitive development. Consequently, we should prefer empiricist theories of cognitive development to nativist theories, and minimal nativist theories to more full-blooded nativist theories, other things being equal. If we *can* offload explanatory responsibility for the development of particular psychological traits onto learning we *should* do so; nativist theories, which place the explanatory responsibility on innate structure are to be avoided if at all possible.

This argument has been extremely influential in the cognitive science literature. Indeed, it is notable that those with more empiricist inclinations have not been the only ones to endorse it. For example, though developmental psychologist Renée Baillargeon is generally friendly to nativist views, and argues that it is plausible that infants posses quite rich innate concepts and expectations about how the physical world operates, she argues that we can only "take seriously"

such nativist claims if it turns out that such concepts are things that "infants' learning mechanisms are ill-equipped to learn" (Baillargeon, 1999a, p121). The reason for this is parsimony:

> Why assume that infants possess from birth physical concepts that we have empirical reasons to believe could be readily learned? Such an assumption would be unparsimonious unless and until additional evidence rendered the learned-continuity hypothesis unparsimonious. (Baillargeon, 1999b, p160)

Similarly, within the field of generative linguistics, which seeks to explain language acquisition by positing an innate Universal Grammar (UG)—a body of principles concerning the grammatical features of human natural languages, which is taken specify the "initial state" of the human language faculty before competence in a particular natural language is acquired—there is disagreement over how much of the apparent grammatical variation between human languages is "parameterized" in UG. So-called "underdeterminists" (e.g., Pinker and Bloom, 1990; Newmeyer, 1998) think that those grammatical features of natural languages that appear open to variation—such as whether the language is head-initial, like English ("throw the ball"), or head-final, like Japanese ("ball throw")—are left open by UG, and are acquired by learning through exposure to a particular language. Hence, the infant is taken to learn at least *some* language-specific grammatical principles during development. "Overdeterminists" (e.g., Chomsky, 1981; Baker, 2001), on the other hand, who advocate what has become known as the "Principles-and-Parameters" (P&P) approach to UG, think that all of these features are "parameterized" in UG: the grammars of all of the possible human languages are, in effect, innately represented in the mind of the infant in the form of a limited number of possible parameter settings, which are set

like toggle switches through exposure to a particular language. On this view, the infant doesn't learn *any* principles of grammar at all; what she learns is to which of set of principles, specified by the range of possible parameter settings, the local language conforms. Underdeterminists have claimed that their position is to be preferred because it is more "parsimonious" than that of the overdeterminist camp, since it is assumed to reduce the overall *quantity* of innate mental representation that must be postulated (see Baker, 2005).

Here again, it seems, even amongst quite strong nativists, the view is that we should, on grounds of parsimony, postulate as little innate structure as possible in order to explain cognitive development. If we *can* off-load explanatory responsibility for the development of particular psychological traits onto learning, we *should* do so.

2.1 Strong Anti-Nativism and explanatory redundancy

The argument for minimal innateness has often been seen as motivating a particularly strong methodological conclusion about the status of nativist theorizing in cognitive science. In this section, I will explain what this strong conclusion is, why it would not follow from the argument for minimal innateness even if the argument were sound, and why there is good independent reason not to accept it in any case. In so doing, I also want to set aside the tempting idea that the argument for minimal innateness is merely a call for the elimination of explanatorily redundant posits.

At a minimum, it seems that the argument for minimal innateness is meant to establish the following:

> Theories of cognitive development that posit less innate structure are, other things being equal, better than theories that posit more innate structure.

This can be seen as following from two claims: first, that theories that are more conservative with the amount of innate structure that they posit are more parsimonious, and, second, that more parsimonious theories are better theories, other things being equal. Ignoring for the moment the question of whether either of these two claims are in fact well motivated, there is an issue about exactly what methodological implications this conclusion has for the business of doing cognitive science. If we accept that less strongly nativist, more empiricist theories are more parsimonious, and that we should prefer more parsimonious theories, what methodological view should we take about the comparative acceptability of more strongly nativist and empiricist theories?

It seems that many of those who endorse the argument for minimal innateness seem to take it to have the following methodological implication:

> *Strong Anti-Nativism*: Empiricism is the default hypothesis for cognitive science. Additional innate structure over and above general learning mechanisms can only be posited if such mechanisms are *incapable* of acquiring the psychological trait in question. Moreover, no more innate structure than is strictly necessary to augment learning should be countenanced.

In effect, Strong Anti-Nativism is the idea that nativism about a particular psychological trait can only be acceptable if we have reason to believe that the trait is something that could not *possibly* be acquired through learning. It is therefore incumbent on nativists to provide a *poverty of the stimulus* (PoS) argument in support of any claim for the innateness of a particular psychological trait—an argument that shows that the trait is not learnable for a given learning mechanism because there is not sufficient information present in the learning environment for the trait to

acquired through the operation of that learning mechanism alone.[6] In the absence of such an argument, we should default to an empiricist learning hypothesis. We are only warranted in increasing the size of our inventory of innate structure in so far as this is required to surmount the inability of unaugmented learning to successfully acquire the trait.

Several of the passages quoted earlier are strongly suggestive of Strong Anti-Nativism. Such a view seems to be exactly what Locke was getting at when he said that "[i]t would be impertinent to suppose, the ideas of colours innate in a creature, to whom God hath given sight". Similarly, Strong Anti-Nativism seems to be at work when Renée Baillargeon argues that Spelke's (1994) nativist theory of infants' conceptions of how the physical world operates can only be taken "seriously" "if it turns out that [these concepts are] something that infants' learning mechanisms are ill-equipped to learn" (Baillargeon, 1999a, p121). A clear symptom of this view is the tendency to infer from the fact that a given psychological trait is learnable in principle that *it is in fact learnt*.

One likely reason why many theorists are attracted to Strong Anti-Nativism is because it is tempting to interpret the argument for minimal innateness as a kind of *anti-redundancy* argument—i.e., as merely calling for the elimination of posits that do *no explanatory work* in our theory of cognitive development. The idea is something like this: both nativists and empiricists tend to agree that humans have some basic general learning mechanisms—for instance, as Prinz emphasizes, capacities for statistical learning and pattern recognition.[7] If these mechanisms have to be posited anyway, and we have no reason to think that such mechanisms are not capable of facilitating acquisition of the psychological trait in question, the postulation of further innate structure seems to be explanatorily redundant—it can explain nothing that the pre-existing postulation of general learning mechanisms cannot explain. Thus, to avoid positing explanatorily

redundant psychological structure, we should only advocate innate structure if general learning by itself could not possibly do the job.

I think that it is this idea that underlies the tendency of cognitive scientists to predicate the acceptability of nativism on the plausibility of a PoS argument, since it suggests that the only cases in which posting innate structure over and above general learning mechanisms will not be explanatorily redundant are cases where general learning mechanisms would be *incapable* of acquiring the trait without it. If this were correct, the argument for minimal innateness would also obtain clear justification, for theoretical posits that do no explanatory work in our theories are also posits for which we have no evidence. Plausibly, unobserved theoretical entities are legitimated by empirical data only to the extent that they can help us to explain why the data is the way it is. If a theoretical entity fails to contribute to this end, the data fails to confirm its existence.[8]

However, we cannot make sense of the argument for minimal innateness in these terms. A simple example should make this clear. Suppose a species of songbird possesses general learning mechanisms, which we have no reason to believe are not capable of acquiring the species-typical song. Suppose also that we have no reason to believe that the information available to the species in its normal developmental environment is insufficient for the trait to be learnable: chicks receive ample exposure to the birdsong as they develop, and possess all the required capacities for acquiring it through general learning. Given all of this, does it remain a coherent possibility that birdsong is nonetheless innate in this species? The answer is clearly "yes". One way of determining whether a trait is innate even if it is learnable is by performing a "deprivation" experiment: we could see whether chicks raised in an environment where they are never exposed to birdsong nonetheless acquire it successfully. If they do, we have powerful evidence for innateness. Note that this sort of test for innateness does not rely on showing that the chick's

learning mechanisms are *incapable* of acquiring the birdsong, as do PoS arguments, which aim to show that the normal developmental environment is too impoverished for general learning to be able to explain the development of the trait. Rather, the test relies on showing that the trait is not *in fact* learnt because development can proceed without significant alteration in a radically impoverished environment.[9]

In this example, it is clearly not appropriate to say that postulating further innate structure is somehow explanatorily superfluous or redundant. Though the postulation of innate structure cannot explain anything that the postulation of general learning mechanisms cannot explain in principle, it may nonetheless be the case learning cannot actually explain the acquisition of the species-typical birdsong because in this species *it is in fact innate*. In such a situation, it would clearly be the innate structure and not learning that does the explanatory work, even though the trait could have been readily learnt. It is not the case, then, that nativists must postulate explanatorily redundant structure in cases where the trait could potentially be acquired by general learning processes alone. Showing that a particular psychological trait could have been learnt is not sufficient for showing that it is learnt. Once this is accepted, we see that Strong Anti-nativism does not follow from the legitimate desire to remove explanatorily redundant posits from our theories. We need a more substantive justification for the claim that we should default to empiricism if we have no reason to believe that the trait is not learnable.

This shows why Strong Anti-Nativism should not be accepted, even if we assume the plausibility of the argument for minimal innateness. Strong Anti-Nativism could only be taken as a reasonable methodological restriction if *only* PoS arguments can provide evidence for nativism. Though many of the debates between nativists and empiricists have concentrated on the plausibility of various PoS arguments—language acquisition being the prime example—it is important to recognize that there are other arguments for nativism. For instance, nativists have

appealed to evolutionary considerations (arguing that natural selection will tend to favour innateness over learning for some psychological traits), to the lack of cultural variability in certain traits (which would be surprising were they acquired solely through learning), to comparative data (arguing that some human psychological traits are homologous to plausibly innate traits in animals), and to genetic data (in particular, to developmental disorders that selectively impair the development of some psychological traits). Hence, alongside PoS arguments, nativists have sought to enlist a wide variety of other sorts of considerations as evidence for the view that many psychological traits have an innate basis. Empiricists have, of course, resisted these sorts of arguments, seeking to find alternative non-nativist explanations or disputing the empirical data on which such claims have been based. But, the point remains that a nativist account of the development of a particular psychological trait can be perfectly well motivated, even if the trait in question could be readily acquired through general learning—i.e., even if there is a "wealth of the stimulus".

2.2 Weak Anti-Nativism

In place of Strong Anti-Nativism, it seems that the methodological implication of the argument for minimal innateness would be better understood as follows:

> *Weak Anti-Nativism*: the fact that a theory has a smaller inventory of innate structure than a rival theory should be considered as counting in favour of the theory. Conversely, having a larger inventory of innate structure should be considered as counting against it. This preference may, however, be overridden by evidence that warrants positing more innate structure.

Weak Anti-Nativism allows that nativists can support their case without providing a PoS argument. Thus, one could potentially endorse a nativist account of the development of a particular psychological trait, even if there is a "wealth of the stimulus" in the organism's normal developmental environment, so long as there is sufficient evidence that overrides the initial preference for a more empiricist account. In what follows, I will assume that the argument is meant to establish only Weak Anti-Nativism.

3. Measuring the relative simplicity of psychological theories

The argument for minimal innateness claims that we should, on grounds of simplicity, prefer theories of cognitive development that posit less innate structure to theories that posit more. Consequently, it is claimed that empiricist theories should be preferred to nativist theories, and minimal nativist theories should be preferred to more full-blooded nativist theories, other things being equal.

However, as we have seen, some nativists have challenged the idea that positing more innate structure must come at the cost of being unparsimonious. For example, Pinker (1984, p33-37) argues that greater parsimony with respect to infants' innate endowment for language typically comes at the cost of making our overall theory of cognitive development *less* parsimonious. According to Pinker, a workable empiricist theory of language acquisition may have to concede to the nativist that a language-specific acquisition device is essential—if natural language grammars are in fact learnable, only mechanisms dedicated to learning language have any hope of succeeding. One way to incorporate language-specific learning into an account that eschews the idea of there being an innate language-specific acquisition procedure is to suggest the following sort of constructivist story (e.g., Elman et al., 1996): infants, as they mature, essentially learn *how to learn* language. The infant starts out with a set of general-purpose,

largely unstructured neural networks, with a small set of innate biases for searching out particular kinds of environmental input, including linguistic input. Over the course of development, these networks grow and self-organize themselves in response to the different kinds of environmental input they are biased to seek out. Differences between kinds of environmental input are taken to drive specialization, so that individual networks form modules specialized for particular domains of input. Once this has occurred, these modules then progressively adapt themselves to the idiosyncrasies of particular domains. In the case of language, the claim is that a specialized set of networks dedicated to language will emerge in the infant's mind, which subsequently facilitate domain-specific learning. This account therefore attempts to buy language-specific learning mechanisms without innateness. The power of endowing self-organizing neural networks with minimal innate biases for seeking out linguistic input is taken to be such that infants can acquire a language-specific acquisition module that can deal with all the idiosyncrasies of language learning, without any innate pre-specification for forming this module.

Pinker, discussing earlier versions of constructivism, argues that such an approach is strikingly "unparsimonious". It amounts to postulating "a superfluous acquisition procedure for the acquisition procedure" (1984, p36). In contrast, the linguistic nativist can just cut to the chase, postulating an innately structured, language-specific acquisition device. Hence, while this sort of constructivist theory may achieve greater parsimony than a nativist theory in terms of the amount of innate cognitive structure that is postulated, this must come at the cost of a much less parsimonious theory of the developmental process itself. In particular, it violates Pinker's main guiding parsimony assumption, the *continuity assumption*: "the null hypothesis in developmental psychology is that the cognitive mechanisms of children and adults are identical" (1984, p7). So, while Pinker is prepared to accept that parsimony considerations may, in some instances,

motivate a preference against nativism, he argues that it is plausible that the most empirically adequate empiricist theory on the market is in fact less parsimonious than a nativist one:

> Either there exists an underlying non-linguistic basis for language acquisition mechanisms [i.e., in terms of general learning] or there does not. If there does, any theorist is free to propose a theory of what those general mechanisms are, and if such a theory really does account for language acquisition and other phenomena in a simpler manner than a language-specific innateness theory, the latter theory can be rejected as unparsimonious. If no such theory exists, then the innateness theory will win out over all competitors. (Pinker, 1984, p37)

Pinker's argument shows that the argument for minimal innateness focuses on one particular sense in which theories of cognitive development might be considered simpler than others—parsimony with respect to the amount of innate structure posited. Yet, there is another way in which we might measure the simplicity of theories of cognitive development that, in Pinker's view, comes down on the side of nativist theories. The Chomskian argument endorsed by Robert Mathews provides another possible measure of simplicity: it is more "parsimonious" to assume that psychological traits are innate and hence like other bodily traits, unless we have good reason to believe otherwise. Such an assumption helps to maximize the unity of psychology with biology. In a similar vein, Margolis and Laurence (2013) have suggested that models of non-human animal psychology tend to be thoroughly nativist, with theorists being quite happy to posit innate structure to explain various aspects of animal behaviour. There is, therefore, an "evolutionary parsimony" argument for adopting a similar nativist orientation with respect to human behaviour.

Though they have not to my knowledge been part of any explicit arguments proposed by nativists, there are a variety of other ways in which nativist theories might be considered to be "simpler" or "more parsimonious" than empiricist theories. For example, another potential way of evaluating the simplicity of theories is by counting the number of adjustable parameters they contain—simpler theories having fewer such parameters. One way in which one might go about counting adjustable parameters here is to compare rival cognitive theories when they are implemented in connectionist neural networks. Cognitive theories may be said to have as many free parameters as the respective networks have adjustable connection weights (Marcus, 2000). Patterns of connectivity in neural networks are said to be learnt if the relevant connection weights are the product of adjustments made in response to input to the network, and innate if the connection weights are pre-specified at the outset. The most extreme nativist theory, where all connection weights are pre-specified, will therefore come out as the simplest, since it will have zero adjustable parameters; the most extreme empiricist theory, where all connection weights are freely adjustable, will come out as the most complex, since it will have as many free parameters as connections.

Consider also the number of auxiliary assumptions that rival theories of cognitive development require. By definition, empiricists try to off-load as much explanatory responsibility as possible for the development of particular psychological traits onto learning. In so doing, it might be suggested that they must make more (and much stronger) auxiliary assumptions about the structure of the infant's learning environment than must nativists. An empiricist theory has to assume not only that appropriate information is present in the environment, so that the trait is potentially learnable, but also that it is available in the *right sort* of way, so that it can be reliably extracted by general learning mechanisms. Thus, there is a direct trade-off between positing less

innate structure and making more auxiliary assumptions about the structure of the learning environment. In this respect, nativist theories might be seen as simpler than empiricist theories.

The point I want to emphasize is that, even if it is granted that empiricist theories tend to be simpler than nativist theories, in the sense that they are more parsimonious with respect to the amount of innate structure that is posited in explaining cognitive development,[10] this does not suffice to show that empiricist theories do tend to be simpler theories overall. This is because there are a number of other respects in which nativist theories may be considered simpler than empiricist theories. I am not suggesting that we *should* measure the simplicity of theories of cognitive development in any of these ways—for instance, I am not advocating Pinker's claim that nativist theories of language acquisition come out simpler overall. Rather, my point is that, in so far as they appeal to a general theoretical virtue of simplicity, proponents of the argument for minimal innateness must either explain why we should ignore these other dimensions of simplicity, or provide us with some way of evaluating how these considerations are to be balanced against each other.

This discussion serves to highlight a long-running concern of philosophers of science about the notion that simplicity ought to be a criterion for evaluating scientific theories: the vagueness of the notion of simplicity itself. In an effort to make the notion more precise, philosophers and, more recently, statisticians and information theorists, have proposed a number of different formal measures of simplicity that can be employed in various contexts (see [AUTHOR 2] for a survey of the literature)—for instance, to measure the simplicity of statistical models—but no adequate general account of how to measure the simplicity of scientific theories has been forthcoming, particularly not one that can be applied to the sorts of qualitative theories at issue in this debate. And what we see in this case is something that has been under-recognized in the philosophical literature on simplicity: the fact that scientific theories can be assessed for

their simplicity in a multitude of different ways, and that for each respect in which T1 can be regarded as "simpler" than T2, there is usually a number of other respects in which T2 can be regarded as "simpler" than T1. This creates a problem in determining which kind of simplicity is to be valued over others. In so far as the argument for minimal innateness focuses on one particular way of assessing the simplicity of theories of cognitive development to the exclusion of others, without our being offered a clear explanation for why we should focus on *that* particular conception of simplicity, we have grounds for doubting that the argument is really a principled one. The same would also appear to be the case for arguments of nativists. Pinker, for instance, gives us no explanation for why the continuity assumption ought to carry more weight than parsimony with respect to innate structure.

These problems would seem to be compounded when we ask the further question about what *justification* we can have for preferring putatively simpler theories of cognitive development. Philosophers of science have also wrestled long and hard with the problem of explaining why, in general, it may rational to favour simpler theories over more complex ones (see [AUTHOR 2] for a survey), but even if we think that an adequate rationale can be articulated for the standard sorts of cases that philosophers of science tend to focus on (Copernicus versus Ptolemy, and so forth)—though that certainly cannot be taken for granted—we might (as I noted in the introduction) expect there to be particular problems in this case, given the known complexity of evolved biological systems. The eminent biologist Francis Crick once complained, "[w]hile Occam's razor is a useful tool in physics, it can be a very dangerous implement in biology. It is thus very rash to use simplicity and elegance as a guide in biological research" (Crick, 1988, p138).

4. Deflating simplicity

Given these difficulties, it is tempting to conclude that all such appeals to simplicity in the nativist-empiricist dispute are completely unprincipled and without merit. In a scientific and philosophical community weaned on uncritical reverence for Ockham's Razor, simplicity, it seems, functions as a convenient methodological cudgel that theorists can use to beat the opposition. Yet, in this case at least, such appeals do not survive closer scrutiny.

While I accept that many appeals to simplicity here and elsewhere in science and philosophy ought perhaps be regarded as empty rhetoric, I want to suggest a different way of thinking about the argument for minimal innateness. In my view, it provides a clear case study in support of a *deflationary* account of simplicity. According to such an account, simplicity shouldn't be seen as a theoretical desideratum in its own right. Rather, when theorists appeal to simplicity, and those appeals do seem to carry some weight, they are better understood as appealing to something else. In this sense, "simplicity" or "parsimony" should be regarded as a stand-in for other, less problematic, virtues of theories that do the real underlying epistemic work. A number of philosophers of science have advocated such views in the context of attempts to make sense of apparent role of simplicity in inductive and abductive inference (e.g., Boyd, 1990; Norton, 2003; AUTHOR 1, 2). Most notably, looking at some common appeals to simplicity in debates in evolutionary biology, such as the debate over the legitimacy of group selection explanations for adaptive traits, Elliott Sober (1994) has argued that such arguments are often best understood, not as appeals to simplicity *per se*, but rather as covert appeals to local background assumptions. The claim that group-level selection hypotheses are to be avoided in evolutionary biology because they are "unparsimonious" compared with hypotheses invoking selection at the level of individual organisms or genes, for instance, makes sense, according to Sober, on the assumption that conditions required for group-level selection to take place are quite onerous and obtain relatively rarely. The parsimony argument against group selection is thus

better understood as an indirect appeal to these particular local background assumptions (which reduce the prior probabilities of group selection hypotheses relative to alternatives), rather than an appeal to some general theoretical virtue of simplicity.

As I have argued elsewhere (AUTHOR 1, 2), a key advantage of this deflationary perspective is that it removes the burden of articulating and defending an account of what simplicity is and how it should be measured, since simplicity *per se* drops out of consideration altogether. In addition, as Sober emphasizes, once the relevant underlying considerations for which simplicity acts as a proxy are made explicit, we can move forward in evaluating what weight such arguments should actually carry in our theoretical decisions—e.g., the extent to which we should actually avoid advancing group selection hypotheses. It remains an open question about how much of the apparent role of simplicity in science can be deflated in this way (particularly, the famous cases from the history of science),[11] but, as I will now try to show, a deflationary analysis make most sense in the present context.

5. "Passing the buck to biology"

Nativists are often accused of "passing the buck to biology" (Dennett, 1980). By claiming that a particular psychological trait is innate, nativists shirk the responsibility of explaining the development of the trait by transferring responsibility from the domain of psychology to the domain of developmental or evolutionary biology. Worse, it is sometimes argued that this makes nativism tantamount to claiming that the development of the trait is a "mystery" (e.g., Cowie, 1998). This sort of criticism seems ill founded. As Chomsky (1980) notes, it only makes sense insofar as one also perceives there to be similar explanatory problems with regarding non-psychological traits, such as the heart, as being innate rather than learnt.

Nonetheless, there is a sense in which the accusation that nativists pass the buck to biology does seem to be relevant to the argument for minimal innateness. For instance, Culicover and Jackendoff worry that:

> Critics of generative grammar… are justified in being suspicious of a learning theory that depends on the child having an innate language capacity that contains, say, an intricately crafted definition of government… This is more than a quibble about scientific elegance. In order for such intricacy to be present in the prelinguistic child, it must be constructed in the brain (somehow) from the human genetic code. In turn, the genetic code ultimately has to be a product of genetic variation and natural selection in prelinguistic hominids (or perhaps earlier, if it serves some purpose more general than language). Granted, we know virtually nothing about how any innate cognitive capacity is installed in a brain by a genetic code, much less the dimensions of variation possible in such codes. But that doesn't absolve us from at least keeping this problem in mind, and therefore trying to minimize the complexity of UG in an effort to set the stage for eventual explanation. (Culicover and Jackendoff, 2005, p13)

Culicover and Jackendoff complain about the tendency of some linguistic nativists to put very large amounts of innate structure into the human language faculty, without appreciating that there are biological costs to this promiscuity with innate structure. All psycholinguistic theories must ultimately provide explanations for how the innate structures that they posit have come about, and how they are passed on from one generation to the next. This suggests that there ought to be two sorts of biological constraint on nativist theorizing. First, the informational load that can be carried by the human genome is finite; it is reasonable to think that only so much mental structure

can be encoded in our genetic endowment. Second, human evolutionary history is also finite, as are the powers of natural selection and other evolutionary processes; it is reasonable to think that evolution could have endowed us with only so much innate structure. These constraints provide upper bounds on the amount of innate structure that nativists can reasonably posit without carrying their theories into the realms of biological implausibility. While these constraints are difficult to quantify, given current knowledge, Culicover and Jackendoff worry that too many nativists have indeed ignored the importance of these constraints.

Culicover and Jackendoff offer such biological concerns as an *additional* consideration in theory evaluation in linguistics, on top of "general grounds of parsimony" and empirical adequacy.[12] However, I suggest that "parsimony" ought to fall out altogether: we will be able to see at least some plausibility in the argument for minimal innateness if we construe it, not as an appeal to simplicity or parsimony *per se*, but rather as a covert appeal to the idea that these (and perhaps other) biological constraints on nativist theorizing are relatively severe. In other words, there will be at least some cogency to the idea that we should regard the fact that one psychological theory posits more innate structure than another as being significant when it comes to evaluating the relative plausibility of these theories, if we assume that there are significant evolutionary and developmental constraints on the amount of psychological structure that can plausibly be innate in humans. Given such a background assumption, parsimony with respect to innate structure may become a potential form of *evidence* that could count in favour of one psychological theory and against another. The degree to which such a difference in "parsimony" will be evidentially significant will, of course, depend on how severe we determine these constraints to be. This is an empirical question, and a very difficult one at that. Evidence from biological constraints will also clearly have to be balanced against other forms of evidence. Empiricists theories do, after all, have to be able explain how particular traits could plausibly be

acquired through learning. It will be no use being parsimonious with innate structure in the name of greater biological plausibility, if this means endorsing an empirically inadequate theory of acquisition. The fact (if it is a fact) that there are significant constraints on the amount of innate structure that can be reasonably posited does not, therefore, automatically entail that empiricism will be more plausible than nativism.

Nonetheless, with these qualifications in hand, this deflationary perspective provides by far the best way of construing the argument for minimal innateness. First, it avoids all the problems with measuring and comparing the relative simplicity of rival theories of cognitive development. If these biological constraints are genuinely severe, parsimony with respect to innate structure will clearly be an evidential factor, irrespective of whether theories that are parsimonious with respect to innate structure come out as simpler (or more parsimonious) theories overall. Second, if we do in fact have good reason to believe that these biological constraints are severe, this would clearly give evidential weight to the argument—evidential weight that would clearly have to be balanced against other factors, but evidential weight nonetheless. Third, this reconstruction is in terms of very real concerns that have featured in the nativist-empiricist dispute—namely, biological constraints on psychological theorizing. The above quoted passage from Culicover and Jackendoff is one example of the manifestation of these concerns on the part of committed nativists. Unsurprisingly, those with more empiricist inclinations have been quick to latch onto these biological issues. Bates and Elman (2002), for instance, have argued that there is a "gene shortage" when it comes to explaining cognitive development: at around 20-25,000 protein coding genes—many fewer than researchers had predicted before the completion of the Human Genome Project—the human genome is too compact to encode all the innate traits that nativists are inclined to postulate. In their view, there is a "mathematical bottleneck" involved in setting up a brain with around 10^{14} connections using

so few genes. Given this, they argue that we can only realistically expect that our innate psychological endowment will consist in a plastic "proto-cortex"—a brain built ready to learn, but with very little innate structure and knowledge wired in.

Broader evolutionary considerations have also figured in discussion. Andy Clark, for instance, has questioned the evolutionary plausibility of Chomskian nativism on the grounds that "[e]volution will surely favour *minimal nativism*":

> Instead of building in large amounts of innate knowledge and structure, build in whatever minimal set of biases and structure will ensure the emergence, under realistic environmental conditions, of the basic cognitive modules and structures necessary for early success and for subsequent learning. (Clark, 1993, p185)

Clark defends this evolutionary claim on the grounds that natural selection is the "laziest of designers" (ibid). If the adaptive knowledge state can be produced as a reliable developmental outcome under realistic environmental conditions merely by the "fixing of a few assorted parameters so as to facilitate the *subsequent* attainment (by learning) of states properly described as involving knowledge" (p186 [emphasis in original]) instead of building in the knowledge state as itself innate, then evolution will favour this strategy: "evolution will, in the laziest possible way, have slightly loaded the dice" (p187).[13] This is a different sort of evolutionary constraint from the one identified by Culicover and Jackendoff, who are more concerned with the ability of evolutionary processes to wire in lots of intricate innate structure, such as abstract principles of UG, over a constrained period of time. Clark's claim is about the *character* of evolutionary engineering: that evolution tends to build organisms that use the environment as a developmental scaffold whenever it can.

My suggestion, then, is that biologizing the argument for minimal innateness in the way I have described allows us to reduce empiricist concerns about simplicity to quite genuine background biological concerns in cognitive science. Of course, taking this route means engaging with very difficult and highly controversial issues in evolutionary and developmental biology. As Culicover and Jackendoff rightly point out, current knowledge in these fields has not reached a state at which solid foundation can be given to any particular conclusions about the nature and severity of the biological constraints on human psychology. Even so, it still seems a more productive way for the empiricist-minded to pursue objections to nativist theories that have so far been couched in terms of problematic appeals to simplicity.

In this vein, I now want to consider how nativists might respond to some of the empirical claims about biological constraints on nativist theorizing just described, looking in particular at the role that similar such concerns have played in the Minimalist Programme in generative linguistics. Once again, my goal is not to defend any particular nativist or empiricist position, but rather to map out some of the issues that need to be taken in account in evaluating the biologized argument for minimal innateness.

6. Nativist responses

In reply to the "gene shortage" and "mathematical bottleneck" claims that have been cited in objection to strongly nativist theories of cognitive development, Gary Marcus (2004) has argued the surprising compactness of the human genome does not, in fact, represent a particularly significant constraint on nativist theorizing. He compares the structure of the genome and the developmental process that it instigates to information compression techniques used in software programming, but argues that the information compression that genes are capable of puts all current human achievements to shame. Even though the number of genes is vanishingly small

compared to the number of cells in the body, the quantity of information that they capture is vast. These few genes work together in very complex ways, controlling the development with enormous informational economy: "There's no gene shortage because nature has figured out how to use the same genes over and over" (Marcus, 2004, p158). Marcus does not by any means claim that there are *no* informational constraints on innate specification. Rather, he claims that the informational constraints on the genome are nowhere near as great as has been speculated, and, consequently, that nativists should not be too concerned about "mathematical bottlenecks".

Marcus, like those posing the gene shortage idea, adopts standard coding/software metaphors for describing the role of genes in the development of phenotypes. Such metaphors assume the primacy of the genome in development and in the transmission of phenotypes from one generation to the next. In recent decades, however, some developmental biologists have urged less a gene-centric perspective, some even emphasizing "parity" between genomic factors non-genomic factors in development and inheritance (Oyama, 2000). Such a view suggests another possible response: nativists needn't place the entire burden on the genome, with each innate psychological trait having to be "coded" for by a limited number of genes; they might be able to bring in resources from the wider developmental system.[14] For instance, a recurring speculation in Chomsky's work is that some innate cognitive features may be the product of basic physical constraints on brain development, and thus, in a sense, "come for free" from physics (e.g., Chomsky, 2005). Along these lines, Cherniak (2010) suggests that some innate brain structures may emerge in development from general principles of neural network optimization, such as minimizing the length of connections between neurons, giving rise to the possibility of a "non-genomic nativism".

The claim that evolution is a lazy designer and will therefore tend to build developmental processes in organisms that leave as much room as possible for the environment to provide

information necessary for the development of adaptive traits, rather than wiring them into the organism innately, seems to ignore lots of important considerations. One obvious response is the suggestion that, at least in some cases, there will be significant fitness advantages to making a trait innate, rather than have the organism rely on environmental scaffolding. Learning is, after all, vulnerable to error, and error may be extremely costly to the organism if acquiring the trait is integral to survival. Innateness might therefore make organisms less vulnerable to perturbations in their environment and lead to stronger developmental canalization of the trait across environments where there relevant information required for learning is corrupted or unavailable. Trail and error learning may also impose higher metabolic costs on the organism than innateness because of extra brain activity brought about by error. If there are often significant fitness advantages to innateness, then we may have good reason to be sceptical about the idea that evolution will generally favour the lazy option.[15]

With regard to evolvability constraints of the sort suggested by Culicover and Jackendoff, there will clearly be significant differences in the extent to which they will motivate parsimony with respect to innate structure when we are concerned with different sorts of psychological traits. Evolvability constraints will be less pressing when we are concerned with psychological traits that we plausibly share with non-human animals. This is because nativists will then have a much longer evolutionary time period to work with. They will be most pressing when we are concerned with distinctively human traits that will have to have a evolved over a much shorter time span. This means that nativists about language might have to be more concerned about how much innate structure they posit than nativists about folk-physical or folk-mathematical concepts, which are likely to be widely dispersed in the animal kingdom (see, e.g., Gallistel, 2010). The key problem for linguistic nativists, it seems, is requiring that too much distinctively human innate architecture will have evolved in the six million years or so that separate us from our

closest ape relatives. Indeed, this problem may be particularly severe if, as is generally (but by no means universally) held, quite recent hominid species such as Neanderthals lacked anything like modern human language, meaning that the capacity for modern language evolved no more than 200,000 years ago, and perhaps within the last 100,000 years.

This brings us to the latest incarnation of Chomskian generative linguistics, the so-called "Minimalist Programme" (MP), which can be seen as a direct attempt to reconcile a thoroughgoing nativist view of language with such concerns about its recent evolution.

7. The Minimalist Programme and the evolution of language

Chomsky's original articulation of MP was as an attempt to answer the question, "[h]ow 'perfect' is language?": "…we seek to determine just how far the evidence really carries us toward attributing specific structure to the language faculty, requiring that every departure from 'perfection' be closely analyzed and well motivated" (1995, p9). "Language" is here understood in terms of the faculty of language (FL), a cognitive system that mediates between articulatory-perceptual systems and conceptual-intentional systems, or, in other words, between sound (or gestures, in the case of sign language) and meaning. FL would be a "perfect" or "optimal" system, according to Chomsky, if all that it contains is what is *conceptually necessary* for relating sound and meaning.

Chomsky has long argued that the key defining feature of human language, which sets it apart from the communication systems of non-human animals, particularly those of other primates, is unlimited *recursion*: the capacity to string lexical items together into complex combinatorial structures of potentially infinite expressive power.[16] The most fundamental computational operation that seems to be conceptually necessary, then, is an operation that Chomsky calls "Merge", which takes a linguistic item, *X*, and combines it together with another

linguistic item, *Y*, to form a new item, *Z*, that constitutes the set of *X* and *Y*. With Merge operating repeatedly on its own outputs, an infinite number of possible structured expressions can be generated from a finite stock of basic items from the lexicon. The primary impetus of research in MP is therefore to see how far all other language-specific architecture can be minimized.

The strongest minimalist approach to language would be an attempt to explain all of the seemingly complex features of linguistic constructions in terms only of the operation of Merge, plus the "interface" conditions imposed by articulatory-perceptual systems and conceptual-intentional systems. The idea is that many of the peculiar features of such constructions may result from this interaction with these other systems—i.e., FL having to translate syntactically structured expressions into a format usable by other systems—rather than being core features of FL itself. For example, consider the phenomenon of displacement, discussed by Berwick and Chomsky (2011, p31-32), where a phrase is pronounced in one position in a sentence, but interpreted in another. In *guess what Mary ate*, for instance, *what* is interpreted as the object of *ate*, but pronounced in a different place in the sentence. They argue that Merge can be held to operate in at least two ways: in *external* Merge two distinct objects are joined together, while in *internal* Merge (regarded as a distinct operation, Move, in early versions of MP, but now assimilated to Merge), one of the merged objects, *X*, is a part of the other, *Y*, which is copied from its occurrence in *Y* and then merged with *Y*. The two types of Merge allow us to go from *Mary ate what* via *what Mary ate what* (internal Merge) to *guess what Mary ate what* (external Merge). Berwick and Chomsky argue that the original occurrence of *what* may then be suppressed when the sentence is externalized in speech, for reasons of computational efficiency, resulting in the sentence that is actually pronounced. Displacement, then, may result from the interaction of Merge with basic requirements of computational efficiency—in this case, operative at the interface with articulatory systems. Similar proposals have been made for other features of

language, the key idea being that apparent surface complexity may turn out to be the product of a surprisingly simple computational system, Merge, interacting with other systems, and constrained by general, language-independent considerations, such as computational efficiency.

The emergence of MP has been met with a fair amount of incredulity. In particular, it has been seen as motivated by an obscure and idiosyncratic notion of theoretical elegance and simplicity that functions in an *a prioristic* way: MP is not the product of choosing the simpler of two equally empirically adequate theories, or an attempt to make an existing empirically adequate account of language simpler, but rather a drive towards simplicity (in the guise of the notion of "perfection") as a fundamental *presumption* about the nature of the language itself (e.g., Lappin et al., 2000; Pinker and Jackendoff, 2005; Kinsella and Marcus, 2009). Lappin et al. (2000) even go so far as to label work in MP "unscientific", and only taken seriously because of Chomsky's status and personal influence in the field. Their concern, in particular, is the apparent inclination of linguists working within MP to get rid of large chunks of the architecture of UG found in Government and Binding Theory, the most developed articulation of the P&P approach, not on the basis of any empirical results, but merely in order to show that language is "perfect". The conception of FL as "perfect" or "optimal" is also seen as at variance with it being a biological structure, since such structures are typically imperfect and sub-optimal (Pinker and Jackendoff, 2005; Kinsella and Marcus, 2009).

There are important questions about the ability of future theorizing in MP to account for all the empirical facts about human languages—particularly, whether an approach based exclusively on Merge-like operations is feasible. However, it seems clear that the conceptual criticisms of the motivations behind MP just described miss the point. As recent articulations have made clear (e.g., Chomsky, 2007; Berwick, 2011; Boeckx, 2011a; Chomsky and Berwick, 2011; see also, Drummond and Hornstein, 2011 for a nice overview), MP is best understood as

an attempt to avoid an apparent trade-off between what linguists call "Plato's problem"—the problem of explaining how linguistic competence can be so quickly and reliably acquired by the infant—and what has more recently been dubbed, "Darwin's problem"—the problem of showing how language (FL) could have evolved. Earlier work in the generative tradition judged that the best way to address Plato's problem—given substantive empirical claims about the poverty of the linguistic stimulus that infants are exposed to—was to develop a theory of FL that minimized what it is that the child has to learn. This was a key driving force behind the notion of UG in the P&P approach (and hence of GB), where the initial state of FL is held to be a rich body of innate principles of grammar, with a few assorted parameters that the infant merely has to adjust to the linguistic input. An adequate response to Darwin's problem, however, seems to pull in the other direction, since if UG is full of such rich, apparently language-specific information, it is hard to see how it could have evolved so quickly in recent human history. MP is represented as aiming to square this circle: i.e., reconciling the generative vision of language with the constraints imposed by Darwin's problem.[17]

Contra Lappin et al., MP does not assert that GB provides a *false* account of UG. Researchers within MP adopt the working assumption that the core syntactic principles posited by GB are approximately correct as descriptions of the "laws of grammar" (Drummond and Hornstein, 2011). Rather, the aim of MP is to determine the extent to which principles of the sort posited by GB may emerge in a bottom-up fashion from the interaction of Merge with other cognitive systems. As Chomsky (2005, p9) puts it, if successful, this would allow us to "no longer assume that the means of generating expressions are highly articulated and specific to language. We can seriously entertain the possibility that they might be reducible to language-independent principles". Moreover, the evolutionary speculation is that UG may have been produced by the minimal addition of a computational system for recursion (Merge) to pre-

existing systems. This would mean that there was no set of historical junctures at which evolution wired-in each principle of UG. These principles were, in a sense, already there in the interface conditions that Merge had to meet.[18] The hope, then, is that a strong nativism, adequate to address Plato's problem as the generative tradition has conceived it, is consistent with a strong evolutionary conservatism of the sort seemingly demanded by Darwin's problem.[19]

The upshot of this seems to be that "simplicity", "perfection", "optimality", "minimality", etc., are not significant in MP purely for their own sake, but insofar as pursuing particular understandings of these notions helps to satisfy the constraints seemingly imposed by Plato's problem and Darwin's problem.[20] Indeed, as Chomsky himself has put it, "[w]ithin the biolinguistic framework, methodological considerations of simplicity, elegance, etc., can often be reframed as empirical theses concerning organic systems generally" (2007, p1).[21] There seems, therefore, to be some convergence between recent thinking about these notions in generative linguistics and the deflationary account of simplicity. Moreover, we can see that MP reveals additional considerations that need to be taken in account in evaluating the biologized argument for minimal innateness—in particular, the possibility that a strong nativism about language needn't necessarily be at variance with it emerging so recently in evolutionary history.

8. Concluding remarks

Appeals to simplicity have played a significant role in the nativist-empiricist dispute in cognitive science. Both nativists and empiricists have sought to defend their respective positions about the structure and ontogeny of the human mind by arguing that theirs is "simpler" or "more parsimonious" than the rival position. The weight theorists have placed on these sorts of arguments is striking. For example, Prinz's (2012, p168-169) appeal to simplicity is pretty much his *only* argument for linguistic empiricism, his discussion largely consisting of attempts to show

that the putative evidence for linguistic nativism is not as strong as nativists have claimed, thus allowing empiricism to win out in a stalemate in virtue its supposedly greater simplicity. On the other side of this debate, Steven Pinker has even gone so far as to suggest that:

> [T]he innateness debate in the study of language acquisition be recast as a parsimony debate concerning rival acquisition theories for language and other cognitive domains. This debate then becomes an empirical one—inasmuch as any empirical debate ultimately hinges on parsimony and elegance, it being commonly accepted that data always underdetermine the theories explaining them. (Pinker, 1984, p36)

One finds similar arguments in many other debates in cognitive science—for example, in the debate over whether non-human primates are capable of mindreading (see [AUTHOR 1]). A reason for the popularity of such arguments might be the speculative nature of the subject matter. Since compelling empirical evidence for one theory over another is hard to come by in debates over the nature of the mind, partisans may feel the need to resort to such methodological arguments to bolster their respective positions. The appropriateness of these arguments is by no means self-evident, however. In particular, as we've seen, they tend to focus on very particular ways in which theories can be regarded as simple or complex, without our being given clear reason to focus on *that* particular type of simplicity at the expense of others that may motivate different theoretical preferences. As a result, it is hard to resist the feeling that these arguments lack a principled basis, and this is before we even ask questions about what justification we can have for preferring simpler theories of cognition.

My claim, however, is that it is more productive to view arguments such as the argument for minimal innateness in a deflationary manner: as really being about something other than

simplicity *per se*—in this instance, presumed background biological constraints on innate structure. Such an analysis clearly brings several advantages in this case: it not only helps to avoid concerns about the argument for minimal innateness being *ad hoc* (since it is not some general theoretical virtue of simplicity that does the real epistemic work, we are relieved of the burden of explaining why parsimony with respect innate structure ought to trump other types of simplicity that may point in different directions), but also shows us how we might move forward in evaluating what weight, if any, the argument for minimal innateness should carry in our theoretical decisions, by focusing attention on the nature and extent of these potential biological constraints on innate structure, and how we might balance them against other countervailing considerations. The same, I suggest, holds for the simplicity arguments of nativists—indeed, as we've just seen in the case of MP, there seems to be some convergence in nativist circles towards deflationary conceptions of such arguments. Even if we find ourselves on difficult and speculative terrain at the intersection of cognitive science and evolutionary and developmental biology, this is still much firmer ground on which to prosecute this particular aspect of the nativist-empiricist dispute.

References

Ariew, A. 2007. Innateness and triggering. In M. Matthen and C. Stephens (eds.), *Handbook of Philosophy of Biology*. New York: Elsevier.

Arnold, K. and Zuberbühler, K. 2008. Meaningful call combinations in a non-human primate. *Current Biology*, 18(5), R202-R203.

Baillargeon, R. 1999a. Young infants' expectations about hidden objects: a reply to three challenges. *Developmental Science*, 2, 115-132

Baillargeon, R. 1999b. Response to Smith and commentators. *Developmental Science*, 2, 157-161.

Baker, M. 2001. *Atoms of Language*. New York: Basic Books.

Baker, M. 2005. Innate endowment for language: underspecified or overspecified? In P. Carruthers, S. Laurence and S. Stich (eds.), *The Innate Mind: Structure and Contents*, Oxford: Oxford University Press.

Barnes, E.C. 2000. Ockham's razor and the anti-superfluity principle. *Erkenntnis*, 53, 353-374.

Bates, E. and Elman, J. 2002. Connectionism and the study of change. In M. Johnson (ed.), *Brain Development and Cognition: A Reader (2nd edition)*. Oxford: Blackwell.

Berwick, R.C. 2011. Syntax facit saltum redux: Biolinguisitcs and the leap to syntax. In A. DiSciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise*. New York: Oxford University Press.

Berwick, R.C. and Chomsky, N. 2011. The biolinguistic program: The current state of its development. In A. DiSciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise*. New York: Oxford University Press.

Berwick, R.C., Okanoya, K., Beckers, G.J., and Bolhuis, J. 2011. Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences*, 15(3), 113-121.

Berwick, R.C., Pietroski, P., Yankama, B., and Chomsky, N. 2011. Poverty of the stimulus revisited. *Cognitive Science*, 35, 1207-1242.

Boeckx, C. 2011a. Some reflections on Darwin's problem in the context of Cartesian linguistics. In A. DiSciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise*. New York: Oxford University Press.

Boeckx, C. 2011b. Approaching parameters from below. In A. DiSciullo and C. Boeckx (eds.), *The Biolinguistic Enterprise*. New York: Oxford University Press.

Boyd, R. 1990. Observations, explanatory power and simplicity. In R. Boyd, J. Glasper and J.D. Trout (eds.), *The Philosophy of Science*. Cambridge, MA: MIT Press.

Carruthers, P. 2006. *The Architecture of the Mind*. New York: Oxford University Press.

Cheney, D. and Seyfarth, R. 2007. *Baboon Metaphysics*. Chicago: University of Chicago Press.

Cherniak, C. 2010. Brain-wiring optimization and non-genomic nativism. In M. Piattelli-Palmarini, J. Uriagereka, and P. Salaburu, *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*. New York: Oxford University Press.

Chomsky, N. 1980. The new organology: replies to peer commentaries. *Behavioral and Brain Sciences*, 3, 42-62.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

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

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry*, 36, 1-22.

Chomsky, N. 2007. Approaching UG from below. In U. Sauerland and M. Gaertner (eds.), *Interfaces + Recursion = Language?* Berlin/New York: Mouton de Gruyter.

Clark, A. 1993. *Associative Engines*. Cambridge, MA: MIT Press.

Cosmides, L. and Tooby, J. 1994. Origins of domain-specificity. In L. Hirschfeld and S. Gelman (eds.), *Mapping the Mind: Domain-specificity in cognition and culture.* New York: Cambridge University Press.

Cowie, F. 1998. *What's Within*? Oxford: Oxford University Press

Crick, F. 1988. *What Mad Pursuit: a Personal View of Scientific Discovery*. New York: Basic Books.

Culicover, P. and Jackendoff, R. 2005. *Simpler Syntax*. New York: Oxford University Press.

Dennett, D. 1980. Passing the buck back to biology. *Behavioral and Brain Sciences*, 3, 19.

Drummond, A. and Hornstein, N. 2011. Basquing in minimalism. *Biolinguistics*, 5, 331-349.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.

Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Forster, M. and Sober, E. 1994. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science*, 45, p1-36.

Gallistel, C. R. 2010. The foundational abstractions. In M. Piattelli-Palmarini, J. Uriagereka, and P. Salaburu, *Of Minds and Language: A Dialogue with Noam Chomsky in the Basque Country*. New York: Oxford University Press.

Jackendoff, R. 2011. What is the human language faculty?: Two views. *Language*, 87, 586-624.

Kinsella, A. and Marcus, G. 2009. Evolution, perfection, and theories of language. *Biolinguistics*, 3, 186-212.

Lappin, S., Levine, R., and Johnson, D. 2000. The structure of unscientific revolutions. *Natural Language and Linguistic Theory*, 18, 665-671.

Laurence, S. and Margolis, E. 2001. The poverty of the stimulus argument. *British Journal for the Philosophy of Science*, 52, 217-276.

Locke, J. 1690/1975. *An Essay Concerning Human Understanding*, P. H. Nidditch (ed). Oxford: Oxford University Press.

Mameli, M. and Bateson, P. 2006. Innateness and the sciences. *Biology and Philosophy*, 21, 155-188.

Marcus, G. 2000. *The Algebraic Mind*. Cambridge, MA: MIT Press.

Marcus, G. 2004. *The Birth of the Mind*. New York: Basic Books.

Margolis, E. and Laurence, S. 2013. In defense of nativism. *Philosophical Studies*, 165, 693-718.

Martens, R. 2009. Harmony and simplicity: aesthetic virtues and the rise of testability. *Studies in the History and Philosophy of Science*, 40, p258-266.

Matthews, R. 2001. Cowie's anti-nativism. *Mind and Language*, 16, 215-230.

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

Newton, I. 1999. *The "Principia": Mathematical Principles of Natural Philosophy*, *A New Translation by I. Bernard Cohen and Anne Whitman*. Berkeley: University of California Press.

Norton, J. 2003. The material theory of induction. *Philosophy of Science*, 70, 647-70.

Oyama, S. 2000. *The Ontogeny of Information: Developmental Systems and Evolution (2nd Edition)*. Durham, NC: Duke University Press.

Papineau, D. 2005. Social learning and the Baldwin effect. In A. Zilhão (ed.), *Cognition, Evolution, and Rationality*, London: Routledge.

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pinker, S. and Bloom, P. 1990. Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-784.

Pinker, S. and Jackendoff, R. 2005. The faculty of language: what's special about it? *Cognition*, 95, 201-236.

Prinz, J. 2002. *Furnishing the Mind: Concepts and Their Perceptual Basis*, Cambridge, MA: MIT Press.

Prinz, J. 2003. Author's response. *Metascience*, 12, 297-303.

Prinz, J. 2012. *Beyond Human Nature*. New York: W.H. Norton.

Samuels, R. 2002. Nativism in cognitive science. *Mind and Language*, 17, 233-265.

Sober, E. 1994. Let's razor Ockham's Razor. In E. Sober, *From a Biological Point of View*, Cambridge: Cambridge University Press.

Sober, E. 2007. *Evidence and Evolution*. New York: Cambridge University Press.

Sober, E. 2008. Empiricism. In S. Psillos and M. Curd (eds.), *The Routledge Companion to the Philosophy of Science*. New York: Routledge.

Spelke, E. 1994. Initial knowledge: six suggestions. *Cognition*, 50, 431-445.

[1] In addition to such a *psychological* empiricism concerning the nature of human cognitive development, Locke was also defending an *epistemological* empiricism concerning the epistemic ground for knowledge. It is only psychological empiricism that is at issue in this paper—though it is perhaps worth noting that more recent epistemological empiricists have not always had an easy relationship with putatively "extra-empirical" methodological principles like Ockham's Razor (see Sober, 2008).

[2] Prinz uses a similar argument for his neo-Lockean theory of concepts, according to which all concepts are built solely from perceptual representations. Prinz takes it as a given that our perceptual faculties provide us with representations of the world, then writing:

> Empiricist theories can be defended by appeal to methodological parsimony. Once we have postulated a certain class of representations for a theory of perception, it is cost effective to see whether those same representations can be used for a theory of cognition. We should try to make use of what we already have rather than overpopulating the mind with other kinds of mental representations. I think parsimony considerations give us a reason for preferring empiricist theories over nonempiricist theories, all other things being equal. (Prinz, 2002, p122)

[3] These terms are typically used interchangeably, though sometimes the term "parsimony" is reserved for a specific type of theoretical simplicity, namely economy with respect to the number of entities or postulates that feature in the theory.

[4] Another important motivation for this view is the role that simplicity considerations have been thought to play in the practice of curve-fitting and statistical inference (see Forster and Sober, 1994).

[5] A note of clarification may be in order here. Though the nativist-empiricist dispute presupposes the legitimacy of the concept of *innateness*, there are notorious difficulties with clarifying exactly what it means to say that a cognitive or physiological trait is "innate" or "acquired" (see, e.g., Mameli and Bateson, 2006). Indeed, some philosophers of biology doubt that the innate/acquired distinction makes any sense at all. Following Richard Samuels (2002), I am inclined to think that the notion of innateness, at least as it is deployed in cognitive science, must make *some sense*, for without it we cannot distinguish between patently different views of cognitive development—for instance, Chomskian and Skinnerian accounts of language acquisition. Moreover, Samuels has provided a useful stopgap explication of this concept that we can use in lieu of a fully worked out account. According to Samuels, we can say that a psychological trait is innate if it is *psychologically primitive*: if there is no satisfactory explanation for the origin of this trait that falls within the bounds of the traditional explanatory resources of psychology. Naturally, this implies that a trait acquired through learning cannot be innate. A limitation of this proposal is that aphasias caused by brain injuries would count as psychologically primitive in this sense, but are clearly not innate. Hence, some rider has to be added, which excludes psychological traits that are the product of "abnormal" non-psychological origins. Nonetheless, this explication will do for the purposes of this paper.

[6] For discussion of the structure and function of PoS arguments, see Laurence and Margolis (2001), Berwick, Pietroski et al. (2011).

[7] This is not universally the case. Evolutionary psychologists and other proponents of the so-called "massive modularity thesis", for instance, who adopt strong nativist positions for most cognitive domains, deny that there is

any such thing as "general learning": natural selection has engineered the human mind to be entirely made up of domain-specific modules; all learning mechanisms are specialized for dealing with specific adaptive problems (e.g., Cosmides and Tooby, 1994; Carruthers, 2006).

[8] This idea allows us to distinguish between two versions of the slogan commonly identified with Ockham's Razor, "entities are not to be multiplied beyond necessity": the *anti-redundancy principle* and the *anti-quantity principle* (Barnes, 2000). Suppose theory, T1, explains the relevant data by positing the existence of entity, A, while theory, T2, posits the existence of A and another entity, B, but this additional posit fails to explain anything in the data. If there are no other grounds for asserting the existence of B, the anti-redundancy principle says that we should prefer T1 to T2. This is because we can regard B as an explanatorily redundant entity; therefore, T2 posits an entity for which we have no evidence. The anti-quantity principle goes further, asserting that we should prefer the theory that posits *the fewest total number of entities*, whether or not the relevant entities do explanatory work. For example, suppose that T1 and another theory, T3, are both equally able to explain the relevant data. T1 does so by positing one entity, A; T3 does so by positing the existence of two entities, C and D. Even if C and D are indispensible to the explanation of the data provided by T3, the anti-quantity principle says that we should prefer T1 to T3 because of its sparser theoretical ontology. Since there is no sense in which T3 posits explanatorily superfluous entities, the evidential justification just described cannot be given for the anti-quantity principle. Moreover, the anti-redundancy principle motivates only an agnostic stance towards the razored-off ontology, since absence of evidence is not (by itself) evidence of absence. It gives us no grounds for preferring a theory that asserts A & ¬B to a theory that asserts A & B. In contrast, the anti-quantity principle would say that we should indeed prefer the former theory. As we will see, the argument for minimal innateness fits better with the more substantive, but also more problematic, anti-quantity principle than with the anti-redundancy principle.

[9] Note that while acquisition under radical deprivation may be sufficient for innateness, it is not be necessary. The development of innate birdsong might still require exposure to birdsong in order to "trigger" its developmental process. For example, some songbirds require exposure to birdsong during development, but still develop the song that is typical of their own species, even if only exposed to different song of another species (Ariew, 2007).

[10] Indeed, it is not clear that off-loading more of the explanatory responsibility for the development of a trait onto learning will always yield greater parsimony with respect to innate structure. Baker (2005) argues that if the parsimony debate between underdetermination and overdetermination theories of UG is seen as hanging on the amount of innate representational structure that is required to implement UG in the human mind, assumptions about the kind of representational format that is postulated for UG are crucial to whether parsimony favours underdetermination or overdetermination. On some representational formats, leaving grammatical rules unspecified may actually require positing *more* innate representational machinery than specifying the options in advance. Consider, a rule specifying where words are to be inserted into sentences to build larger phrases. If the representational format imposes no *intrinsic* word ordering in mental representations of sentences, then UG can remain neutral between different rules about how words should be orderedm and it will be more parsimonious to leave the choice unrepresented. However, if the representational format does impose an intrinsic ordering, UG cannot leave word ordering unspecified without adding *more* representational machinery. In this latter instance, overdetermination will actually be more parsimonious than underdetermination. Sometimes adopting a stronger nativism, means positing *less* innate representational structure.

[11] Sober has presented an account of the role of simplicity in curve-fitting and statistical inference that fits broadly with this deflationary outlook, based on the work of the statistician, Hirotugu Akaike (Forster and Sober, 1994; Sober, 2007).

[12] Culicover and Jackendoff argue that such considerations support their "Simpler Syntax Hypothesis": "the most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning" (2005, p5). Although very much in the generative tradition (and thus largely nativist in orientation), their approach departs in significant ways from what they call "mainstream generative grammar", associated with Chomsky and followers. The key difference is that rather than focusing on minimizing the number of distinct *principles* of grammar, so that syntactic variation between languages is kept to a minimum, the Simpler Syntax Hypothesis emphasizes that the syntactic structure of linguistic expressions (as mentally represented in the minds of speakers and listeners) is only as complex as it needs to be to facilitate interpretation. Much of what is taken to be the explanatory responsibility of syntax on the mainstream approach is thus passed on to principles of semantic interpretation, supposedly doing away with much (but not all) of the complex hidden syntactic structure imputed to linguistic expressions by mainstream generative grammar. Although this may lead to a multiplication of distinct rules of grammar, which have to be learnt by the language learner, Culicover and Jackendoff claim that this produces a

much leaner and more evolutionarily tenable theory of UG—albeit not as lean as that found in Chomsky's more recent Minimalist Programme, which I'll discuss in Section 7.

[13] Prinz makes a similar appeal to evolutionary conservatism in defence of his parsimony argument for an empiricist theory of concepts:

> Parsimony arguments have even more force in psychology. If a mechanism that evolved for one function can do another, there would be no selection pressure on the evolution of another mechanism. If mechanisms of categorisation can serve as mechanisms for thinking about things, then dedicated thinking mechanisms would be unlikely to evolve. (Prinz, 2003, p301-302).

[14] Such views of development have sometimes been seen as undermining claims about innateness. However, they only require that we avoid the common tendency to equate "innate" with phrases like "genetically specified". It is here that Samuels' (2002) primitivism has a clear advantage over more gene-centric accounts of innateness.

[15] The emergence of trait in a lineage through learning may indeed *lead* to the trait becoming innate in later generations, via something like the Baldwin effect. For instance, a trait that initially spread through a population via social learning might subsequently become innate (or scaffolded by innate structure) in later generations, if that would give individuals a fitness advantage by, for instance, speeding up acquisition or reducing the costs of error in the acquisition process. It seems plausible that Baldwin effects may have taken place in the course of human evolution after the emergence of complex cultures (Papineau, 2005).

[16] Many primates have vocalizations (such as predator alarm calls) that plausibly carry semantic content, but only humans seem to be able to string words together into meaningful combinatorial structures (Cheney and Seyfarth, 2007; but see Arnold and Zuberbrühler, 2008). Interestingly, Berwick et al. (2011) argue that some species of songbird produce what look like hierarchically organized syntactic structures, but that their songs lack word-like elements. Only humans seem to have words *and* syntax.

[17] Chomsky (2007) has argued that the P&P approach was, in this respect, an advance on earlier generative theories, since the switch-box of parameters idea lessened the amount of internal complexity that had to be attributed to FL. MP seeks to take this goal of minimization further.

[18] This would imply that the principles of GB are not explicitly mentally represented in FL, which *is* a significant change from the P&P approach. There is also some doubt about whether the notion of syntactic parameters (parameterized principles, and the "overdeterminist" view of grammar) can be saved in such a bottom-up approach when it comes to explaining variation between languages, which most work in MP has attributed to features of the lexicon (Boeckx, 2011b).

[19] This story is meant to fit with Chomsky's long-held view that FL is not an evolutionary adaptation *for* communication. Recently, he has argued that Merge evolved to facilitate personal thought, with externalization of language in speech coming later (e.g., Berwick and Chomsky, 2011). In explicating this view, Chomsky has also suggested that recursion is both unique to, and new with the evolution of, Merge. As Jackendoff (2011) points out, this goes against a common view, held by many sympathizers with Chomsky, that recursion/combinatorial structure is ubiquitous in cognition—e.g., a fundamental feature of an independent language of thought (Fodor, 1975)—and that there is strong evidence for such structure in non-human animals that lack syntactically-structured communication. However, Chomsky (2007, p7) does seem to be at least open to the alternative possibility that recursive/combinatorial structure was appropriated from other systems. Either way, the basic evolutionary speculation of MP would remain the same: the key innovation that brought about FL and UG was recursion at the interface between articulatory-perceptual and conceptual-intentional systems.

Of course, something else that would also have to have evolved to give rise to full human language is the capacity for a much richer lexicon (presumably the result of changes to conceptual-intentional systems), along with a capacity for vocal imitation, seemingly absent in other extant primates (Berwick and Chomsky, 2011).

[20] This point undermines Culicover and Jackendoff's (2005) criticism that the MP view of language is actually not very simple because the tree structures generated by Merge end up being quite complex—the goal of their own Simpler Syntax Hypothesis being to reduce the amount of covert structure in mental representations of linguistic constructions. If we read MP concerns about simplicity as just suggested, the complexity of resultant mental representations seems beside the point.

[21] Much the same holds for the various "economy" principles that work in MP has invoked in efforts to characterize the operation of Merge. For instance, Merge is taken to abide by principles of computational efficiency, such as the No-Tampering Condition and Minimal Search. These are explicitly motivated by what Chomsky calls "third factor" considerations, basic biological or physical constraints on neural structure and organization that shape the nature of all cognitive processes, not just FL (see Chomsky, 2005).