1 Minimalism and Darwin’s Problem

1.1 Introduction

Contemporary generative theorists are united by (at least) one conviction and divided by (at least) one other. What unites everyone is the understanding that grammatical knowledge is rule based. Native speakers of a given language L have mastered rules for L that allows them to generate an unbounded number of tokens of L (i.e. sentences, phrases, etc.). Rules are required because the tokens of L are for all practical purposes infinite and thus cannot possibly be stored individually in a finite organism. The rule-based character of linguistic knowledge is, thus, not controversial among generative grammarians.¹

What is controversial is how these grammars are structured; what kinds of rules they allow, what kinds of primitive relations they exploit and what kinds of elements they involve. Here there is a lot of controversy. One line of inquiry with which I am very partial, the Minimalist Program, takes it as a boundary condition on inquiry that the basic operations of UG be simple and that the attested complexities of natural language be the result of the interactions of simple subsystems. This vision gains teeth when the meaning of “simple” is filled out. Here is how I understand the term.

There are several dimensions to simplicity.

First, simple systems are non-redundant. Redundancy arises in grammars when different operations can independently generate the same structural relations or different principles independently exclude them. An example (which is developed in more detail in later chapters) can serve to illustrate my meaning. Many current grammars postulate both a Move operation and an AGREE operation capable of operating over long distances.² Both serve to relate remote

¹ Which does not mean to say that it is not still controversial. There are many in the connectionist world who appear to deny the rule-based nature of grammatical knowledge. Such dissenters are happy enough to concede that natural language objects display patterns, but patterns are not rules. The problem with this view, I believe, is that it is quite clear that the number of possible patterns is likewise unbounded and that only rules will do. For discussion of this basic point see Jackendoff (1994).

² Note the qualification. That grammars involve agreement, i.e. some form of feature checking, is clear. The interesting operation is the non-local version of AGREE as it covers much the same
elements (non-sisters) to each other. All things being equal, grammars should not contain both kinds of operations as they can cover a great deal of the same empirical territory. This is not a good thing for at least two reasons. First, a UG with multiple routes to the same end gains an undesired flexibility, which adversely affects its explanatory potential. Methodologically speaking more brittle theories are more easily falsified and thus preferable. Further, more brittle UGs restrict the learner’s options more than more flexible ones do. If there are two ways of covering the same data set, then the learner must choose between them, seldom a good thing given the logic of Plato’s problem. Of course, things may not be equal and both operations might be required, but a good working hypothesis is that grammars are not redundant in this way.3

Second, in simple theories of UG the basic operations are as sparse as possible. Fewer is better. Ockham is right. All things being equal theories that employ a sparser inventory of principles and basic operations are better than those with an ampler armamentarium.4 Of course, oftentimes things are not equal. In such cases, I am inclined to a somewhat stronger allegiance to Ockham. It is a truism that the richer a theory’s apparatus the wider its empirical coverage. This means that sparser theories are expected to face empirical challenges that more ample theories will avoid. I understand this truism to mean that the latter should face more stringent explanatory demands before winning the day. Precisely because their data coverage is expected to be wider, more ample theories should either cover a hell of a lot more territory than their more restrained competitors or should do so in such ways that do not sacrifice explanatory insight. My version of Ockham strongly prefers the leaner meaner account and requires substantial advantages before it is abandoned!

Third, in simple accounts the basic operations and principles are natural. Just what makes such operations and principles “natural” is a subtle question.
1.1 Introduction

However, this has not prevented generative grammarians from arguing for and against proposals in just such terms over the years. For example, to the degree grammars facilitate “computation” they are natural, e.g. locality conditions (like subjacency or minimality) are “nice” properties from a computational point of view given the burden that distance imposes on computational efficiency and memory.\(^5\) Another example; feature checking and copying are natural computational operations for the faculty of language (FL) to exploit as they are almost certainly operative in other cognitive domains, albeit with different expressions being copied and different features being checked. Given the rather late emergence of FL in humans it is evolutionarily natural that FL should import operations from other parts of the cognitive system. This suggests one more mark of “naturalness,” namely generality; operations and principles at work in other parts of the cognitive economy are natural resources for linguistic computations. A further mark of the “natural” is the “atomicity” of the computational operations. Merge (join two expressions) and copy (duplicate an expression) are reasonably taken as computationally “atomic” operations.\(^6\) They contrast with more complex language specific rules like “passive” which are reasonably analyzed as compiled combinations of more basic operations. This conception of “simple” and “atomic” casts a furtive glance towards implementation in brain like material. Whatever operations grammarians propose must ultimately be embedded in brain circuitry. It is reasonably clear how one could build a merge or copy circuit, and this is one reason that primitive operations like these are attractive.

I would like to stress this last point. David Poeppel and colleagues have recently emphasized that any grammatical process we propose must be embodied in brain circuitry if it is really operative in our FL. However, the linking hypotheses between language and brain are “most likely to bear fruit if they make use of computational analyses that appeal to generic [my emphasis, NH] computational subroutines” (Poeppel and Monahan in press). Thus, keeping basic operations simple and generic comes with the advantage of conceivably being implementable.\(^7\)

In sum, FL will be natural if it is based on principles and operations that promote computational tractability, that are built from parts that are cognitively general and atomic, and that are basic enough to be (plausibly) embodied in neural circuitry.

\(^5\) See Chomsky (1977) for discussion along these lines for subjacency. See too Berwick and Weinberg (1984).

\(^6\) I would be inclined to say that they are primitively recursive, the building blocks for possibly more complex combinations. For discussion, see Chapter 7.

\(^7\) For some further discussion of how primitives of grammar should relate to primitives of neuroscience, see Embick and Poeppel (2005a).
As should be evident, even given the desiderata above, there remains plenty of room for diverging views on how to interpret these guidelines and, not surprisingly, there is a large pool of potential candidates for the inventory of basic operations and principles. Nonetheless, I believe that these guidelines can play a more than rhetorical role in the construction and evaluation of grammatical proposals. More concretely, I believe that the search for simple operations and principles suggests an interesting minimalist project: the construction of grammatical models based on a small inventory of operations and principles that are at once evolutionary and neurologically plausible and from which the basic properties of natural language grammars can be qualitatively derived. The reason for this is best articulated in an evolutionary idiom.

1.2 Minimalism and Darwin’s Problem

Over the last 50 years of research generative grammarians have discovered many distinctive properties of natural language grammars (NLG). For example: (a) NLGs are recursive, viz. their products (sentences and phrases) are unbounded in size and made up of elements that can recur repeatedly; (b) NLGs generate phrases which display a very specific kind of hierarchical organization (viz. that described by X’ theory); (c) NLGs display non-local dependencies (as in Wh-movement, agreement with the inverted subject in existential constructions, or reflexive binding), which are subject to hierarchical restrictions (e.g. binding relations are subject to a c-command requirement) and locality restrictions (e.g. controllers are subject to the minimal distance requirements and anaphors must be bound within local domains). These properties, among others, are universal characteristics of natural language and thus reasonably construed as universal features of human grammars. A widely adopted (and to my mind very reasonable) hypothesis is that these characteristics follow from the basic organization of FL, i.e. they derive from the principles of UG.

Given this, consider a second fact about FL: it is of recent evolutionary vintage. A common assumption is that language arose in humans in roughly the last 50,000–100,000 years. This is very rapid in evolutionary terms. It suggests the following picture: FL is the product of (at most) one (or two) evolutionary innovations which, when combined with the cognitive resources available before the changes that led to language, delivers FL. This picture, in turn, prompts the following research program: to describe the pre-linguistic cognitive structures that yield UG’s distinctive properties when combined with the one (or two) specifically linguistic features of FL. The next three chapters try to outline a version of this general conception.8

8 This clearly echoes the program outlined in Hauser, Chomsky and Fitch (2002).
1.2 Minimalism and Darwin’s Problem

The approach, I believe, commits hostages to a specific conception of FL. It does not have a high degree of internal modularity. The reason for this is that modular theories of UG suppose that FL is intricately structured. It has many distinct components that interact in complex ways. On the assumption that complexity requires natural selection and that natural selection requires time to work its magic (and lots of it: say on the order of (at least) millions of years), the rapid rise of language in humans does not allow for this kind of complexity to develop. This suggests that the highly modular structure of GB style theories should be reconsidered.

Fodor (1998) puts the logic nicely:

If the mind is mostly a collection of innate modules, then pretty clearly it must have evolved gradually, under selection pressure. That’s because . . . modules contain lots of specialized information about problem-domains that they compute in. And it really would be a miracle if all those details got into brains via a relatively small, fortuitous alteration of the neurology. To put it the other way around, if adaptationism isn’t true in psychology, it must be that what makes our minds so clever is something pretty general . . .

What holds for the modularity of the mind holds for the modularity of FL as well. A highly modular FL has the sort of complexity that requires adaptation through natural selection to emerge. In addition, adaptation via natural selection takes lots of time. If there is not enough time for natural selection to operate (and 50,000–100,000 years is the blink of an evolutionary eye), then there cannot be adaptation, nor this kind of highly modular complexity. The conclusion, as Fodor notes, is that the system of interest, be it the mind or FL, must be simpler and more general than generally thought.

Lest I be misunderstood, let me make two points immediately.

First, this reasoning, even if sound (and it is important to appreciate how speculative it is given how little we know about such evolutionary matters in the domain of language) does not call into question the idea that FL is a distinct cognitive faculty. What is at issue is not whether FL is modular with respect to other brain faculties. Rather what we are questioning is the internal modular organization of FL itself. The standard view inherited from GB (and

---

9 The assumption that complexity requires natural selection is a standard assumption. For example, Cosmides and Tooby (1992), Dawkins (1996) and Pinker (1997) quoted in Fodor (2000: 87). Dawkins’s words serve to illustrate the general position:

whenever in nature there is a sufficiently powerful illusion of good design for some purpose, natural selection is the only known mechanism that can account for it. (p. 202)

10 Fodor (2000) might not accept this inference as he takes the program in linguistics to only be interested in knowledge not mental mechanisms. I am inclined to think that Fodor is incorrect in his characterization of Chomsky’s position. However, what is relevant here is that grammars are construed as interested in the mechanics of linguistic mentation. The inventory of rules and principles describe real mechanisms of the mind/brain.
I believe still with us today) is that FL itself is composed of many interacting grammatical subsystems with their own organizing principles. For example, the Binding Theory has its proprietary locality conditions (i.e. Binding Domains), its own licensing conditions (i.e. Principles A, B and C), and its own special domain of application (i.e. reflexives, pronouns and R-expressions). So too for Control, Case Theory, Theta Theory, etc. It is this kind of modularity that is suspect as it requires FL to have developed a lot of complicated structure in a rather short period of time both internal to FL itself and internal to each module of FL. If this is not possible because of time constraints, then rich internal modularity is not a property of FL.

Second, I assume that the generalizations and “laws of grammar” that GB discovered are roughly empirically correct. In my opinion, one of the contributions of modern generative grammar to the study of language has been the discovery of the kinds of properties encapsulated in GB. Reconsidering the internal modular structure of GB does not imply rejecting these generalizations. Rather it takes as its research goal to show that these generalizations are the products of more primitive factors. The proposal is to add to the agenda of grammatical theory the aim of deducing these “laws” from more basic principles and primitives.

A picture might be of service here to get the main point across.

(1) Pre-linguistic principles and operations → ?? → (roughly) GB laws

This picture is intended to invoke the more famous one in (2).

(2) Primary Linguistic Data (of L)→ UG → Grammar (of L)

The well-known picture in (2) takes the structure of FL as a black box problem, dubbed “Plato’s Problem” or the logical problem of language acquisition. The goal is to study what UG looks like by constructing systems of principles that can bridge the gap between particular bits of PLD to language particular grammars consistent with that PLD. Generativists discovered that the distance between the two is quite substantial (as the information provided by the PLD significantly underdetermines the properties of the final state of FL) and so

11 The generalizations characteristic of GB have analogues in other generative frameworks such as LFG, GPSG, Tag Grammars, Relational Grammar etc. In fact, I consider it likely that these “frameworks” are notational variants of one another. See Stabler (2007) for some discussion of the inter-translatability of many of these alternatives.

12 There is a term in the physical sciences for the status I propose for GB. The roughly correct theory whose properties are targets for explanation is called an “effective theory.” Being an “effective theory” is already a mark of distinction for to be one, a theory must have good empirical credentials. However, the term also implies that the structural properties of an effective theory need further elucidation and which will come from being subsumed in a more general account. As such, treating GB (and its analogues, cf. note 11) as an effective theory is to at once praise its accomplishments and ask for more theoretical refinement.
1.2 Minimalism and Darwin’s Problem

requires considerable innate mental structure (including the principles of UG) to bridge the gap. GB is one well-articulated proposal for the structure of UG that meets this “poverty of stimulus” concern.

An important feature of the GB model is its intricate internal modularity as well as the linguistically dedicated aspects of its rules and principles. The modules in a GB system are specifically linguistic. By this I mean that their structures reflect the fine details of the linguistic domains that concern them rather than being reflections of more general cognitive mechanisms applied to the specific problems of language. On this conception, FL is a linguistically dedicated system whose basic properties mirror the fine structures of problems peculiar to language; problems related to antecedence, binding, displacement, agreement, case, endocentricity, c-command etc. These latter are specifically linguistic in that they have no obvious analogues in other cognitive domains. It is fair to say that GB is cognitively exceptional in that its principles and operations are cognitively sui generis and very specific to language. In other words, GB endorses the view that FL is cognitively distinctive in that its internal structure displays few analogues with the principles and operations of other cognitive modules. In Chomsky’s (2005a) terminology, GB reflects the view that linguistic competence is replete with first factor kinds of ingredients and that third factor processes are relatively marginal to explaining how it operates.

The picture in (1) is modeled on that in (2). It proposes taking the reasoning deployed in (2) one step further. It relies on the belief that there is an analogy between learning and evolution. In both cases development is partially a function of the environmental input. In both cases it is also partially a function of the prior structure of the developing organism. In both cases the “shaping” effects of the environment on the developmental processes requires reasonable

13 Fodor (1998) characterizes a module as follows:

A module is a more or less autonomous, special purpose, computational system. It’s built to solve a very restricted set of problems, and the information it can use to solve them with is proprietary.

This is a good characterization of GB modules. They are autonomous (e.g. to compute case assignment one can ignore theta roles and similarly licensing binding relations can ignore case and theta properties) and special purpose (e.g. case vs. theta vs. binding). The problems each addresses are very restricted and the concepts proprietary (e.g. binding, control).

14 As Embick and Poeppel (2005a) observe, this is a serious problem for those aiming to find brain correlates for the primitives of FL. They dub this the granularity problem. They propose that one aim of linguistics and neuroscience should be to solve this problem by finding a level that can serve to relate the basic conceptions of each. Their concrete proposal is that an appropriate level of abstraction is the “circuit.” Circuits are brain structures that compute simple operations. The aim is to find those primitive operations that are at once empirically grounded and that could be embodied in neural wet-ware. Given this, the goal for the minimalist will be to find a class of very basic primitive operations that plausibly underlie linguistic computations for consideration as candidates for possible neural circuits.
time during which the environment can “shape” the structures that develop.\textsuperscript{15} (1) takes the evolution of the principles of UG as a function of the pre-linguistic mental state of “humans” and something else (“??”). Moreover, we know whatever “??” is, it must be pretty slight – a new kind of operation or principle – given that FL/UG emerged quite rapidly. We can investigate this process abstractly (let’s call it the logical problem of language evolution or “Darwin’s Problem”) by considering the following question: what must be added to the inventory of pre-linguistic cognitive operations and principles to deduce the principles of UG?\textsuperscript{16} We know that whatever is added, though pretty meager, must be sufficient when combined with the resources of non-specifically linguistic cognition to derive a system with the properties summarized by GB. In other words, what we want is an operation (or two) that once added to more general cognitive resources allows what we know about FL to drop out. On this conception, what is specifically linguistic about FL’s operations and principles is actually rather slight. This is in strong contrast to the underlying ethos of GB, as noted above. The logic of Darwin’s Problem argues against the cognitive exceptionalism of FL. Its basic operations and principles must be largely recruited from those that were pre-linguistically available and that regulate cognition (or computation) in general. FL evolved by packaging these into UG and adding one novel ingredient (or two). This is what the short time frame requires. What (1) assumes is that even a slight addition can be very potent given the right background conditions. The trick is to find some reasonable background operations and principles and a suitable “innovation.”

Once again, the sense of the program is well expressed in Fodor (1998):

\ldots it’s common ground that the evolution of our behavior was mediated by the evolution of our brain. So what matters with regard to the question whether the mind is an adaptation is not how complex our behavior is, but how much you would have to change an ape’s brain to produce the cognitive structure of the human mind\ldots Unlike our minds, our brains are, by any gross measure, very like those of apes. So, it looks as though small alterations of brain structure must have produced very large behavior discontinuities from the ancestral apes to us.

This applies to the emergence of linguistic facility as well, surely the most distinctive behavioral difference between us and our ape ancestors.

Note two more points: First, evolutionary explanations of behavior, as Fodor rightly insists, piggy-back on changes in brain structure. This is why we would like our descriptions to be descriptions (even if abstract) of mechanisms and

\textsuperscript{15} These analogies between learning and evolution have long been recognized. For an early discussion in the context of generative grammar, see Chomsky (1959). As Chomsky’s review makes clear, the analogy between learning and evolution was recognized by Skinner and was a central motivation for his psychological conceptions.

\textsuperscript{16} The term “Darwin’s Problem” is taken from Boeckx (forthcoming).
processes plausibly embodied in brains (see note 14). Second, as Fodor correctly observes, much of this talk is speculative for very little (Fodor thinks “exactly nothing”) is known of how behavior, linguistic or otherwise, supervenes on brain structure. In the domain of language, we know something about how linguistic competence relies on grammatical structure and one aim of the Minimalist Program as I understand it is to investigate how properties of grammars might supervene on more primitive operations and principles that plausibly describe the computational circuitry and wiring that the brain embodies.

Many minimalist proposals can be understood as addressing how to flesh (1) out. Chomsky (2005a) is the prime text for this. As he notes, there are three kinds of principles at work in any specific grammar: (i) the genetic endowment (specific to language), (ii) experience, and (iii) principles that are language or even organism independent. Moreover, the more that any of these can explain a property of grammar, the less explanatory work needs to be done by the others. What modern generative grammar has investigated is the gap between experience and attained linguistic competence. What minimalism is studying is the gap between the third factor noted above (non-specifically linguistic principles and operations) and the first factor (what UG needs that is not already supplied by third factor principles). The short evolutionary time scale, Chomsky (2005a: 3) suggests, implicates a substantial role for principles of the third kind (as do Fodor’s 1998 speculations noted above). The inchoate proposal in (1) is that this problem is fruitfully studied by taking the generalizations unearthed by GB (and its cognates, cf. note 11) as the targets of explanation (i.e. by treating GB as an effective theory).

Before moving on, I would like to emphasize one more point.\(^{17}\) As conceived here, the Minimalist Program is clearly continuous with its GB predecessor in roughly the way that Darwin’s Problem rounds out Plato’s. GB “solves” Plato’s problem in the domain of language by postulating a rich, highly articulated, linguistically specific set of innate principles. If successful, it explains how it is that children are able to acquire their native languages despite the poverty of the linguistic input.\(^{18}\) This kind of answer clearly presupposes that the sorts of mechanisms that GB proposes could have developed in humans. One source of skepticism regarding the generative enterprise is that the structures that UG requires if something like GB is correct could simply not have arisen by standard evolutionary means (e.g. by natural selection given the short time period involved). But if it could not have arisen, then clearly human linguistic facility cannot be explained by invoking such mechanisms. Minimalism takes

---

\(^{17}\) This addition owes a lot to discussions with Paul Pietroski.

\(^{18}\) As the reader no doubt knows, this overstates the case. Principles and Parameters accounts like GB have not yet accounted for how children acquire language. The problem of how parameters are set, for example, is very difficult and as yet unresolved. See Chapter 7 for some additional discussion.
Minimalism and Darwin’s Problem
d this concern to heart. It supposes that FL could arise in humans either by the shaping effects of experience (i.e. through natural selection) or as a by-product of something else, e.g. the addition of new mechanisms to those already extant. For natural selection to operate requires considerable amounts of time. As it appears that FL emerged recently and rapidly as measured in evolutionary time, the first possibility seems to be ruled out. This leaves the “by-product” hypothesis. But a by-product of what? The short time scale suggests that the linguistic specificity of FL as envisaged by GB must be a mirage. FL must be the combination of operations and principles scavenged from cognition and computation in general with possibly small adventitious additions. In other words, despite appearances FL is “almost” the application of general cognitive mechanisms to the problem of language. The “almost” signals the one or two innovations that the 50,000–100,000 year time frame permits. The minimalist hypothesis is that FL is what one gets after adding just a little bit, a new circuit or two, to general principles of cognition and computation. If this is “all” that is distinctive about FL it explains how FL could have rapidly emerged in the species (at least in embryonic form) without the shaping effects of natural selection. The Minimalist project is to flesh this picture out in more concrete terms.

1.3 Two more specific minimalist research projects

To advance this theoretical goal two kinds of projects are currently germane. The first adopts a reductive strategy. Its goal is to reduce the internal modularity of UG by reducing apparently different phenomena to the same operations. This continues the earlier GB efforts of eliminating “constructions” as grammatical primitives by factoring them into their more primitive component parts. Two examples will illustrate the intent.

An important example of reduction is Chomsky’s (1977) proposal in “On wh-movement.” Here Chomsky proposes unifying the various kinds of constructions that display island effects by factoring out a common movement operation involved in each. In particular, Wh-movement, Topicalization, focus-movement, tough-constructions, comparative-formation and Relativization all display island effects in virtue of involving Wh- (or later, A′-) movement

---

19 This way of stating matters does not settle what the mechanism of evolution is. It is compatible with this view that natural selection operated to “select” the one or two innovations that underlie FL. It is also compatible with the position that the distinctive features of FL were not selected for but simply arose (say by random mutation, or as by-products of brain growth). This is not outlandish if what we are talking about is the emergence of one new circuit rather than a highly structured internally modular FL. Of course, once it “emerged” the enormous utility of FL would insure its preservation through natural selection.

20 See Chomsky (1983) for discussion.