Chapter Twelve

What is language like? Moving On

12.1 Where have we gotten?

Three fundamental insights of early generative grammar were (a) that language is to be regarded as a mental capacity, (b) that the creativity or unlimitedness of language is one of its most important features, and (c) that there is a lot more to linguistic structure than is apparent on the surface. For nearly fifty years, these insights, echoing the Cartesian linguists, have guided inquiry into the nature of language. Under their influence, linguistics and the related cognitive sciences have flourished.

The present study, without relinquishing these insights, proposes a substantial revision of the basic organization of language laid out by mainstream generative grammar. Many aspects of our revision have been in currency in one or another of the alternative frameworks such as LFG, HPSG, Cognitive Grammar, Construction Grammar, Role and Reference Grammar, and so on, and we have tried to bring out the similarities when we are aware of them. In fact we have a sense of an implicit consensus among the alternative frameworks – not a monolithic consensus by any means, but one with more of a family resemblance character. We highlight important elements of this consensus in this summary of our conclusions.

A useful starting point in laying out our revision is the status of syntax. From the beginning of generative grammar, syntax was seen as the distinctive characteristic of language – the component that provides language with its creative properties, and the one in which resides the lion's share of complexity and abstractness. It has been in terms of syntax that the central problems of acquisition and innateness have been framed and argued. Phonology and morphology, while recognized as important to language, have decidedly played second fiddle (and viola) to syntax. Insofar as semantics (the mysterious cello?) cannot be coded in syntactic terms, mainstream generative grammar leaves it either to the formal semanticists (who for the most part do not concern themselves with the mental character of language and language acquisition) or to vague invocations of unformalizable “knowledge of the world”, “about which little is known.”

This view of syntax certainly served the field well in its early days. Our sense, though, is that it is time to move on. Syntax, on our view, is simply one component of language among several, each of which contributes to the creativity, complexity, and abstractness of language, and each of which poses problems for acquisition. Syntax is still central “geographically”, in that it serves as a major conduit between semantics and phonology. But it is no longer central
conceptually. It is no longer to be considered a success to show that some bizarre phenomenon can be accounted for by expressing it in terms of a sufficiently abstract theory of syntax. Success lies in showing how the properties of the phenomenon are properly parcelled out among syntax, semantics, phonology, morphology, and the interfaces among them.

The alternative frameworks concur with this judgment in varying degrees. HPSG, Role and Reference Grammar, and some versions of Functional Grammar and Construction Grammar divide the work of building linguistic structure into syntactic and semantic components, and both components contribute liberally to the licensing of phrases and sentences (there is less discussion of phonology). Cognitive Grammar and related versions of Functional Grammar and Construction Grammar also reject the hegemony of syntax, but they downplay its importance in favor of semantics. We find this an overreaction: it is not necessary to reject syntax altogether in order to dethrone it. As we have said at many points in this study, we doubt that semantics does all the work, and we stress that the balance is an empirical matter to be decided by careful weighing of alternatives in each case.

Like most of the alternative frameworks, we have made a major break with the formal tools of mainstream syntax. There are three parts to this break. First, rather than seeing syntactic structure as built up algorithmically through a sequential derivation, we see it as licensed by constraints, all of which can be applied simultaneously. In many cases this approach permits more flexible analyses in terms of “structure-sharing”: a particular element meets disparate constraints simultaneously -- or represents a compromise among conflicting constraints -- instead of having to be put together from disparate abstract parts, each of which embodies a single constraint. This difference has come up, for example, in our analysis of constituent order in the VP (chapter 4), in the interaction of raising and passive in the GF-tier (chapter 4), in our treatment of chains in long-distance dependencies (chapter 6), and in our discussion of simultaneous GF- and CS-binding (chapters 7 and 8).

The second part of this break with mainstream tradition, shared universally with the alternative frameworks, is more radical: the rejection of syntactic movement. Syntactic movement is of course the most prominent formal tool of generative grammar, and the one that was touted in the early days of the field as what most clearly distinguished generative grammar from more traditional “taxonomic” grammar. Hence a rejection of movement has often been seen as repudiating everything that generative grammar holds most dear. “Transformational grammar” and “generative grammar” were taken as more or less coreferential terms, and for this reason many alternative frameworks have been loath to call themselves varieties of generative grammar. Yet in the sense of “generative grammar” originally intended -- namely a formally explicit account of grammar, part of whose brief is to explain the unlimited expressive capacity of
language – all the alternative theories are every bit as much generative grammar as the mainstream.

In the present study we have stressed that the choice between movement and non-movement accounts is an empirical one, not just a matter of preference or ideology. In one case after another, we have counterposed syntactic analyses in terms of movement with alternative accounts in terms of a combination of phrase structure, semantic structure, and the syntax-semantics interface. In every case we have found the movement approach insufficient. We have proposed a variety of formal mechanisms that account for the “illusion of movement”, including alternative linearizations of constituents (chapter 4), manipulations of the GF-tier (chapter 4), indirect licensing (chapter 5), and direct generation of chains and signatures of long-distance dependencies (chapters 6 and 11). There is no rearrangement of constituents, only arrangement.

One of the most seductive features of syntactic movement is that it makes it possible to posit a “hidden level” of syntactic structure in terms of which syntactic regularities can be coded. In the Aspects theory this level was Deep Structure; in GB there were two such levels, D-Structure and LF; in the Minimalist Program there is LF. The promise held out by these theories was that the hidden level of syntax could provide a direct mapping into meaning – what we have called here Interface Uniformity. Thus “deep” syntax could provide a key to the structure of thought.

Without movement, of course, such hidden levels of syntax are impossible, and with their banishment goes all hope of a simple mapping from syntax to meaning. Again the issue is not a matter of preference or dogma, it is empirical. Throughout this study, especially in chapters 2, 3, 5, and 6, we have traced how the logic of Interface Uniformity and the tools of movement have led mainstream theory into ever more abstract syntactic structures and ever more elaborate derivations, aided and abetted by an aesthetic that considers such results to be a conceptual advance. What we (along with the alternative theories) have shown is that these results are often brittle: they do not generalize as they should to the full range of evidence. A prime case studied here was the treatment of Sluicing and Sluice-Stranding (chapters 1 and 5), which look like and are interpreted like elliptical wh-questions, which therefore by Interface Uniformity must be wh-questions in all respects, and whose violation of extraction constraints is therefore an unexplained anomaly. We have shown here that the most insightful generalizations for these constructions are with Bare Argument Ellipsis, for which an account in terms of Interface Uniformity is unworkable once one considers the full range of data. Likewise, we have spent considerable effort (chapters 7-11) showing that anaphoric binding and control cannot receive an insightful account in terms of syntactic structure alone. And we have shown (chapters 8 and 10) that covert movement in LF, originally developed to syntacticize quantifier scope, cannot account for the full range of phenomena, because quantifier scope depends
on lexical semantics, on information structure, and on semantic relations of
subordination that cannot be insightfully coded in syntactic terms.

Interface Uniformity has a methodological advantage: you don't have to
know anything about the structure of meaning in order to do syntax. Rather, every
semantic judgment simply is rendered into “hidden” syntactic terms by one
stratagem or another, and then the principles of movement are adjusted so as to
allow a derivation to the surface form. By contrast, a theory that eschews
movement lacks this shortcut and has to make concrete proposals about the
structure of meaning. In the early days of generative syntax, not enough was
known about meaning for this to be possible, so it was necessary basically to
handwave about meaning. For this reason, the “Interpretive Semantics” of the late
1960s and early 1970s never took off. However, 35 years later, proposals about
semantic structure have a great deal more independent motivation and the tools of
semantic analysis have a great deal more precision. Hence it is now possible to
use semantic analyses to ground syntactic investigation.

The theory of Conceptual Structure used here is only one of a number of
possibilities available for this purpose. However, we think it is ideally suited for
investigating the extent to which the syntax-semantics interface is uniform and
the extent to which it is littered with mismatches. In particular, because it assumes
a continuity between the generative mechanisms behind word meanings and
behind phrasal composition, it allows us to explicitly show the interplay between
syntactic phenomena and details of word meaning, as we did especially in our
studies of argument structure (chapter 4), prepositional passives (chapter 4),
binding (chapters 7 and 8), and control (chapter 9). In this respect it differs from
most versions of formal semantics and concurs with the spirit of treatments in
Cognitive Grammar.

We conclude, then, that the notion of syntactic movement can no longer
be considered the most promising formal device for revealing how language
works. Fifty years ago there was no coherent notion of meaning that could
function as the necessary hidden level, and Deep Structure looked like a suitable
proxy. Movement and deletion therefore were an altogether natural way to
simulate the flexibility of the syntax-semantics interface within syntax alone.
Indeed, had transformational grammar not been invented, the field would have
had no tools for clearing away the debris of syntax-semantics mismatches pointed
out for centuries by philosophers of language, and, crucially, there would have
been no way to get at the underlying regularities of grammar. Transformational
grammar has been a phenomenally effective vehicle for motivating the explosion
in research on language of the past half century – not only theoretical syntax but
also phonology, semantics, psycholinguistics, neurolinguistics, and first and
second language acquisition. We think, though, that the time has come for the
field to find ways to live without syntactic movement as its most productive
metaphor.
There is a third way in which we (and some of the alternative frameworks, especially Construction Grammar and HPSG) have broken with mainstream generative grammar. This one is more subtle, but perhaps even more consequential than the rejection of movement. It concerns a feature of traditional grammar that has been preserved throughout all incarnations of mainstream generative grammar: the strict distinction between grammar and lexicon. Chomsky 1965 (p. 214, n. 16) approvingly cites Bloomfield 1933 (p. 274): “The lexicon is really an appendix of the grammar, a list of basic irregularities” and Sweet 1913 (p. 31): “grammar deals with the general facts of language, lexicology with the special facts.” As far as we can tell, this conception has persisted throughout the history of mainstream generative grammar. The emphasis has been on finding ever larger regularities in grammar, to the point that the “essence” of grammar has been distilled down in the Minimalist Program to Merge and Move, or perhaps only to Internal and External Merge.

In the course of this distillation of the grammar, the constructions of traditional grammar such as the passive and wh-questions have been subsumed into the larger generalizations of movement. This has been considered a virtue, as it removes more of the superficial idiosyncrasy of language. Chomsky has asserted (e.g. Chomsky 2002: 95, 105) that this is a greater advance over early transformational grammar than the latter was over traditional grammar. Yet much of the fine detail of the traditional constructions has ceased to garner attention – recall our discussion of the passive in chapter 2, for instance. The appeal to parameterization has accounted for some of the cross-linguistic differences; but as we have seen, the intricacies in the “syntactic nuts” of English defy plausible finite and innate parameterization. The loss of empirical coverage has also been rhetorically justified by appeal to an ever-shifting distinction between “core” and “periphery,” the latter considered (at least for the foreseeable future) of comparably little interest in comparison to the real quarry, the maximally general and perhaps “perfect” core.

The alternative frameworks will have none of this (and we share this sense). They continue to value the idiosyncrasies of language, often in the interest of demonstrating the effects of semantics and pragmatics on syntactic structure, but often in the interest of just seeing how a language works down to the last little detail. A particular focus has been the behavior of individual words in governing the structures in which they reside, for example how the intricate specifics of argument structure are coded in the lexicon and implemented in phrase structure. It has turned out that constraint-based formalisms lend themselves readily to encoding this material. This is because, within these formalisms, a word can carry pieces of contextual structure in its lexical entry (what we have called here contextual variables), and these can be used directly to constrain or even build the context in which the word appears. By contrast, the mainstream derivationally-
based formalism treats a lexical item as completely self-contained, and it interacts with its context only by such formal devices as moving to check features.

But now there is a striking consequence, which various constraint-based approaches have come upon independently. Once one sees a word as carrying external structure with it, it is a short step to constructing lexical entries for idioms, even discontinuous idioms, in the same format. And from there it is another short step to seeing complex syntactic structures as constructions; that is, as abstract idioms that may or may not carry phonological or morphological markers. Like other idioms, they may carry meaning – or (at least in our approach but not in some others) they may be purely structural. Thus the traditional strict divide between grammar and lexicon has been breached, and replaced with a multi-dimensional continuum between the idiosyncrasy of individual words and general principles such as X-bar theory, verb position, and the possibility of long-distance dependencies. We brought this point out in chapter 1, and it has been a theme throughout, especially in our discussion of argument structure, (chapter 4), VP constructions (chapter 4), noncanonical utterance forms (chapter 5), ellipsis constructions (chapter 5), long-distance dependency constructions (chapter 6), coercions (chapters 7 and 9), and paratactic constructions (chapters 10 and 11).

We believe that this emerging "constructionist" view represents a deep and true discovery about the nature of language. Like traditional grammar but unlike mainstream generative grammar, it acknowledges and captures the rich interaction of semantics with the use of constructions. And, as we have mentioned in several places, it has deep consequences for the problem of acquisition, for now the acquisition of rules can be seen as a principled extension of the acquisition of words, guided by generalizations over the contextual restrictions of individual words. The mystery of classical generative grammar was how the child could acquire all the strange and idiosyncratic rules. Principles and Parameters theory cut the Gordian knot by advocating very general universal principles of structure and movement, differentiated by parameters. This was taken to be a major advance. However, the constructionist approach offers a different and potentially better way to cut the knot, namely by finding a way to make rules learnable – without relegating any of their idiosyncrasy to an ill-defined periphery.

We find it significant that this discovery became possible in large part because of the adoption of constraint-based formalism, whose affordances for representation lend themselves to discovering and formalizing the continuum between words and rules. Thus, as stressed in the early foundational work in generative syntax and phonology, the choice of formalism plays an important role in theory development – in what it leads the theorist to regard as natural and unnatural. The wrong formalism can hobble development. Again, we think it is time to stop thinking in terms of derivations – a conceptual holdover both from the traditional theory of computation in algorithmic Turing-machine-like terms,
and also from traditional morphology, which thinks in terms of complex words being “derived” by putting together smaller morphemes.

We stress that none of this makes the overarching issues of mental representation, creativity, and acquisition go away. They are the same problems as ever, just approached with a different technology that affords different sorts of insights – insights that we judge to bring us closer than mainstream generative grammar to what language is really like.

12.2 Special properties of the present approach

Within this broad constructionist view of language, the approach we have developed here has a number of special properties. Like everyone else, we have phrase structure built out of lexical and phrasal categories that observe some version of X’ theory. Like many of the alternative frameworks, our phrase structure is relatively flat and multiple branching. In particular, phrasal categories are only two nodes deep: NP immediately dominates N, AP immediately dominates A, and so on. The only exception is S, which dominates VP, which dominates V. There is no “Chomsky-adjunction”; there are no phrasal nodes projected from functional categories, especially covert functional categories; and S=IP=CP (i.e. no separate complementizer phrase). Chapter 3 argued extensively for this approach in contrast to the articulated Spec and Comp levels of mainstream theory, and especially in contrast to the strictly binary branching favored over the last decade.

In addition, like the alternative frameworks, we have strenuously avoided the use of covert syntactic elements that are present only to satisfy semantic constraints or to guide the organization of overt elements. We have argued against such null elements in ellipsis (chapters 3 and 5), coercion (chapters 4, 7, and 9), operators in long-distance dependencies (chapter 6), binding (chapters 7 and 8), and control (chapter 9), showing not only that one can do without them, but that their use leads ultimately to impossible complications. The only null elements we have retained from mainstream generative grammar are the traces of long-distance dependencies. Such dependencies are a crucial defining characteristic of human language, and traces are a convenient formal device for characterizing them.

The upshot is Simple(r) Syntax, a syntactic component that does maintain the standard insights about phrase structure but lacks movement, null elements, projections of functional categories, and all other excess nodes. By virtue of its simplicity and particularly its overtness, it presents fewer problems for acquisition that force one into positing complexity in the innate genetic component of language. Thus it also sharpens and simplifies the problems of the evolution of the language capacity – though it hardly makes them go away.
Whereas our constructionist position brings us closest in many respects to HPSG and Construction Grammar, we differ from them in our views on the autonomy of syntax. HPSG stresses the Saussurean sign—a complex of phonological, syntactic, and semantic features—as the fundamental building block of linguistic structure, and Construction Grammar stresses the meaningfulness of all constructions. We have taken a more flexible view, allowing lexical entries to constrain/license only one or two of the parallel structures of language. Thus not every construction involves phonology (e.g. coercions, raising, the sound+motion construction); some involve the interaction of prosody and information structure (e.g. marking focus by intonation); some have phonology and syntax but no semantics (pleonastic reflexives, default of in NPs); some involve the interaction of syntactic ordering and prosody (Heavy Shift). In short, we have found it useful to take the Saussurean sign as simply the stereotypical case of a lexical item, and to allow a broader range of possibilities. Among these possibilities are constructions that involve the constituency and/or ordering of syntax alone (e.g. V final or initial in VP). These are the classic locus of autonomous syntax in mainstream generative grammar, and we see no reason to rule them out on general principle (say just because mainstream theory has stressed them too much). As always, we are seeking the right balance of power between components, on empirical grounds.

Any theory of syntax without movement requires a syntax-semantics interface that takes over much of the work that mainstream theory attributes to movement. The minimal conception of such an interface has surface syntactic structure, semantic structure, and a set of interface principles that establishes the correspondence between the two. Among the interface rules, on a constructionist conception, are the words of the language, which establish relatively idiosyncratic correspondences. Beyond the words we seek the most general account possible of phrasal correspondences. In the beginning of chapter 4 we skimmed off the relatively easy cases of head/modifier relations, the treatment of adjuncts, and the order of non-NP arguments. We were then confronted by a tougher problem in the position of NP arguments of verbs. This could not be solved entirely in semantic terms, because dummy arguments behave in many respects just like semantic arguments. We were therefore forced to introduce the Grammatical Function Tier, a small "hidden level" that deals only with the positioning of NP arguments of verbs.

A grammar incorporating a GF-tier is by no means an a priori optimal position. But we observed that every theory of syntax has a counterpart—most notably in LFG and Relational Grammar, where it is the central innovation, but also in the obliqueness hierarchy of HPSG and (for the most unwieldly treatment) abstract Case in mainstream generative grammar. And in each of these theories this little system is responsible for the prominent constructions of raising and passive. In our approach it is also responsible for the core cases of intraclause
Authors’ note: This document is an uncorrected prepublication version of the manuscript of *Simpler Syntax*, by Peter W. Culicover and Ray Jackendoff (Oxford: Oxford University Press. 2005). The actual published book will have a different organization, in that the longer chapters are broken into two chapters. Please do not cite this version, but the published version.

binding, and it differentiates the most rigid cases of reflexives such as direct objects from more flexible cases such as picture-noun constructions and logophoric reflexives.

Given that no one has found a way to avoid these problems, and that everyone has a solution with the same flavor despite radically different formalisms, we are inclined to think that the GF-tier and its counterparts in other theories are also capturing something deep and true about language. It does have the uncomfortable consequence of letting back in some of the classical problems of acquisition: how can the child be expected to figure out that there is a GF-tier, if linguists can't agree on it? Thus we are pressed into the position of claiming there is something innate about it – it is one of those things that has to be carried on the genome. But this is the way empirical science goes. Maybe there will be a deeper explanation for it and maybe there won't. Our conjecture is that it is a late evolutionary add-on to the syntax-semantics interface, hardly inevitable but affording adaptive advantages in efficiency and reliability of communication (see below). At the same time, we have pared its structure down to a bare minimum, compared to the fully fleshed-out f-structure of LFG, so there is less for evolution to account for.

Another important feature of our approach is the introduction of Indirect Licensing into the theory of grammar, with Bare Argument Ellipsis as its simplest exemplar (chapter 5). The basis of Indirect Licensing is the grammatically primitive process of interpreting fragmentary utterances in terms of the pragmatics of the discourse and nonlinguistic context. This is probably something that children do prior to the acquisition of grammar, and that adults do in interpreting these children's utterances. Pidgins too make heavy use of such pragmatic effects. What makes Indirect Licensing grammatical rather than just pragmatic is that it transmits syntactic properties such as case, gender, and syntactic (GF-) binding properties from the discourse antecedent to the fragment, giving rise to connectivity effects. Indirect Licensing is further grammaticized by being incorporated into more complex constructions as diverse as Gapping, Sluicing, VP ellipsis, *ones*-ellipsis, and Left Dislocation.

The most grammaticized instances of Indirect Licensing are those involving long-distance dependencies, where the fragment is the “fronted” phrase and the antecedent or “target” is a trace to its right (chapter 6). Here Indirect Licensing interacts with another important interface device, the binding of an operator-variable configuration in semantics to a chain in syntax. Following especially the approach of HPSG, a chain associated with a particular long-distance dependency construction consists of a trace in a position associated with the bound variable plus a “signature” phrase structure configuration at the position associated with the operator (e.g. a wh-phrase followed immediately by tense in main clause questions, *that* in one variety of relative clause, *what a* in exclamations, and so on). The characteristic island constraints are not products of
movement, as in MGG, but a consequence of the construction of chains in syntax (as in HPSG) plus semantic constraints on operator-variable configurations (as in many different non-mainstream schools of thought).

Finally, we have shown that binding theory has components both in syntax – specifically in the GF-tier – and in semantics. Chapter 4 developed the position that GF-binding is limited to NP arguments of verbs within a single clause; all other cases are CS-binding alone. Chapter 7 developed the position that GF-binding and CS-binding can come into conflict, leading to ungrammatical reflexives in positions where they would be expected – both in cases of coercion and in cases where the semantic violations come from the verb meaning. We do not feel anywhere near solving the problem of binding, but we believe we have established parameters for a proper solution, again following in the footsteps of many others.

12.3 Two aesthetics

We end up with a rather heterogeneous theory. It includes a considerable variety of tools in place of the simplicity of Merge and Move: the machinery of constructions, which is an extension of the machinery for words and idioms; the linearization of hierarchical structure; the “hidden level” of the GF-tier; Indirect Licensing; the linking of operator-variable complexes to chains with all their peculiar syntactic properties. On the other hand, each tool is relatively constrained, and we have managed to do without complex and abstract auxiliary assumptions, including movement and all the constraints on it, LF, null elements, and parameters. And we have been able to maintain a very lean notion of syntactic structure.

We have learned, mostly through the grapevine, that our approach to syntax has met in some quarters with the reaction: “If this approach is right, then syntax isn’t interesting.” We find this reaction disconcerting and feel it calls for some speculative exegesis.

We begin by looking at the rhetoric behind the predicate interesting. It belongs to a class of psychological and evaluative predicates that we mentioned briefly in section 4.6. The inherent semantic argument structure involves two characters: the entity characterized as interesting and the individual in whom this entity evinces interest. The latter can be expressed overtly in Syntax isn’t interesting to me or Syntax doesn’t interest me, or it can be left implicit, in which case the implicit argument is generic, roughly Syntax isn’t interesting to people. When used this latter way, it presents the uninterestingness of syntax as though it is an objective inherent property of syntax, not a property of an interaction between syntax and an observer. Thus the invited inference of …then syntax isn’t interesting is that if you happen to be interested in this approach to syntax, there is something wrong with you. We reject this invited inference and submit that
interestingness is to some degree in the eyes of the observer – baseball cards and poodles interest some people and not others, and the same is true of Simple(r) Syntax.

This reaction of course typically comes from someone committed to the Minimalist Program, which has cultivated a particular aesthetic of what is interesting. The basic question of the Minimalist Program is “How optimal is the design of language?” where “optimal” is defined in terms of providing the “best possible” mapping between sound and meaning. And the conjecture is that if we work hard enough at it, we will find that language is indeed optimal, despite all its apparent imperfections. What constitutes the “best possible mapping”? One that uses the fewest and most general principles and that squeezes out every bit of redundancy.

Chomsky (2002) says that our ability to meaningfully ask this question is another major breakthrough in our understanding of language. He implies that any approach that does not take this question as basic misses the point of the last 50 years of linguistics and, being mired in mere description, simply cannot be compared with the Minimalist Program. This is the sense, we think, in which our minimalist critics find other approaches “not interesting”: only the questions, style of argumentation, and ultimately the style of solution of the MP are “interesting.”

We have several lines of reply. First, theories can and should be compared for their results, whatever their stated goals of explanation. As we observed in chapter 2, the Minimalist Program is notably lacking in empirical coverage compared even to its predecessor, Government-Binding Theory. The response that it is still just a “program”, a way of asking questions, may have been justifiable in 1995, but a decade later it rings hollow.

Second, as we observed in chapter 1, there are other definitions of optimality besides formal elegance and lack of redundancy. To be sure, scientific standards demand predictiveness and generality, and Occam’s Razor should always be on the ready to slice away unneeded machinery. However, another criterion for optimality, the one we have explored here, is minimization of structure, especially covert structure. As we have stressed in the course of this study, this is not always consistent with minimization of principles. Moreover, optimization must be consistent with adequate coverage of the facts; we have been at pains to show that the derivations of late PPT and MP are not up to the task. Here we concur with the many critics of the MP who have come forward especially in recent years (e.g. Johnson and Lappin 1999; Lappin et al., 2000; Newmeyer 2003; Postal 2004; Pullum 1996; Seuren 2004; Pinker and Jackendoff 2004).

Yet another criterion for optimality is effectiveness for actual use by humans as a system for communicating thoughts to each other. Yes, Chomsky has denied that adaptivity for communication has anything to do with the design of
language, but we find this position almost perverse; (see Pinker and Jackendoff 2004 for discussion). A system designed for optimal communication will most likely be full of redundancy in order to enhance reliability; and indeed language features alternative systems for conveying the same thing, such as word order, agreement, and case. Still another criterion for optimality is optimality for acquisition, which again would favor redundancy. And another is optimality for brain storage and processing, about which we know little, but it is a good bet that it again favors redundancy. Thus the design criteria for language are quite different from the austere logical optimality that the Minimalist Program seeks.

A third line of reply to the goals of the Minimalist Program is to acknowledge the importance of asking how optimal or perfect language is, but also to recognize that a language is under very different constraints than a physical system like a galaxy or a molecule. It has to be carried in brains, and its structure has to come from somewhere. The reason there are lots of languages rather than a single Universal Language is that a human language requires too much information to be carried on the genome. In particular, there are too many words for them all to be innate. Or, seen from a different angle, for language to be entirely innate would have required a far more extensive evolutionary process than humans have undergone. The innate language faculty has evolved just enough tricks for learners to be able to pick up the rest relatively quickly, and for the result to be a communication system that does the job more effectively than if the innate parts had not been there.

Languages may have adventitious, nonoptimal properties for a number of reasons. They are products of imperfect transmission through a society over historical time, subject to cultural forces, language contact, and so on. These effects may be “peripheral” in Chomsky’s sense, but the languages still are learned. Moreover, speakers may be tolerant of “imperfections” such as ambiguity, because they have pragmatics to help guide interpretation. A tolerance for ambiguity in turn may make it possible to communicate in less time and with less effort to achieve precision.

A deeper reason for “nonoptimality” in languages is that the learning of languages is guided by brains whose language capacities are the product of adaptation from an ape prototype over evolutionary time. Adaptation may result in more or less optimal structures in response to environmental constraints, for example the streamlined shape of fish and dolphins. But it is hard to argue that, say, the human reproductive tract is “optimal” in any engineering sense, apart from the fact that it is more or less the best achievable compromise, given the form from which it evolved (e.g. the baby’s head has to fit through the mother’s pelvis, often endangering the lives of both).

The genetic component of language might be like that: a collection of innovations, each of which improved the kind of communication system that could be learned through cultural transmission. Such a conception is very much in
line with the Toolkit Hypothesis. Following the spirit of the sequential evolutionary scenario in Jackendoff 2002a: If you start with words alone, it's an improvement to have phrase structure so you can group words reliably. If you have phrase structure, it's an improvement to add the possibility for canonical order, case marking, and agreement so you can tell the semantic roles of each phrase. From there, it's an improvement to add the GF tier, because you have an even more reliable way to mark the roles of NPs, the quintessential arguments. In another, possibly simultaneous line of development, if you have words, it's an improvement to have indirect licensing so you can connect the sense of one utterance to the next. If you have indirect licensing plus syntax, it's an improvement to add syntactic properties to indirect licensing, so as to make elliptical utterances less ambiguous. If you have indirect licensing plus syntax, it's an improvement to use long-distance dependencies with traces, because that frees you up to express both focus and thematic role at the same time by using a chain.

The redundancy of linguistic systems in this light is perhaps not so surprising. We know that the visual system is full of partially redundant tricks for accomplishing the same goal, for example depth and shape perception, and that these tricks can be teased apart by experiment. The language capacity has done the experiment for us by giving us languages with an abundance of morphology and without, languages with tone and without, languages with overt long-distance dependencies and languages without—and they all permit us to say the same things.

We see no inevitability in this series of developments, especially its later stages. If language didn't have a GF-tier, would it be less perfect, less optimal? Is there another way that the same or greater advantage might have been gained? If languages had a GF-tier but no GF-binding, would that be better or worse? We can't say. But perhaps, following the Minimalist mode of inquiry, these questions are worth at least keeping in mind.

Is this bad? It is not the beautiful elegant "Galilean" outcome envisioned by Chomsky. We agree that a simple Grand Unified Theory of language would be an amazing scientific achievement. On the other hand, given that language is situated in the brain, itself a product of evolution, we find ourselves willing to back off from a Grand Unified Theory and to acknowledge that language is as riddled with eccentricities as every other part of the biological world, arising through contingencies of evolutionary history. This leads to a different aesthetic for the form the theory should take. We submit that, along with derivations, movement, and the lexicon-grammar distinction, it's also time to give up on the Galilean aesthetic and learn to live with this new one.

Of course, good science still demands that we try to reduce the internal resources for learning language to a minimum, with a minimum of eccentricity. But if at some point we have tried in many possible ways, and we keep coming up with the same somewhat eccentric resources, just reconceived in different
technologies, it is scientifically most responsible to recognize this clearly as the best tentative answer, and not try to hide it in the inner workings of a putatively Galilean system. This is what we have tried to do here. Taking such a tack does not preclude doing one's best to make each component and each interaction among components maximally general and beautiful — again, this is just good science. But good science also involves judging that the situation is not ideally beautiful, and then searching for the most productive compromise. In our experience, making such judgments and thereby deciding to break with tradition are among the most difficult choices in one's professional life, taking on an almost moral character. Through the present work we hope to have laid down some new paths to explore and to draw attention to some relatively neglected ones. If we have also encouraged others to create still more new paths, that would be a bonus.

12.4 Where do we go from here?

The present study, despite its undoubtedly excessive length (whew!), is only a beginning at reconceiving syntax and its interfaces with the rest of grammar. We have drawn on everything we know from nearly forty years each in the field, but the full enterprise is way beyond the capacities of two guys who are nearing retirement. Here is a wish list of directions in which we could envision research proceeding. The list is hardly exhaustive.

- We have looked at many phenomena on which there is a vast literature that goes far beyond our discussion. To what extent can the results of other frameworks be translated into ours, and when there are differences, what is the best way to resolve them?
- We have from time to time mentioned phenomena in other languages (and in English) without being able to provide a serious analysis. How does the present framework fare with such well-known syntactic phenomena as agreement within phrases, clitic and contrastive left dislocation, clitic climbing, coherent infinitives, ergativity, floating quantifiers, internally headed relatives, pro-drop, quirky case, scrambling, small clauses, subject-aux/verb agreement, superiority, unaccusativity, verb second, verbless sentences, weak crossover?
- We have looked not at all at morphology. How does it integrate into the system? What are the implications for the structure of the lexicon?
- We have repeatedly brought up prosody as a force on syntactic structure. How influential is it? (Steedman 1996 and Hartmann 2000 come to mind as the sort of work we would welcome.)
- We have sketched a ramified space of possibilities for binding theory but hardly have made concrete proposals for the licensing conditions for
binding. Do the possibilities we have laid out make it possible to make further progress on an adequate binding theory?

- We have speculated on the possibility of unifying the treatment of quantification, anaphora, and the variables of long-distance dependencies through a uniform notation for binding. Can this be carried through and what does that tell us about the nature of language? We would guess that the answer would come through a more fundamental treatment of referential dependencies, via some variant of Discourse Representation Theory or the referential tier of CS in Jackendoff 2002a.

- We have shown repeatedly that syntactic distribution is driven by the lexical semantics of heads and the combinatorial semantics of phrases. The current theory of conceptual structure has taken us a certain distance, but there are many important words and constructions whose semantics we do not understand at all. (The Experiencer verbs, which have played a role here and there in our discussion, are perhaps at the cusp of our understanding.) In order to go deeper, we also need a far subtler theory of semantic combination, including coercion, integration of context, and cocomposition in the sense of Pustejovsky 1995.

- How does Indirect Licensing work? We have suggested that it involves principles of conversational coherence, which take us still further into the details of semantics and pragmatics.

- We have not explored the implications for processing and for acquisition here, although Jackendoff 2002a and Culicover and Nowak 2003 have suggested some plausible directions. However, we suspect that our approach ought to lend itself to experimental investigation, which we would welcome.

On another plane, our wish list would include some change in the way people do syntax. Above all, we think we have shown that syntax cannot be studied without simultaneously studying its interaction with semantics (not to mention prosody) It is true that semantics is hard and much of it is ill-understood, and the way has often been obscured by philosophers who have quite different goals in mind than linguists (see Jackendoff 2002a, chapters 9-10). Syntax seems so much more concrete, tractable, like a hard science. But science is not about playing it safe.

We have suggested as well that an adequate theory of syntax should connect in a natural way to an account of how humans produce and understand sentences. It has been a dogma in the field that a syntactic theory should not be viewed as a theory of how language is processed, and that a theory of linguistic competence should be free of performance considerations. But we believe that it is essential to understand the role that processing plays in explaining the form of language and the character of speakers’ judgments. We do not want to substitute a competing dogma that all of the properties of a language can be explained in
terms of how it is processed. But not to look into processing as a possible source of explanation is short-sighted, to say the least. The interaction between syntactic theory and psycho-/neurolinguistics should be a two-way street.

Finally, we think there must be more of an effort to transcend the technology. The standard genre for papers in syntax gives an account of some phenomenon within a particular framework, and perhaps suggests some tweaks within the framework. We would like to see more investigation that compares frameworks dispassionately, for it is only by doing such comparisons that we pit them against each other scientifically rather than merely sociologically. In particular, most of the alternative frameworks conceive of themselves primarily in opposition to mainstream generative grammar and not in comparison to one another – despite the evident conceptual affinities among them that we have reviewed above. The goal of such comparison would be to distill out of each framework the essence of what it thinks language is like, free of the peculiarities of its formalism. Again, this is hard: it's tough enough to master one framework, much less know two or more well enough to do a comparison. But it can be done, especially in the context of collaboration. Given the fragmentation of frameworks that prevails today, such integration is crucial for the future of the field. Again, we hope the present work has helped set the stage.
Notes

1 HPSG admits such possibilities as well, by leaving one or more components of a sign unspecified. We are uncomfortable less with the formalism than with the emphasis on the sign. For instance, it is hard to think of a coercion as a sign – if anything, a coercion licenses the absence of a sign. Similarly, it is hard to think of a rule such as V-final in VP or a preference for heavy prosodic constituents at the end as signs, since they do not signify anything; they only linearize the other signs. Yet all of these are natural constraints within the formalism of the parallel architecture.

2 This point is missed by Baker (2001), who sees linguistic diversity as socially adaptive for group identification. He therefore conjectures that evolution selected for the system of innate parameters in the service of increasing linguistic diversity. Our argument is that it would be impossible to select for linguistic uniformity, because the genome isn't big enough and evolution hasn't had enough time. Linguistic diversity is an inevitable consequence of the insufficiency of the genome. The use of language as a badge of group membership is a natural exaptation: people use anything they can (even badges) as badges of group membership.