PHILOSOPHICAL PSYCHOLOGY, VOL. 16, NO. 2, 2003

Is C_{HL} linguistically specific?

NIRMALANGSHU MUKHERJI

ABSTRACT *C_{HL} is Noam Chomsky's shorthand for "Single Computational System of human language." C_{HL} is that part of the faculty of language (FL) that integrates lexical information to form linguistic expressions at the interfaces where language interacts with other cognitive systems. In this paper, I am asking whether the elements of FL are dedicated to language alone, or whether significant parts of FL might apply beyond language. From a close examination of the properties of the principles of C_{HL}, I argue that they might well apply to a class of natural symbol systems that includes language and other cognitive systems. This issue of linguistic specificity differs from a similar issue raised recently by Chomsky. For Chomsky, while the "elements" of the linguistic system* per se *are drawn from all over nature, general principles of computational efficiency control the operation of the system. Currently, there is little empirical motivation for this vast generalization to all of nature. The more restricted generalization proposed here looks better suited to the current form of inquiry on language and related system.*

1. The issue

According to Chomsky, Universal Grammar (UG) postulates the following provisions of the faculty of language (FL) that enter into the acquisition of language (Chomsky, 2000b, 2000c):

(1) a set of features,
(2) principles for assembling features into lexical items, and
(3) operations that apply successively to form syntactic objects of greater complexity.

C_{HL} incorporates (3) in that it integrates lexical information to form linguistic expressions at the interfaces where language interacts with other cognitive systems of the mind: roughly, the sensori-motor systems access representations of sound (Phonetic Form, PF), and the conceptual–intentional systems access representations of "meaning" (Logical Form, LF). It is generally assumed that provisions (1)–(3) are specific to language. The GLOW manifesto—which represents the guiding spirit and motivation of current linguistic work—states explicitly that, "It appears quite likely that the system of mechanisms and principles put to work in the acquisition of the

Nirmalangshu Mukherji, Department of Philosophy, Delhi University, 5/21 University Road, Delhi 1100007, INDIA, email: somanshu@bol.net.in

ISSN 0951-5089/print/ISSN 1465-394X/online/03/020289–20
DOI: 10.1080/09515080320103364

knowledge of language will turn out to be a highly specific language faculty" (Koster *et al.*, 1993, p. 342) [1].

GLOW thus applies to (1)–(3). In general, Chomsky (1986, xxvii) has consistently held that, even if the "approaches" pursued in linguistic theory may be extended to study other cognitive systems, the *principles* postulated by the theory are likely to be specific to language. I will assume that (1) and (2), which concern lexical organization, are specific to FL. The point of interest here is that the idea of linguistic specificity is advanced for (3) as well, i.e., for the principles and operations that constitute C_{HL}. Although there has been significant progress in recent decades on discovering principles of lexical organization, linguistic theory has been primarily concerned with the properties of C_{HL}. Thus, the main thrust of GLOW is on the linguistic specificity of C_{HL}.

Nonetheless, I am asking whether the elements of FL are dedicated to language alone, or whether there is some motivation for thinking that significant parts of FL might apply beyond language. I am suggesting that the most reasonable way to pursue this motivation, if at all, is to focus on the combinatorial part of the system to ask whether some of the central principles and operations of this part could be used for other cognitive functions. For the purposes of this paper, therefore, the term C_{HL} is taken to be a rigid designator that picks out a certain class of computational principles and operations, notwithstanding the built-in qualification regarding human language. However, the concept of C_{HL} is also viewed as restricted to language and some other *human* cognitive systems, especially those which may be viewed as "language-like." In this formulation of the issue, the human-specificity of these systems is not denied although the domain-specificity of some of the central organizing principles of these systems is questioned.

The formulation arises out of the fact that, besides language, there are many other cognitive systems in which combinatorial principles play a central role: arithmetic, geometry, music, logical thinking, interpretation of syntactic trees, maps and other graphical representations, to name a few. If the elements of FL are to be used elsewhere at all, it is quite likely that they reappear in some of these systems; that is the step of generalization I have in mind [2]. We will see later that some of these assorted systems fall into a coherent class around language.

To my knowledge, there has been little discussion on whether and to what extent the principles *actually* discovered in the study of language can be extended to other cognitive systems. Clearly, the issue under discussion here arises only for those cognitive systems for which a fairly abstract body of principles is already in hand. In other words, if the principles postulated for a cognitive system are too directly tied to the phenomena they cover, then their very form will resist generalization across phenomenal domains. For example, questions of generalization could not have been interestingly asked for the system of rules discussed in the *Aspects* model of language (Chomsky, 1965). For the cognitive systems under consideration here, it is generally acknowledged that a sufficiently abstract formulation has been reached, if at all, only for a few cognitive systems including language, and that too *very* recently.

These systems, perhaps, include parts of the visual system as well. However, as Chomsky (1980, 1988) suggested, the visual system does not have anything like principles of binding and head parameter, which are basic properties of the language system; the language system does not have the rigidity principle, which is a basic property of the visual system. Since these two systems are not likely to form a family, issues of generalization do not interestingly arise. This is not to deny that they interact: we can report what we see. Thus, given the lack of sufficient advance in studies on other "language-like" cognitive systems, the question that concerns us here has not been routinely asked.

2. Chomsky on linguistic specificity

The issue of generalization to other systems of the mind differs from a similar issue raised recently by Chomsky. Chomsky (2000a, fn. 6; 2000b, p. 2; 2001a, fn. 1) has suggested that FL as a whole, i.e., as a system, is dedicated to language, even if some of its elements are "recruited" from, or used for, other functions. Although Chomsky has not given any specific example of these "elements" in the cited work, he does explain the general idea: "Some other organism might, in principle, have the same I-language (= brain state) as Peter, but embedded in performance systems that use it for locomotion" (2000b, p. 27). If such an organism is found, it will illustrate the idea of the same system being used for other (very different) functions by (very different) organisms.

It will follow as well that neither the elements of the system nor the fabricated structure could be viewed as dedicated to language since there is no difference in substance between this system used for locomotion and the human linguistic system. Therefore, FL is to be viewed as dedicated to language only insofar as it is integrated with the specific performance systems such that the functioning of the *integrated* system results in articulation and interpretation of speech, rather than in locomotion. This striking idea is naturally motivated within the program of what Chomsky is calling *biolinguistics*. Chomsky is suggesting that, since FL is a product of evolution, it will be a deficiency in the program if it postulates principles solely for linguistic explanations without anchoring them to the (possible) workings of organic systems; hence, the task is to eliminate such linguistically specific principles, if any, from the program.

Chomsky (2001a) has also suggested that the linguistic system *per se* is to be distinguished from what he has called the "third factor," which consists of general properties of organic systems that endow efficiency to the computational systems employed by organisms. So, if we find such computational principles while studying a cognitive system, we ought to conclude that they belong to the third factor, not to the cognitive system itself. Again, this idea is naturally motivated by Chomsky's adoption of the "Galilean style," which views nature as perfect: natural systems work under principles of least effort (Chomsky, 1995, 2001a). Accordingly, the general principles of computational efficiency are now viewed as instantiating the concept of least effort, perhaps for the entire class of organisms if the study of organisms is to fall under the Galilean style. In fact, Chomsky and others argue that third factor principles are best viewed as *physical* principles, on a par with such principles as least-energy requirement or minimal "wire-length" (Chomsky, 2002;

Chomsky, 2001a,b). Combining the twin ideas of biolinguistics and perfection, the net picture is as follows: while the elements of the linguistic system *per se* are drawn from the organic part of nature for the fabrication of the computational system, third-factor principles, drawn from the rest of nature, control the operations of the system.

There is no tension between this view and the view advocated in this paper. For Chomsky, elements of the linguistic system, including those of the computational system, are scattered across cognitive systems of organisms (Hauser *et al.*, 2002). My restricted interest is to see if parts of the computational system of language are located in some human non-linguistic cognitive systems. In fact, my view is sharpened by the possibility that elements of the (wider) linguistic system, except for the computational system, may be found in non-human organisms. So, Chomsky's view trivially contains mine, and there is nothing in my view which rules out Chomsky's. In this sense, the two views are likely to supplement each other. However, it is unclear if the stage is set for implementing Chomsky's view.

As noted, the basic thrust of Chomsky's view is to adopt both the notions of biolinguistics and perfection: the notion of biolinguistics places linguistic inquiry within the study of organic systems; the notion of perfection places that study within the form of scientific inquiry illustrated in post-Galilean physics (Chomsky, 1980, pp. 8–9; Chomsky, 2002). Hence, biological systems are viewed as perfect, in contrast to the standard view that these systems are "strange and messy." Although this alternative view of biological systems is beginning to gain some acceptance (Jenkins, 2000; Leiber, 2001), the view is at least severely controversial. Chomsky himself held, not too long ago, that, "Biology and the brain sciences, which, as currently understood, do not provide any basis for what appear to be fairly well-established conclusions about language" (1995, pp. 1–2; I will presently mention some of these "established conclusions"). This is because "biological systems usually are … bad solutions to certain design problems that are posed by nature—the best solution that evolution could achieve under existing circumstances, but perhaps a clumsy and messy solution" (Chomsky, 2000d). Furthermore, Chomsky (1999) suspects "that current understanding falls well short of laying the basis for the unification of the sciences of the brain and higher mental faculties, language among them, and that many surprises may lie along the way to what seems a distant goal."

Notwithstanding the intellectual appeal of Chomsky's proposals, therefore, they remain distant goals given current understanding. In particular, there is a wide gap, currently unbridgeable, between the terms of biology in which its firm results are expressed, and the terms in which the established conclusions about language are stated. For example, we do not know what it means for an insect or a non-human mammal to have a computational system—whose operations generate sound-meaning correlations in humans—that is used for locomotion or for mating display. In fact, we do not know what it means for an insect to have a computational system in the first place in the sense in which the human language system is computational (Gallistel, 1997).

It seems more realistic to concentrate on those parts of nature which are likely to respond to the current form of inquiry on language. The first choice here will naturally be the array of human cognitive systems that interface with articulatory systems and something like "thought"-systems, rather than systems of insect navigation or reptilian mating display. This step will rule out even the human visual system from further consideration since the visual system is a "passive" system. Next, we focus on that subclass of these systems where it may not be immediately implausible to extend what we know to be the established conclusions about language. This will rule out the human (non-linguistic) gesture system although it is an articulatory system, since one of the established conclusions about language is that it is a system of discrete infinity, while the gesture system is not. Further study might rule out all systems except, of course, language; or, it might include other articulatory systems that are systems of discrete infinity. A choice between these options depends on whether there is anything linguistically specific about the principles of the linguistic system such that they operate only on linguistic information; hence the title of this paper.

To anticipate, we may ask if the operations and the principles of linguistic computation, especially those which I call "purely computational principles" (PCPs) below, have anything specifically linguistic in them to bar them from applying to other articulatory systems of discrete infinity. It may well turn out that the class of PCPs constitutes exactly what Chomsky calls the "third factor". In that case, we may proceed further in the direction that Chomsky outlines. But then this direction will be motivated from what we already know about the generality of the linguistic system without stepping out of the current mode of inquiry: formal, top-down, and basically geared to the intuitions of users, rather than to cellular organizations or biochemical processes.

3. Natural symbol systems

The established conclusions on language fall broadly into two groups: (a) general properties that characterize the overall nature of the system, and (b) specific properties that pick out the internal details. In that sense, the property of discrete infinity belongs to the category of general properties; it characterizes the magnitude of the outputs of the system. The operations and principles of C_{HL} belong to the second category.

According to Chomsky, discrete infinity of human language is an unusual property of organisms. However, several human cognitive systems certainly have this property: arithmetic, music, and logical thinking, for example. Following Fodor (2000), Chomsky has stressed another general property of language. As he puts it, "language is different from most other biological systems, including some cognitive systems, in that the physical, external constraints that it has to meet are extremely weak. … The innate system of object recognition … has to be attuned to the outside world; if you had a system that had objects going through barriers and so on, you couldn't get along in the world" (2002, p. 147). No doubt, each of the systems of music, arithmetic, and logic (i.e., natural, "mental" logic) has this property as well;

for example, there doesn't seem to be any control from the world on the human musical system.

This suggests that general properties of language belong to cognitive systems other than language as well; we need not worry about the exact enumeration of the members of this broad class to reach this general conclusion [3]. It is natural to ask: do the combinatorial principles of language, i.e., the specific properties of language, belong to each member of this class as well? At this point, I wish to distinguish this class from other classes of systems whose members have combinatorial properties.

First, no doubt, various natural systems can be described in combinatorial terms: DNA sequences, visual representations, crystal formation, chemical affinities, interaction of particles, and so on. But, none of these objects are "languages" themselves; *we* use symbols to describe their structures. Our understanding of these symbols certainly requires that we have internalized combinatorial principles underlying the use of these symbol systems; but the objects of these symbol systems themselves are not formal objects. The fact that some people can invent them, and that most understand them after suitable training is of great psychological interest. But that interest probably attaches to the study of what Chomsky calls the human "science-forming faculty," not to FL. So, notions like "grammar of vision," "grammar of face recognition," "expression of genes," etc. can only be metaphorical. Second, the symbol systems just mentioned are artificial devices in the sense that they are invented for various special purposes. Importantly, this also applies to musical scores. Most cultures have not developed explicit musical notations, just as many cultures have not developed written language. Even for cultures that have developed musical notations, the understanding and use of these notations is restricted to very few because it takes considerable training and expertise to decipher a musical piece from its score. Sometimes conscious decisions are taken—accompanied by elaborate explanations—to significantly alter a notational scheme to suit specific forms of music. Similar remarks apply to more prominently artificial systems such as programming languages and logistic systems [4].

These considerations suggest the concept of *natural symbol systems* (NSS): "natural" to distinguish them from artificial symbol systems, and "symbol systems" to indicate their formal, articulated nature. Provisionally, language, music, arithmetic, and logic seem to fall in this category; but, as noted, an exact enumeration of cognitive systems is not the immediate task. The concept of NSS gives some rough idea as to how the examination of linguistic specificity of C_{HL} might proceed.

We may think of four kinds of rules and principles that a linguistic theory may postulate. First, the formulation of some rules may be tied to specific languages; call them *language-specific rules* (LSR): relative clauses in Japanese, passivisation in Hindi, and so on. Second, some rules may refer to specific constructions without referring to specific languages; call them *construction-specific rules* (CSR): NP-preposing, VP $\Rightarrow$ V NP, and the like. I am introducing this group for expository purposes. In practice, these rules often refer to language typologies; for example, VP $\Rightarrow$ V NP holds only for head-first languages. It does not affect the discussion that follows. Third, we may have rules that refer neither to specific languages nor to specific constructions, but to general linguistic categories; call them *general linguistic* *principles* (GLP): a lexical item may have a θ-role just in case it has Case, an anaphor must be bound in a local domain, there is a head parameter, and the like. Finally, we may have rules that simply signal combinatorial principles and general principles of interpretation without any specific mention of linguistic categories; call them *purely computational principles* (PCP): all elements in a structure must be interpretable, the shorter of two converging derivations is valid, etc.

It is obvious that if parts of linguistic theory are to apply to NSS at all, only PCPs count. From that point of view, the four rule-kinds basically form two groups: linguistically specific (LSR, CSR, GLP), and linguistically non-specific (PCP). If PCP is empty, then NSS is single-membered. If PCP is non-empty but "poor," then NSS is uninteresting beyond language. Thus, the real question is whether PCP is rich. In other words, how much of the working of C_{HL} can be explained with PCPs alone?

Before we proceed, notice that, from the fact that a computational principle does not mention linguistic categories, it does not necessarily follow that the principle applies to cognitive systems other than language [5]. Nevertheless, two points have emerged. First, it *is* a necessary condition for a principle to apply to varied cognitive systems that it cannot make reference to system-specific categories. Second, *if* a principle is non-specific in the stated sense, and *if* there is a need for a similar principle in another domain (as the general concept of NSS suggests), then it is counter-intuitive that a non-specific principle is geared solely to a specific system. It is more likely that there is a significant generalization here across cognitive systems; otherwise, the non-specific nature of the principle becomes a mystery. However, ultimately, it is a matter of discovery, not stipulation.

Similar remarks apply to Chomsky's idea that certain principles of the linguistic system simply enforce general computational efficiency (see Section 2). Here as well there is a considerable gap between the claim that certain principles of the linguistic system endow computational efficiency to the system and the claim that these principles are general properties of organisms. Nevertheless, following the lead from language research, Chomsky is making the plausible assumption that cognitive systems of organisms are open to investigation in the Galilean style. If so, then we expect economy conditions to be available across the board. The existence of principles without specific linguistic content in the linguistic system surely suggests that these may be the ones that foster economy in the class of systems at issue, even if we do not yet know if they in fact do so. In that sense, both Chomsky and I are offering possible explanations of what these intriguing principles may be doing in cognitive systems. Hopefully, these explanations will converge around a single set of principles.

4. Principle and parameters: G-B

The remarkable thing about current linguistic theory is that there is a real possibility that rules of the first two kinds, viz., LSR and CSR, may be totally absent from linguistic theory. The general theoretical framework that made this vast abstraction possible is called the "principles and parameters framework" (P-P) (Chomsky, 1981). In slightly different terms than mine, Chomsky (1991, pp. 23–24) brings out the basic features of a P-P theory as follows. Consider two properties that descriptive statements about languages might have: a statement may be language-particular or language-invariant [± lp, where + lp = LSR], or, it could be construction-particular or construction-invariant [± cp, where + cp = CSR]. Then, according to Chomsky, a P-P theory contains only general principles of language that are [-lp] and [-cp], and a specification of parameters which is [+ lp] and [-cp]. Traditional grammatical constructions such as active–passive, interrogative and the like are "on a par with such notions as terrestrial animal or large molecule, but are not natural kinds" (1991, pp. 23–24). Once the parameters are set to mark off a particular language, the rest of the properties of the expressions of this language follow from the interaction of language invariant principles: "the property [± cp] disappears."

In our terms, a linguistic theory under the P-P framework postulates just two kinds of principles, GLP and PCP. One of the principal moves that ultimately allowed this abstraction was to shift much of the language-specific information that enters into the properties of expressions to the lexicon itself (Chomsky, 1972). To consider just one simple example among many difficult ones, it was seen that some of the phrase structure rules just repeated lexical properties, e.g., the sub-categorization properties of verbs. The verb (V) *hit* requires a noun-phrase (NP) complement—e.g., *hit the ball.* A language-learner acquires these properties in acquiring the lexicon in any case. To that extent, the phrase structure rule VP ⇒ V NP simply repeats the information. Thus, a natural way to remove the redundancy just noted is to list these properties in the lexicon, and eliminate the corresponding phrase-structure rules from the system (Chomsky, 1986). All that is now needed is to postulate a language-invariant principle—the projection principle—that requires that lexical properties be represented in syntactic structures. Once the computational system was thus freed from language-specific rules, and was replaced with language-invariant principles, it was possible to reduce all transformational operations to just one, viz., move-α, where α is any syntactic category. In effect, the system allowed any syntactic category to be moved anywhere, provided no language-invariant principle is violated.

The form of linguistic theory that emerged from such considerations is justly thought to be a "radical break from the rich tradition of thousands of years of linguistic inquiry," including the early phases of generative grammar (Chomsky, 1995, p. 5). This led naturally to the Galilean style mentioned above. Indeed, the very theme of this paper could not have been seriously contemplated without the P-P framework in hand. However, it is clear that just the framework is not enough for our purposes, since the framework allows both GLP and PCP. Therefore, unless a more abstract scheme is found within the P-P framework in which PCPs at least predominate, no interesting notion of NSS can emerge. The issue obviously is one of grades: the more PCPs there are (and less GLPs) in C_{HL}, the more suited it is for NSS.

The task then is to examine the short internal history of the P-P framework itself to see if a move towards progressively PCP-dominated conceptions of C_{HL} can be discerned. The broad research program under the P-P framework divides roughly into two phases: the earlier phase, usually called "Government Binding (G-B) theory," and the current "Minimalist program" (MP). I will briefly discuss each of these phases from the direction just suggested. I must mention that these are *phases* in which theoretical ideas change constantly. Since I am not concerned with the internal details of shifting research within a phase, what follows are brief descriptions of overall pictures, not expositions of specific theories. Some familiarity with the basic theoretical ideas of these phases will be assumed (see Sells, 1985; Radford, 1997).

The organization of grammar in G-B theory is schematically represented in Figure 1 below. We may think of the picture as representing two basic parts: the lexicon and C_{HL}. Lexical information is mapped onto C_{HL} and computation begins. First, computations progressively yield two *inner* levels of representation: d-structure and s-structure. Then, the computation branches at s-structure to yield two *outer* levels of representation, PF and LF, which feed into (language-external) systems of interpretation. We concentrate on the lexicon to LF computation.

Once a selection of lexical items is mapped onto the computational system in the general format prescribed by X-bar theory, computation proceeds by repeated applications of move-α to satisfy various principles at the levels of representation. The names of these principles are shown in boxes. Some of these apply exclusively at d-structure, s-structure and LF respectively, while others apply at more than one level. This distribution of principles over the system plays a major role in what follows. Recall that the P-P framework postulates just GLPs and PCPs. Let us now see how the principles postulated by G-B theory fall into these classes [6]. The

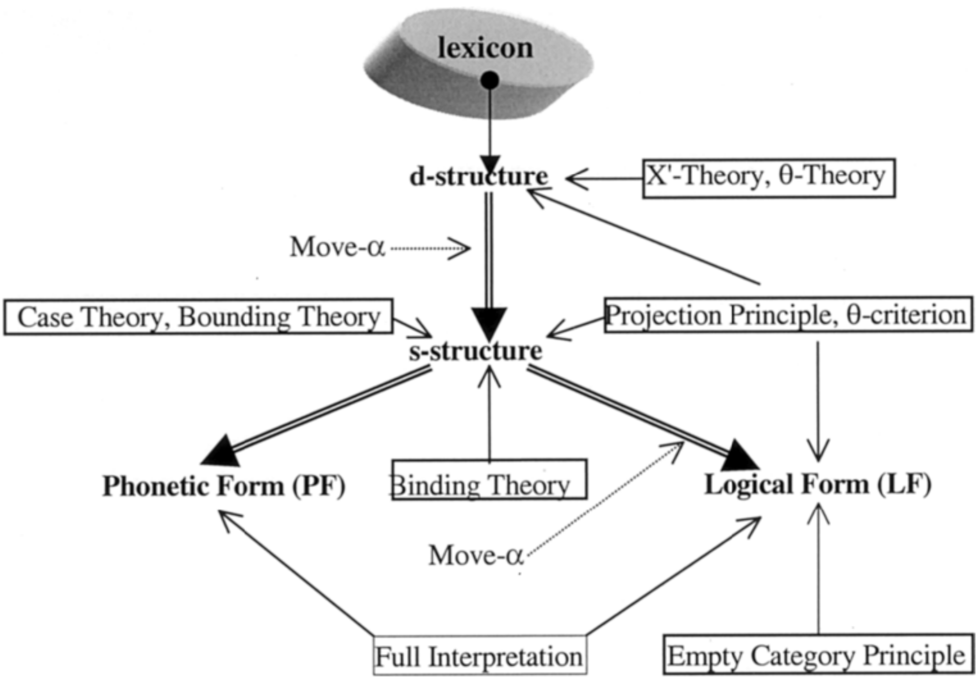

FIG. 1: Government Binding Theory.

classification is going to be slightly arbitrary; we will see that it will not affect the general argument.

The projection principle, as noted, stipulates that lexical information is represented at all syntactic levels. No doubt, it is phrased in terms of lexical properties, but all it does really is to guarantee that input-information not be lost to the system; since the input happens to be lexical information, the principle guarantees that. A cognitive system is designed to solve the perception problem: what is the interpretation of a percept, i.e., what information does it convey? Any computational system that is designed to interact with representations of experience to solve the perception problem requires that none of the representations that encode information are lost to the system until a complete interpretation is reached. To that extent, the broad category of NSS will require the effect of the projection principle, if only implicitly. However, the formulation of the principle in G-B suggests an intermediate category of principles—call it *quasi-PCP*s (Q-PCP): linguistically specific in formulation, but PCP in intent.

X-bar is a universal template, with parametric options, that imposes a certain hierarchy among syntactic categories such as Specifier, Head and Complement within a phrase: [*the* (specifier) [*observation* (head) [*that John is sick* (complement)]]]. Again, it stands to reason that any member of NSS will require some notion of hierarchy if a sequence of its elements is to meet conditions of interpretation; for example, elements of music certainly form tonal and event hierarchies (Lerdahl & Jackendoff, 1983). Still, it is not obvious that every symbol system must have the rather specific hierarchy, as noted, that X-bar theory invokes for each phrase. In that sense, the principle falls somewhere between GLP and Q-PCP. Given the uncertainty, let us assume the worst case that X-bar theory is GLP.

θ-theory seems linguistically specific in that it is exclusively designed to work on s-selectional properties of predicates. The θ-criterion ("each argument must have a θ-role"), the main burden of this theory, is phrased in terms of these properties. But what does the criterion really do, computationally speaking? As Chomsky *et al.* (1982, pp. 85–86) observe, a selection of predicates from the lexicon projects a set of arguments onto syntactic structure. In order to determine precisely the relations between these arguments (in the X-bar format), two kinds of information are needed: an enumeration of arguments, and the order of arguments. Thinking of thematic roles as lexical properties of predicates (e.g., the verb *give* selects an object and a recipient), the θ-criterion checks to see if elements in argument-positions do have this lexical property. In other words, the criterion captures the enumerative part by checking for a close fit between certain positions and a certain lexical property, without caring which argument has which θ-role.

This line of reasoning brings out a fundamental feature of symbol systems. Syntactic operations, being syntactic, are sensitive only to aspects of configuration, such as designated positions and "spatial" relations between them. Input items on the other hand also have properties that are ultimately realized in terms of their semantic significance once a legitimate, interpretable syntactic object has been formed. For the computational system to generate interpretable sequences, therefore, a close fit is needed between aspects of syntactic configuration and lexical properties. To the extent that the θ-criterion accomplishes this task, it is a PCP. Yet, as noted, it is phrased in GLP terms. In my opinion, it ought to be viewed as Q-PCP.

The Case filter ("each lexical NP must have Case")—the main burden of Case theory—is also linguistically specific in exactly the same way: it cannot be phrased independently of linguistically specific properties. Structural Case is assigned during computation, and the system specifically needs this property to activate the Case module. It is hard to imagine that every symbol system must embed a subsystem of Case. Yet, as for the θ-criterion, the Case filter serves a purely computational purpose to check for the ordering part of the set of arguments (*ibid.*); the system does not care which lexical NP has which Case as long as it has a Case. In that sense, it is a Q-PCP as well.

Binding theory explicitly invokes such linguistically specific categories as anaphors (*himself, each other, ...*), pronominals (*them, he, ...*) and r-expressions (*John, Bill, ...*) to encode a variety of dependency relations between NPs. There is a principle for each of these categories. For example, principle A says that anaphors must be bound in a certain narrow domain: *John likes himself, *John thinks that Bill likes himself* (*himself* = *John*). It is hard to make a general computational sense of principles (1)–(3) of Binding theory in non-linguistic terms. In other words, the elements of a symbol system need to have dependency relations of a rather specific sort in order to activate this theory. It is implausible to think of, say, musical quantifiers, anaphors and pronominals, just as it makes no sense to look for subject-object asymmetries in music. Notice the problem is *not* that other symbol systems may lack dependency relations in general; they cannot. The issue is whether they have relations of *this* sort. Similar remarks apply to Empty Category Principle (ECP). These are then GLPs.

This brings us to the principle of Full Interpretation (FI) and Bounding theory. Bounding theory contains the Subjacency principle, which stipulates the legitimate "distance" for each application of move-α. These distances are defined in terms of bounding nodes, which in turn are labelled with names of syntactic categories such as NP or S. In that sense, it is linguistically specific. Abstracting over the particular notion of bounding nodes, it is an *economy principle* that disallows anything but the "shortest move" and, as such, it is not linguistically specific; it is Q-PCP. Finally, the principle of Full Interpretation does not mention linguistic categories at all in stipulating that every element occurring at the levels of interpretation must be interpretable at that level; in other words, all uninterpretable items must be deleted. FI, then, is PCP.

The preceding discussion of the central features of the computational system in G-B is not exhaustive, as noted. Nevertheless, with respect to the issue at hand, some salient points are already available. First, most of the principles cluster at the inner levels of representation: d-structure and s-structure. Second, the principles discussed are a mixed bag of GLPs, Q-PCPs, and PCP; predominantly Q-PCPs, in my opinion. Thus, although PCP is non-empty, it is poor; hence, the system is unsuitable for other members of NSS. But, the predominance of Q-PCPs, and the relatively meagre set of GLPs, suggests that there are large PCP-factors in the system which are concealed under their linguistic guise. If these factors are extracted and explicitly represented in the scheme, G-B theory can turn into one that is more suitable for NSS. I will argue that the scheme currently under investigation in the minimalist program may be profitably viewed in that light.

5. The minimalist program

As noted, the second phase in the P-P framework, the minimalist program, is more directly motivated by the Galilean assumption that nature is perfect. Two basic concepts, *legibility conditions* and *conceptual necessity*, are introduced to capture this assumption. Language as a cognitive system is required to meet two sets of legibility conditions—one each for the phonetic and semantic interfaces, as we saw—such that systems external to language can access the information stored by the language system. For language to be put to use these conditions have to be met; hence these are conceptually necessary.

The Galilean condition is now captured in two steps. First, we stipulate that the theoretical machinery to be used for describing the system be restricted to meet just the conceptually necessary conditions. Second, we show that once these conditions are optimally met, then "sound-meaning relations … will follow. The best theory that considers just satisfaction of the legibility conditions would remain the best theory when you add all the other conditions" (Chomsky, 2000d, p. 20).

The broad picture that emerged has the following basic features (see Figure 2). I have adopted an intermediate position between what was originally proposed (Chomsky, 1995) and what is currently under investigation (Chomsky, 2001a). For example, I have left it open whether there is a uniform computation before the phonetic and semantic computations branch, or whether computations are parallel throughout; similarly for the distinction between overt and covert computations. A decision on these will not affect the discussion that follows. One immediate consequence of pursuing the Galilean steps is that the inner levels of representation of G-B theory, d-structure and s-structure, are no longer motivated since no performance system has access to these levels. We saw that most of the complex array of principles of G-B theory was clustered on these inner levels. Therefore, some drastic reordering is needed to redistribute their effects.

Three basic steps were taken to achieve this. First, the lexicon was further "strengthened" to capture some of this information. For example, Case is now viewed as a universal lexical feature that automatically accompanies nominals. In general, the lexicon is assumed to be a complex system of features, subsets of which cluster in various ways to generate what could be called "lexical items." These features are classified into three broad groups: phonetic, formal, and semantic. As soon as a collection is formed, or soon after, the phonetic features are stripped away to form the phonological object PHON. Language-specific features are assumed to be restricted to the sound part of the lexicon. Hence, differences between languages are captured entirely in PHON computation. The rest of the features proceed to the semantic interface via SEM computation, which is assumed to be invariant across languages. As usual, we focus on this part.

Second, some of the subsystems that are required essentially for interpretation

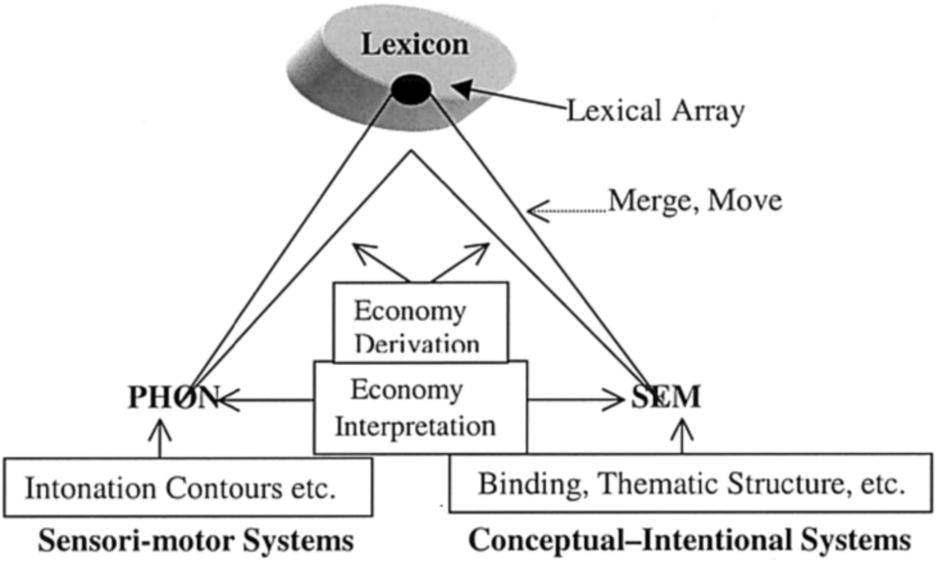

FIG. 2: Minimalist Program.

of syntactic objects are shifted outside of language. As noted, outputs of language (PHON, SEM) are accessed by two broad classes of external systems: sensori-motor systems access PHON, conceptual–intentional systems (C-I) access SEM. The shifting of interpretive principles such as Binding theory, θ-theory, ECP etc. outside of language means that these principles now occupy a location at the "front end" of C-I. If the output of language—SEM—fails to meet these principles, it results in gibberish; but it is a legitimate syntactic object otherwise, provided it meets the principles of C_{HL}.

Third, the computational system obeys the *inclusiveness condition* such that no new objects are added in the course of computation apart from rearrangement of lexical properties; in effect, "any structure formed by the computation … is constituted of elements already present in the lexical items selected" (Chomsky, 1995, p. 228). In particular, X-bar and Binding theory of G-B required that elements such as bar-levels and indices are introduced during computation. Since these are now barred, phrase structure theory needs to be redesigned [7]. Chomsky (1994) proposed a *bare phrase structure* theory, which, among other things, essentially dispensed with X-bar theory. The computational system no longer has a universal template that determines the hierarchy of syntactic elements. Rather, phrase structure is constructed "on-line" as lexical items are individually selected from a lexical array by the operation Select, and put together by the operation Merge. Plainly, both Select and Merge are conceptually necessary. Phrase structure is a consequence of these operations.

The combined effect of these steps makes the computational system itself free of linguistically specific principles and operations (GLPs). Recall that we viewed X-bar theory, Binding theory and ECP as GLPs. We just saw that the C_{HL} in MP does not contain any of them. Further, Projection principle, Subjacency, θ-theory and Case theory were viewed as Q-PCPs. While the Projection principle as such is no longer required in the system, Case theory is basically shifted to the lexicon. The interpretive role of θ-theory is shifted to an external system, as we saw. This leaves Subjacency, a Q-PCP. Q-PCPs are essentially PCPs under linguistic formulation. This raises the possibility, as noted, that just the PCP-factor may be extracted out of them, and explicitly represented in the system. The MP principle *minimal link condition* (MLC) serves that purpose with respect to Subjacency. The actual formulation of this principle is rather technical, but the basic idea, in the spirit of minimalism, is to enforce an *economy condition* that "shorter moves are preferred to longer ones" (Chomsky, 1995, p. 295). Similar remarks apply to the G-B condition on chains (not discussed earlier as it is too technical). This condition is now replaced by an economy condition called the *shortest movement condition* (SMC), which requires that, in case there are competing derivations, the derivation with the least number of steps is chosen (Chomsky, 1995, pp. 130, 182).

Thus, the only G-B principle that is fully retained in MP is Full Interpretation (FI), which, we saw, is a PCP. FI is also naturally viewed as an economy condition that requires that representations at the interfaces are minimal. In sum, in so far as the G-B principles are concerned, all linguistically specific factors have been either removed from the system in MP, or they have been replaced by economy conditions and conceptually necessary operations. As I attempted to show, all the principles and operations of MP have been factored out of those of G-B; that is, no fundamental operation or principle has been added to the system in MP. This applies to the apparently new MP operation Merge as well. Merge, we saw, is the operational part of just the minimal combinatorial aspect of X-bar. Similar remarks apply to various versions of Move, such as Affect and Attract, with respect to the original operation Move-α, in my opinion. The general picture, therefore, is that the C_{HL} in MP is predominantly constituted of PCPs. This applies to (a) the operations of the system such as Select and Merge, (b) constraints on derivation such as MLC and SMC, and (c) constraints on interpretation such as FI. We examine the operation Move in Section 6.

The preceding discussion of MP is not exhaustive. Let us also concede that the rendition of some of the individual principles and operations, regarding the presence or absence of linguistically specific factors in them, could be contentious. Yet, when compared to the G-B framework, the overall picture is unmistakably one of greater generality and abstraction *away* from linguistic specificity and *towards* NSS. Recall that the only issue currently under discussion is whether we can discern a progressively PCP-dominated conception of C_{HL}. In fact, if the preceding analysis is broadly correct, it will need considerable effort to find a *non*-PCP factor in C_{HL}. Supposing, for the sake of argument, that C_{HL} still contains some residual linguistically specific factors, these factors are now naturally viewed as exceptions, and are, therefore, against the direction of progress. Plainly, major theoretical innovations are needed to bridge the current gap between studies on language and other cognitive systems falling under NSS. All we have in hand for now is a possible general picture.

6. Displacement

A variety of objections may be raised against the general picture itself. They fall into two groups: general objections that question the coherence of the picture, and specific objections that concern some individual pieces of the picture. A general objection is that, granting that successive phases of linguistic theory do show a movement from GLPs to PCPs, the link between PCPs and NSS is at best tenuous and at worst metaphorical; even PCPs are to be understood in the context of linguistic explanation.

The objection is trivially true if its aim is to draw attention to a certain practice. There is no doubt that these PCPs were discovered while linguists were looking only at human languages. This paper need not have been written, of course, if someone also discovered them in the course of investigating music or arithmetic. But the future of a theoretical framework need not be permanently tied to the initial object of investigation. As Chomsky observed in the past, a sufficiently abstract study of a single language, say, Hidatsa, can throw light on the entire class of human languages; hence, on FL. This observation cannot be made if it is held that the non-Hidatsa-specific principles that enter into an explanation of Hidatsa cannot be extended to Hindi because Hindi was not in the original agenda.

Nonetheless, the laws and principles postulated by a theory need to be understood in their theoretical context. For example, the notions of action and reaction as they occur in Newton's force-pair law ("every action has an equal and opposite reaction") have application only in the context of physical forces even if the law does not mention any specific system. We cannot extend its application to, say, psychological or social settings, such as two persons shaking hands. Global limits on theoretical contexts, however, do not prevent theoretical frameworks from evolving and enlarging within those limits. The force-pair law does not apply to social situations, but it does apply to a very large range of phenomena, perhaps beyond Newton's original concerns in some cases. For instance, the law has immediate application in static phenomena like friction; but it also applies to dynamic phenomena such as jet propulsion. The theoretical context in which we are asking whether principles of C_{HL} apply to other members of NSS is the general context of the mental. Given the motivation and the theoretical possibilities discussed above, the question is more like asking whether the force-pair law applies to jet propulsion, rather than to people shaking hands. The burden is surely on the linguist now to tell us what exactly the boundaries of the linguistic enterprise are.

A specific objection to the framework proposed here arises as follows. It is a fact about human languages, as contrasted to artificial languages, that an element is often interpreted in a position different from where it is sounded. The elements *John* and *the book* receive identical interpretations in markedly different structures such as *John read the book* and *the book was read by John*. It is the task of a transformational generative grammar to show the exact mechanism by which the element *the book* moves from its original semantic position (as the direct object of *read*) to the front of another structure without altering semantic interpretation. This is just one of many examples falling under the general phenomenon of *displacement*. A basic operation, variously called Move-α or Affect-α in G-B and Move or Attract in MP, implements displacement. We will briefly look at the operation Move in a moment.

Now the objection is that nothing is more linguistically specific than the phenomenon just described. A major part of C_{HL} is geared to facilitate instances of movement in an optimal fashion. Thus, even if the requirement of optimality leads to PCPs, the *reason* why they are there is essentially linguistic. In that sense, the phenomenon of displacement blocks any clear conception of NSS beyond language.

I postponed a discussion of this issue to the end because its early mention would have forestalled any attempt to get at the rest of the general picture. Having allowed a general picture to emerge first, we are in a position to raise the issue of displacement as a specific problem, if at all, within that picture. "If at all" because, as science progresses and attentions are shifted, there is no guarantee that what counted as a fundamental problem at an earlier stage remains so throughout. Initial problems are often posed by pre-theoretical data whose relevance and analysis changes as inquiry becomes more theoretical, and wider perspectives are reached. The process is almost routine in the mature sciences, including the short history of linguistics (Larson & Segal, 1995, pp. 8–9).

Be that as it may in the future, for now there may be ways of accommodating the phenomenon of displacement within the picture itself. To that end, I will outline a number of directions to suggest that the issue of displacement (hopefully) breaks down into research questions that are compatible with the general picture. If the attempt fails, we will be left with a specific problem that we do not know how to handle. This too happens in science all the time.

First, suppose displacement is specific to human languages. In that case, the general picture will not be disturbed if the phenomenon is linked to other linguistically specific aspects of the system. From one direction, that seems to be the case. We saw that the lexicon, which is a collection of features, is certainly linguistically specific in the sense under discussion here. One of the central ideas in MP is that the lexicon contains uninterpretable features such as Case. Since the presence of these features at the interfaces will violate FI, C_{HL} wipes them out during computation. A natural mechanism for that is to let an uninterpretable feature form a local relation with another instance of the feature in the structure. As the instances match, the unwanted feature is cancelled out. When the items to be matched are not originally in a local relation, they need to move to form one, if possible. The operation that executes this complex function is called Move. Details apart, it is clear that Move is activated once uninterpretable features enter C_{HL}; displacement is entirely forced by elements that are linguistically specific [8].

There are several ways of conceptualizing this point within the general picture. If Move is an elementary operation in the system, then we may think of this part of the system as remaining inert until linguistically specific information enters the system. The rest of C_{HL} will still be needed for computing non-linguistic information for Select and Merge to form complex syntactic objects. In effect, only a part of the system will be put to general use and the rest will be reserved for language.

Alternatively, Move may not be viewed as an elementary operation but a special case of existing operations. There is at least one suggestion in which Move is viewed as specialized Merge (Kitahara, 1997) [9]. A compatible evolutionary story could be that the "basic" part evolved first without Move; frills were added later to facilitate the evolution of language. The speculation—nothing more—is that a basic system initially emerged (or was inserted) to implement just the species-specific property of discrete infinity; progressive fine-tuning of the system led to more complex applications [10].

A simpler view will take Move to be Merge itself such that the "frills" are made available as soon as Merge is available. According to Chomsky (2001b), Merge is simply an operation that puts together two elements α and β to form, say, $\{\alpha, \beta\}$. Unless we make the special assumption that α and β are necessarily distinct, β could be a part of α. Since special assumptions are "costly," we allow the latter since it comes "free." In that case, (Internal) Merge can repeatedly put parts together as long as other things are equal. As a result, the original part will appear as copies conjoined to other parts: *the book seems* [*the book*] *to have been stolen* [*the book*]. Here, displacement of *the book* just means that only one of the copies—viz., the left-most one—is sounded for reasons of economy in the phonological component; others are left as covert elements to be interpreted by the C-I systems.

Second, we may ask whether displacement in fact is linguistically specific. We saw a C_{HL}-internal reason for displacement triggered by uninterpretable features. However, there is another reason for displacement. As noted, external systems impose certain conditions on the form of expressions at the interfaces. For example, (efficient) semantic interpretation often requires that items be placed at the edge of a clause to effect a variety of phenomena such as topicalisation, definiteness, and the like. One of the most interesting ideas currently pursued is to see whether these external requirements are satisfied in tandem with the satisfaction of internal requirements (Chomsky, 2002). In other words, the elimination of uninterpretable features takes an element exactly where it receives, say, a definiteness interpretation.

Arguably, some of these external requirements are also imposed on cognitive systems other than language; there could be a need for taking items to the front of a structure for focus, highlight, continuity, and the like. Given that linguistic notions such as topicalisation, definiteness, etc. are viewed as special cases of more general notions such as above, we are now asking whether the external systems that require topicalisation, etc. are themselves linguistically specific. Parts of the C-I system are not even human-specific; hence, not linguistically specific. Higher mammals, for example, seem to possess parts of the conceptual system, including some of the thematic roles (Premack, 1986, p. 16). We just have to know more about the enumeration of external systems and their distribution across species and their cognitive systems.

Suppose that the parts that impose the concerned legibility conditions are not linguistically specific. These parts could be viewed as enforcing conditions on structures that are met in different ways by different members of NSS in terms of the internal resources available there (e.g., language achieves these conditions by drawing on uninterpretable features). This will make the *implementation* of displacement specific to the cognitive system in action; but the *phenomenon* of displacement need not be viewed as specific to any of them.

Acknowledgements

I am indebted to Roberto Casati, Probal Dasgupta, Steven Davis, Pierre Jacob, Adriano Palma, B. N. Patnaik, Jean-Yves Pollock, and, especially, Noam Chomsky for helpful comments on earlier versions and on the issues raised here.

Notes

[1] GLOW stands for Generative Linguists of the Old World, an international organization of linguists based in Europe.

[2] The approach pursued here has some bearing on the modularity issue (Fodor, 1983, 2000; Chomsky, 1984). See Mukherji (2000, pp. 24–32) for an involved discussion of the modularity issue from the suggested direction.

[3] This is not to say that individual cases do not merit further attention. Consider music. First, music, like human language, is a completely universal—species-specific—phenomenon. Infants just a few months old can spot dissonant notes at the end of a melody, and archaeologists have discovered flutes made from animal bones by Neanderthals living in Eastern Europe more than 50,000 years ago (Cromie, 2001). Every culture develops at least vocal music even in adverse environmental conditions. Sami people of Lappland developed extremely complex vocal music although they failed to develop instrumental music because of scarcity of material in such latitudes (Krumhansl *et al.*, 2000). Second, outside of language, music is the only cognitive system for which a fairly detailed generative theory has been proposed on the model of research in generative grammar (Lerdahl & Jackendoff, 1983). Although the theory is focused on Western tonal music, the structures so isolated suggest an underlying basis to musical experience that looks invariant across a vast range of music in that genre. Some parts of the theory were subsequently verified in terms of actual audience response (Jackendoff, 1992; Krumhansl, 1995). Further theoretical and experimental work has now yielded a theory, some of whose predictions can be quantified (Lerdahl, 1996). Third, experimental research on various aspects of music cognition, across a wide spectrum of musical traditions and cultures, are beginning to provide evidence that support "a core set of psychological principles underlying melody formation whose relative weights appear to differ across musical styles" (Krumhansl *et al.*, 2000, pp. 13–14). For example, in the paper just cited, it was reported that in studies on melodic expectancy and tonal hierarchies, considerable agreement was found between listeners from the music's cultural context or from outside it. Thus, "the inexperienced listeners were able to adapt quite rapidly to different musical systems" (Krumhansl *et al.*, 2000, p. 14). Finally, there is some evidence that "tonal syntax is closely analogous to the part of language we call grammar," as Krumhansl interprets a recent study on Broca's area of the brain (Maess *et al.*, 2001). Broca's area has been traditionally thought to be specifically implicated in the understanding of language. Their work indicates that "the Broca's area is involved in processing complex, rule-based information associated with sounds, more fundamental than the rules of music or language alone." See Mukherji (2000, ch. 4) for some reflections along these lines.

[4] An almost exclusive concern with artificial systems is perhaps the greatest problem with Douglas Hofstadter's (1979) otherwise brilliant engagement with a similar topic.

[5] I am indebted to an anonymous referee of this journal for raising this issue.

[6] For convenience of informal exposition, I have not included principles such as Extended Projection Principle, Control theory, Chain Condition etc. I hope this does not affect the generality of the discussion. These are treated very differently in the current program anyway.

[7] Recall that Binding theory, which requires introduction of indices as we saw, is no longer internal to language.

[8] There seem to be some exceptions involving long-distance agreement, currently under study, to this general phenomenon.

[9] Roughly, Kitahara takes Concatenate to be an elementary operation and defines Merge in terms of Concatenate for application at root. He introduces another elementary operation, Replace. This leads to an extended definition of Merge (= Concatenate + Replace) that has non-root applications, especially embedding. Move is now defined as extended Merge for operations in subphrase-markers. The burden thus shifts to Replace: is it linguistically specific? See also Chomsky (2001a) for the idea that Move = Agree + Pied-piping + Merge.

[10] It is sometimes suggested that music evolved before language (Cromie, 2001).

References

CHOMSKY, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

CHOMSKY, N. (1972). Remarks on nominalization. In his *Studies on semantics in generative grammar* (pp. 12–61). The Hague: Mouton.

CHOMSKY, N. (1975). *Reflections on language*. New York: Pantheon.

CHOMSKY, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.

CHOMSKY, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

CHOMSKY, N. (1984). *Modular approaches to the study of the mind*. San Diego, CA: San Diego State University Press.

CHOMSKY, N. (1986). *Knowledge of language*. New York: Praeger.

CHOMSKY, N. (1988). *Language and the problem of knowledge: The Managua lectures*. Cambridge, MA: MIT Press.

CHOMSKY, N. (1991). Linguistics and adjacent fields: a personal view. In A. KASHER (Ed.) *The Chomskyan Turn*. Oxford: Basil Blackwell.

CHOMSKY, N. (1994). Bare phrase structure. In G. WEBELHUTH (Ed.) *Government and binding theory and the minimalist program*. Oxford: Basil Blackwell.

CHOMSKY, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

CHOMSKY, N. (1999). Language and brain. Address at *European Conference on Cognitive Science*, 27–30 October, University of Siena.

CHOMSKY, N. (2000a). Minimalist inquiries: The framework. In R. MARTIN, D. MICHAELS, & J. URIAGEREKA (Eds) *Step by step: Essays in honor of Howard Lasnik* (pp. 89–157) Cambridge, MA: MIT Press.

CHOMSKY, N. (2000b). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

CHOMSKY, N. (2000c). Derivation by phase. *MIT Occasional Papers in Linguistics*, *18*, http://web.mit.edu/mitwpl.

CHOMSKY, N. (2000d). *The architecture of language*. N. MUKHERJI, B.N. PATNAIK, & R.K. AGNIHOTRI (Eds). New Delhi: Oxford University Press.

CHOMSKY, N. (2001a). Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics*, *20*, http://web.mit.edu/mitwpl.

CHOMSKY, N. (2001b). Language and the rest of the world. *Bose memorial lecture in philosophy*, 4 November, Delhi University.

CHOMSKY, N. (2002). *On nature and language*. Cambridge: Cambridge University Press.

CHOMSKY, N., HUYBREGTS, R. & RIEMSDIJK, H. (1982). *The generative enterprise*. Dordrecht: Foris.

CROMIE, W.J. (2001). Music on the brain: Researchers explore the biology of music. *Harvard University Gazette*, March 22. http://www.news.harvard.edu/gazette/2001/03.22/

FODOR, J. (1983). *The modularity of mind: An essay in faculty psychology*. Cambridge, MA: MIT Press.

FODOR, J. (2000). *The mind doesn't work that way: Scope and limits of computational psychology*. Cambridge, MA: MIT Press.

GALLISTEL, C.R. (1997). Neurons and memory. In M.S. Gazzaniga (Ed.) *Conversations in the cognitive neurosciences*. Cambridge, MA: MIT Press.

HAUSER, M.D., CHOMSKY, N. & FITCH, W.T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.

HOFSTADTER, D.R. (1979). *Gödel, Escher, Bach: An eternal golden braid.* New York: Vintage Books.

JACKENDOFF, R. (1992). Musical parsing and musical affect. In his *Languages of the mind: Essays on mental representation* (pp. 125–156). Cambridge, MA: MIT Press.

JENKINS, L. (2000). *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.

KITAHARA, H. (1997). *Elementary operations and optimal derivations*. Cambridge, MA: MIT Press.

KOSTER, J., VAN RIEMSDIJK, H. & VERNAUD, J.-R. (1993). In C. OTERO (Ed.) *Noam Chomsky: Critical assessments, vol. 1*. London: Routledge & Kegan Paul.

KRUMHANSL, C.L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, *17*, 53–90.

KRUMHANSL, C.L, EEROLA, T., TOIVIAINEN, P., JÄRVINEN, T. & LOUHIVUORI, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *76*, 13–58.

LARSON, R. & SEGAL, G. (1995). *Knowledge of meaning: An introduction to semantic theory*. Cambridge, MA: MIT Press.

LEIBER, J. (2001). Turing and the fragility and insubstantiality of evolutionary explanations: A puzzle about the unity of Alan Turing's work with some larger implications. *Philosophical Psychology*, *14*, 83–94.

LERDAHL, F. (1996). Calculating tonal tension. *Music Perception*, *13*, 319–364.

LERDAHL, F. & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

MAESS, B., KOELSCH, S., GUNTER, T.C., FRIEDERICI, A.D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, *4*, 540–545.

MUKHERJI, N. (2000). *The Cartesian mind: Reflections on language and musi*c. Shimla: Indian Institute of Advanced Study.

PREMACK, D. (1986). *Gavagai: Or the future history of the animal language controversy*. Cambridge, MA: MIT Press.

RADFORD, A. (1997). *Syntax: A minimalist introduction.* Cambridge: Cambridge University Press.

SELLS, P. (1985). *Lectures on contemporary syntactic theories*. Stanford, CA: Stanford University CSLI.