DISCUSSION NOTE

What is the human language faculty? Two views

RAY JACKENDOFF

Tufts University

In addition to providing an account of the empirical facts of language, a theory that aspires to account for language as a biologically based human faculty should seek a graceful integration of linguistic phenomena with what is known about other human cognitive capacities and about the character of brain computation. The present discussion note compares the theoretical stance of biolinguistics (Chomsky 2005, Di Sciullo & Boeckx 2011) with a constraint-based PARALLEL ARCHITECTURE approach to the language faculty (Jackendoff 2002, Culicover & Jackendoff 2005). The issues considered include the necessity of redundancy in the lexicon and the rule system, the ubiquity of recursion in cognition, derivational vs. constraint-based formalisms, the relation between lexical items and grammatical rules, the roles of phonology and semantics in the grammar, the combinatorial character of thought in humans and nonhumans, the interfaces between language, thought, and vision, and the possible course of evolution of the language faculty. In each of these areas, the parallel architecture offers a superior account both of the linguistic facts and of the relation of language to the rest of the mind/brain.*

Keywords: narrow faculty of language, recursion, parallel architecture, Merge, Unification, lexicon, consciousness, evolution

1. ISSUES FOR LINGUISTIC THEORY. The human language faculty is a cognitive capacity shared by all normal humans but no other species on the planet. Addressing its character scientifically involves engaging (at least) three issues. The first two have been themes of generative grammar for fifty years; the third has become more prominent in the past two decades.

(i) An account of speakers' ability to create and understand an unlimited number of sentences of their language(s). ('knowledge of a language' or 'competence')

(ii) An account of how speakers acquire knowledge of a language. (acquisition)

(iii) An account of how the human species acquired the ability to acquire the knowledge of a language. (evolution)

These three issues are obviously closely interwoven. Addressing acquisition requires a theory of the character of the linguistic knowledge that is acquired. Addressing the evolution of the language faculty requires a theory of the evolved acquisition mechanism in modern humans. And at the same time, a theory of linguistic knowledge is constrained by the need for languages to be learnable, and a theory of the acquisition mechanism is constrained by the need for it to have evolved (to the extent that we have any grasp of what is evolutionarily possible).

The present essay compares my own view of the character of the language faculty (Jackendoff 2002, 2007a, 2010, Culicover & Jackendoff 2005) with that of the MINI-

* This discussion note originated in my opening keynote address for the summer school on the Origins of Language, at the Université du Québec à Montréal in June 2010. I wish to thank Claire Lefebvre, director of the summer school, for the invitation. I am grateful to many people for valuable comments on earlier versions: Peter Culicover, Lila Gleitman, Heike Wiese, Eva Wittenberg, Neil Cohn, Gina Kuperberg, Tecumseh Fitch, Ari Goldberg, Naomi Berlove, members of the Moses Seminar at MIT, and three anonymous referees. Some of this material was developed while I was a visiting faculty member at the Santa Fe Institute in 2009; I am immensely grateful to SFI and especially to Chris Wood for making such a splendid experience possible.

MALIST PROGRAM (MP) and biolinguistics.[1] I focus on basic foundational assumptions of the two views, as well as on their bearing on the nature of linguistic knowledge and the evolution of the language faculty. I begin with points of agreement, and then move on to divergences.

1.1. KNOWLEDGE OF LANGUAGE. A speaker's ability to use a language requires a systematic mapping between an unlimited number of thoughts or meanings and an unlimited number of sound sequences (or, in the case of signed languages, gesture sequences). Since this mapping must be stored in a finite brain, and since it must be learned in a finite amount of time, it must be encoded in finite terms: a set of stored units and/or complexes, plus a means to combine them in unlimited fashion.

Two speakers are mutually intelligible if their own personal mappings are close enough to enable them to achieve mutual understanding on the basis of exchanged signals. We typically idealize this situation by assuming that the mental systems in the members of a speech community are uniform and thereby can be regarded as a single language. But we readily drop the idealization when dealing with differences of vocabulary, dialect, and register. We also drop the idealization, almost without noticing, when dealing with young children's speech, which only approximates adult language.

An account of the ability to use language must therefore answer this question: What units (or complexes of units) and what combinatorial principles are stored in speakers' brains that enable them to map systematically between thoughts/meanings and sounds? The inventory of these units and combinatorial principles constitute speakers' 'knowledge of their language'.

1.2. ACQUISITION. The second part of an account of the language faculty is a theory of how speakers acquire their knowledge of language. Generative linguistics made a crucial contribution to the emergence of cognitive science in the 1960s and 1970s by raising the question of what cognitive resources are necessary in order to acquire and use language. Chomsky (1965, 1968, 1975, 1981) used to call these resources universal grammar or UG.[2]

The issue has been sharpened by Hauser, Chomsky, and Fitch (2002), who distinguish what they call the 'broad language faculty' (FLB) from the 'narrow language faculty' (FLN). FLB includes the narrow language faculty plus whatever other mental machinery is necessary for the acquisition and use of language but also serves other cognitive purposes, such as:

- an auditory system
- a motor system
- working memory
- long-term memory[3]

[1] Critiques of MP from other viewpoints, with varying degrees of politeness, include Johnson & Lappin 1999, Newmeyer 2003, 2008, Postal 2004, Seuren 2004, and Kinsella 2009, among others. For one view on the history of how mainstream theory progressed from the standard theory to MP, see Culicover & Jackendoff 2005:Chs. 2–3.

[2] Universal grammar unfortunately is frequently confused with universals of language. For example, Evans and Levinson (2009) argue that there are no language universals (which might or might not be true) and erroneously jump to the conclusion that there is no universal grammar. This does not follow. For instance, UG might provide a set of options from which languages choose (such as parameters), or it might be a set of attractors toward which languages tend statistically without necessarily adhering to them rigidly (Jackendoff 2002:Ch. 4, Pinker & Jackendoff 2005).

[3] Hauser, Chomsky, and Fitch specifically exclude memory from FLB, for reasons unclear to me.

- attention
- 'general intelligence'
- a capacity for joint attention (present to some degree in chimpanzees (Povinelli 2000))
- a capacity for vocal imitation (not present in chimpanzees but present in some birds and cetaceans (Fitch 2011))
- a capacity for voluntary, fine-scale control of the vocal tract (also not present in monkeys (Cheney & Seyfarth 2005) but again present in some songbirds and cetaceans (Fitch 2011))
- (possibly) a THEORY OF MIND (presence in chimpanzees is controversial (Povinelli 2000, Tomasello et al. 2003)).

Since the earliest days of generative grammar (Chomsky 1959, 1965, 1968), an important claim has been that these capacities alone are not sufficient for language acquisition: some aspects of language acquisition require unlearned capacities specific to the linguistic modality, and these constitute the narrow language faculty or UG. Although many people, linguists and nonlinguists alike, find the notion of UG implausible, it is well known that unlearned special-purpose capacities are found throughout the animal kingdom, for instance bee dances, nest-building and specialized vocal behavior in birds, elaborate social organization in many species of insects, birds, and mammals, and patterns of aggression and sexual behavior ubiquitously (Tinbergen 1951, Lorenz 1952, Dawkins 1976, Heinrich 1989, Marler & Slabbekoorn 2004). Therefore, on general biological grounds, it should not be objectionable to posit unlearned special-purpose aspects of the human language capacity.

Chomsky's earlier work (e.g. 1965, 1975) argued for a very rich UG. Over the years UG has been pared down more and more (e.g. Chomsky 1995, 2000), to the point that some people (e.g. Goldberg 2003) see Chomsky as actually having recanted on the claim that there is a UG. Of course, there are good reasons to want to minimize UG. At the very least, it is simply good science to posit as little as possible. But a more important reason is that any special mental capacity has to be encoded somehow on the genome, such that the genome builds a brain with special information-processing capability. At the moment we have very little idea how the genome leads to the construction and the functioning of the brain. And we also have no idea how special mental capacities (linguistic or otherwise) can be realized through details of brain wiring (Gallistel & King 2009, Gazzaniga 2010). But we do know that, in principle, the problem has to be solved, and a leaner UG gives us a better chance of succeeding. This is one important issue where linguistics makes contact with biology, taking the study of language beyond just the description of languages and language universals.

1.3. EVOLUTION. This leads to the third prong in an account of the language faculty. Anything special in the genome places a burden on evolution. The genetic differences responsible for UG/FLN (if any) could have arisen through a sequence of small changes, or, as Chomsky (2005, 2010) has advocated, they could have arisen through some single mutation that led to radical phenotypic change. In a single-mutation scenario, it is assumed that individuals endowed with a language faculty had some adaptive advantage over those who lacked one, so natural selection played a role in perpetuating the language faculty. In an incremental scenario, it is assumed that individuals endowed with a more highly developed language faculty had some adaptive advantage over those with a less highly developed one, so natural selection played a role in each step. (The nature of that adaptive advantage is a matter of dispute; I touch on it briefly in §9.)

On the one hand, Chomsky's view of a single mutation is a priori simpler, of course. On the other hand, consider bird songs. The literature (e.g. Hauser 1996, Hultsch et al. 1999, Marler & Slabbekoorn 2004, Podos et al. 2004) explores in detail the balance of innate vs. learned structure in bird songs. But here, in a system far less complex than human language, I find no suggestion that the innate aspects of any particular species' variety of song are the product of a single mutation from a nonsongbird ancestor. Rather, cross-species comparative evidence has been used to help infer incremental routes to the present-day capacities of different species (Price & Lanyon 2002). Of course, we have no cross-species evidence for variation in the language faculty, so all we have to go on is a few fossils and a lot of reverse engineering. But absence of concrete evidence does not preclude the possibility that the capacity evolved incrementally—it only makes it harder to determine the route that evolution took.

One crucial problem in accounting for the origins of human language is how an ability to acquire language could have arisen when there was no language around to learn. But this problem is not unique to the evolution of language. It arises in trying to account for any communicative capacity in any organism (Fitch 2011). So this is not a definitive argument against the language capacity having evolved incrementally.

Explaining the origins of the language faculty depends of course on what one thinks the human language faculty IS—the main issue to be addressed here. At the moment the evidence from the evolution of language is so slim, and the evidence from biological development is so far removed from the details of language structure, that I find it hard to give these sorts of evidence priority over far more clearly established facts of modern linguistic structure in deciding which theory of linguistic knowledge to pursue. Evolutionary considerations do lead us to seek a theory that minimizes demands on the genome—but not at the expense of a rigorous account of the modern language faculty.[4]

2. ISSUES FOR BIOLINGUISTICS. In recognition of the goal of interfacing linguistic theory with biology, practitioners of the minimalist program have begun calling the enterprise 'biolinguistics' (e.g. Jenkins 2000, Larson et al. 2010, Di Sciullo & Boeckx 2011, and the online journal *Biolinguistics* (www.biolinguistics.eu)). The biolinguists have for the most part focused on three issues. The first is the genetic basis of the language faculty, including extensive discussion of the possible role of the FOXP2 gene, as well as work (e.g. Stromswold 2010) on genetic language impairments. This issue has been part of the literature for decades (even if much of the evidence is new) and is not specific to the biolinguists. The second issue is what principles of evolutionary and developmental biology could have led to a language faculty. Here, it seems to me, the distance between the nature of linguistic structure and what is known about the genetic basis of biological development is so great that the best that can be attempted at the moment is informed speculation, and I have nothing to say about this issue here. The third issue is the 'first-principles' question posed in Chomsky 1995: 'How perfect is language?', where perfection is defined in terms of elegance, lack of redundancy, and com-

[4] For instance, Chater, Reali, and Christiansen (2009) claim that there cannot possibly be a language-specific universal grammar. Their evidence is that, in their simulations of the evolution of language, it proves impossible for language-specific aspects of the language capacity to develop in the face of language change; their argument is that language change creates a 'moving target' that cannot be tracked quickly enough by innate biases in language acquisition. Their simulations are built around highly simplified assumptions about language, however, and, more importantly for the present argument, the authors do not address any of the voluminous evidence from synchronic grammars, from language acquisition, and from language impairment that there must be some innate predispositions. See Jackendoff 2002:Ch. 4 for discussion of putative arguments against universal grammar.

putational efficiency. Of the three, this issue has been the most central to biolinguistic inquiry into the actual form of the language faculty.

In stressing the deep questions of genetics and optimal design, however, the biolinguists have bypassed an important biological and psychological issue. Evolution probably did not come up with radically new kinds of neurons; if special kinds of neurons had been discovered in the language areas of the brain, we would have heard about them. Moreover, although it is important to bear in mind how little we know about how assemblies of neurons encode and compute information, we might still be inclined to guess that evolution did not come up with radically new kinds of mental computation, mental processes, and mental architecture in inventing a language faculty. (See Gallistel & King 2009 for extensive discussion of the conservatism of neural mechanisms.) So to the extent that a theory of language permits a graceful integration with a plausible picture of the structure and function of the rest of the mind/brain, it places fewer demands on the genome, and therefore it is a better theory. I would contend that this too should be considered a criterion of linguistic theory from a biolinguistic perspective.

We see throughout this essay that the criterion of graceful integration conflicts in many respects with the criterion of perfection. For a first case, consider the notion that perfection implies the absence of redundancy. As it turns out, brain computation appears to be full of redundancy. For instance, the visual system contains multiple spatial maps, each slightly different from the others, but with considerable overlap (Marr 1982, Koch 2004). Also within the visual system, the process of depth perception relies on multiple cues: the convergence of the eyes, the contraction of the lens, binocular disparity, occlusion, and high-level knowledge of how big particular things are supposed to be (Hochberg 1978, Marr 1982). In many situations, these processes are redundant, and that makes visual judgment more robust and efficient. But in other situations they are not completely redundant: each process has a range of situations in which it is more effective, and that makes the system as a whole more flexible, capable of operating under a wider range of conditions.

In turn, if redundancy is characteristic of the brain, this changes the desiderata for linguistic theory: one should not necessarily be trying to squeeze all the redundancy out of linguistic representations and linguistic processing. Consider, for instance, the apparently semi-independent systems for cuing thematic roles, such as agreement, case, word order, and morphology (as in *bus* being the patient and *driver* being the agent of *bus driver*). Similarly, information structure can be marked by prosody, by topic and focus particles and affixes, and by special constructions such as clefts; and often these devices are used redundantly. Each of these phenomena has the flavor of the multiple mechanisms for depth perception. So a biolinguistic approach that takes graceful integration seriously might well not try to reduce either of them to a single underlying mechanism, as would be dictated by considerations of nonredundant (or 'perfect') computation.

Turning to the lexicon, Chomsky (1965) asserts without argument that it is a nonredundant repository of idiosyncrasies; this assumption has been carried through the tradition ever since. However, psycholinguistic evidence (e.g. Baayen et al. 2002) shows that speakers store some completely redundant material in their mental lexicons, for instance, high-frequency regular plurals. This demonstrates that the lexicon as stored in the brain is inherently redundant. Therefore, by the criterion of graceful integration, a linguistic theory that acknowledges this redundancy should be preferred to one predicated on a nonredundant lexicon. (For theoretical arguments for the redundancy of the lexicon, see Jackendoff 1975, 1997a:§5.6; for a critique of nonredundancy as a guiding principle of the minimalist program, see Kinsella 2009.)

It is true that, as Chomsky reminds us, 'the guiding intuition that redundancy in computational structure is a hint of error has proven to be productive and often verified' (2005:10). This intuition lies behind good scientific practice, and too easy an acceptance of redundancy can be a source of theoretical complacency. However, good science also ought to be able to consider redundancy as a solution when the facts patently point that way.

3. RECURSION IN LANGUAGE AND ELSEWHERE IN THE MIND/BRAIN. Another case where 'first principles' and empirical facts about the mind/brain diverge is the issue of recursion. Chomsky (1995, 2005, 2010; Berwick & Chomsky 2011), seeking 'efficient computation', asserts that we should assume that language uses the simplest possible means of creating combinatoriality. The absolute simplest way of achieving combinatoriality, he argues, is by putting two things together into a larger unit: taking units *a* and *b* and creating the set $\{a, b\}$. This operation, Merge, can apply recursively to its own output to build binary branching structures of arbitrary size. In some expositions, Merge also labels the set with the label of one of its constituents: $\{a, b\}_a$ or $\{a, b\}_b$, thereby creating the headedness of syntactic constituents. In addition, by taking as one of the merged elements a constituent of the other element, so-called Internal Merge, we get structures like $\{c, \{a, \{b, c\}\}\}$. Such structures, it is argued, form a basis for the copy theory of movement, which has been a central aspect of mainstream generative grammar for over fifty years (Chomsky 2005).

The minimalist program reconstructs syntactic theory around Merge as the central computational operation. Without specifically mentioning the minimalist program and Merge, Hauser, Chomsky, and Fitch (2002) propose that FLN consists solely of the ability to build syntactic structures recursively, plus the mappings from syntax to the 'sensory-motor interface'—the auditory and motor systems in language perception and production respectively—and to the 'conceptual-intentional interface'—the formation of thought and meaning, or what I termed 'general intelligence' above.

I take issue with every part of this conception of the language faculty. In the remainder of this discussion note, I first show that recursion is not unique to language and therefore cannot constitute FLN. I next compare the step-by-step nature of Merge with constraint-based rule application, and then with a different fundamental combinatorial operation, Unification. A discussion of the roles of phonology and semantics follows, and finally all of this is synthesized into an alternative theory of the language faculty, the PARALLEL ARCHITECTURE, concluding with discussion of evolutionary issues.

To begin, let us consider what is meant by 'recursion', and what it means to say an organism 'has recursion' in its cognitive repertoire.[5] Evaluating whether a particular domain of behavior or cognition is recursive requires a theory of mental representations in that domain. And unfortunately, outside of linguistics, there are virtually no theories of mental representation (with some exceptions that I mention just below). In the absence of such theories, it is premature to assert that only language requires recursion.

In traditional formal terms (at least in linguistics), a set of rules is called 'recursive' if the rules can apply to their own output an unbounded number of times and thereby can produce an unlimited number of expressions from a finite set of primitives. Let us call such a set of rules FORMALLY RECURSIVE. (The recursion can occur in a single rule, as in S → A (S) B, or it can be distributed through a rule set, for instance S → A C;

[5] See Kinsella 2009 and Fitch 2010 for more extensive discussions of recursion that overlap in some respects with the present exposition and diverge in others.

C → (S) B.) In practice, however, the term is often applied to a domain of representation such as the syntax of English, in terms of the repertoire of structures it contains: we call a domain 'recursive' if it has constituent structure, and if constituents can be embedded in others to an unlimited depth. Let us call such a domain of representation STRUCTURALLY RECURSIVE.

I suggest that for purposes of cognitive science, the notion of structural recursion is more useful than the notion of formal recursion. When we are dealing with artificial systems such as formal logic and computer languages, where the rule system is given, formal recursion is easy to establish. But when we are dealing with natural systems such as language and cognition, we first have to establish what the structures are, and then from them we must induce the rule system. If we can establish that a domain is structurally recursive, we can conclude that the rule set that defines the domain must be formally recursive, EVEN IF WE DO NOT YET KNOW WHAT THE ACTUAL RULES ARE. In such cases, our only evidence for formal recursion is structural recursion.

The crucial issue, then, is to establish the presence of structural recursion on the basis of the organism's behavior. For instance, recursive constituent structure such as the tail recursion in 1a must be distinguished from simple unlimited concatenation (as in the organization of a shopping list), which has no constituent structure and no embedding, as in 1b.

(1)	a. Recursive structure	b. Unlimited concatenation
	[A B]	[A B]
	[[A B] C]	[A B C]
	[[[A B] C] D]	[A B C D]
	…	…

To demonstrate that recursion is genuinely involved, it must be shown that the strings have constituent structure, for instance, that in the string [A B C D], the substrings [A B] and [[A B] C] form replicable units that can occur in other configurations. Otherwise, there is no evidence that the strings in question are not generated by mere concatenation (the 'Kleene star', *W**). This distinction has always been stressed in generative grammar (e.g. Chomsky 1963, 1965): the crucial issue for a theory of grammar is the set of structures the rules generate (strong generative capacity), not the set of strings (weak generative capacity).

Similarly, a set of strings such as 2a might be generated by center-embedded recursion. But it might equally be generated by a procedure that requires that A be followed by A or B, that B be followed by B, and that the number of *A*s and *B*s be the same—but with no internal constituent structure. The computational requirements of the two procedures are quite different.

(2)	a. Recursive center-embedded structure	b. Unlimited concatenation plus counting	
	[A B]	[A B]	(1 A, 1 B)
	[A [A B] B]	[A A B B]	(2 As, 2 Bs)
	[A [A [A B] B] B]	[A A A B B B]	(3 As, 3 Bs)
	…	…	

Again, to show that 2a is the correct structure, one must demonstrate that the bracketing delimits constituents, for instance that [A B] forms a constituent in the second and third lines of 2a.

I raise this particular case because of the widely cited claim (Gentner et al. 2006) that starlings can learn a recursive center-embedded grammar, when what they actually learn is to identify strings of the form shared by 2a and 2b. Gentner and colleagues give no argument about the constituent structure of the strings the birds learn, so we are not entitled to draw any conclusions about the birds' capacity for structural (hence formal) recursion. (Fitch (2010) makes a similar argument against the conclusions of Gentner and colleagues.)

Another reason to give structural recursion theoretical priority over formal recursion is that a traditional recursive rule set, whose output is linear strings of terminal symbols, cannot be applied to domains with two- and three-dimensional organization, such as the configuration in Figure 1. This array displays unlimited possibilities for embedding. It can be seen as rows of five *x*s or *o*s, combined into arrays of fifteen elements, which further pair up to form arrays of thirty elements. These arrays in turn are arranged in a 3 × 3 matrix. Two or more copies of this matrix can obviously be combined into a larger configuration, which in turn can be embedded into a still larger configuration.

```
xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx
ooooo  ooooo    ooooo  ooooo    ooooo  ooooo
xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx

xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx
ooooo  ooooo    ooooo  ooooo    ooooo  ooooo
xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx

xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx
ooooo  ooooo    ooooo  ooooo    ooooo  ooooo
xxxxx  xxxxx    xxxxx  xxxxx    xxxxx  xxxxx
```

FIGURE 1. A two-dimensional structurally recursive array.

A notion of structural recursion defined in terms of unlimited hierarchical embedding is appropriate to Fig. 1, but a notion in terms of traditional formal grammars is not. In particular, binary Merge is incapable of producing this sort of recursion, for two reasons. First, the spatial relation of elements being combined matters: the *x*s are combined horizontally, the rows of *x*s and the rows of *o*s are combined vertically, and so on. Merge, which creates unordered sets, does not incorporate this degree of freedom. Second, since Merge is binary, it cannot produce simple unstructured groups of three and five; and there is no evidence that the first two or three or four in each group together form a constituent.

I conclude that structural recursion is not unique to language: the visual faculty contains a form of recursion considerably richer than what binary Merge can produce. I know of no evidence bearing on whether other primates are also able to understand visual configurations of the sort illustrated in Fig. 1, but I would hazard a guess that this aspect of cognition is of considerably earlier ancestry than the human lineage.

Variants of Fig. 1 appear in Culicover & Jackendoff 2005 and Pinker & Jackendoff 2005, with similar arguments about the limitations of binary Merge. Chomsky (2005:16), possibly in response to these arguments (though without citation), admits the possibility of *n*-ary Merge, then says 'it has been generally assumed that these units are mostly, maybe always, binary'. In fact, as far as I can determine, aside from this one instance, within the minimalist program it has been ALWAYS assumed that these units are ALWAYS binary, including in Chomsky's subsequent work.[6]

Also possibly in response to the arguments concerning Fig. 1 (but again without citation), Chomsky (2010:53) says, 'Suppose the single item in the lexicon is a complex object, say some visual array. Then Merge will yield a discrete infinity of visual patterns, but this is simply a special case of arithmetic, and therefore tells us nothing new about recursion beyond language'. I find Chomsky's reference to arithmetic baffling, and he does not address the problems for Merge raised by the two-dimensionality and nonbinarity of the array.

Since recursion is a feature of mental computation manifested in the visual system, the possibility arises that the language faculty did not develop binary Merge out of the blue, but rather adapted the richer form of recursion displayed in Fig. 1 to linguistic purposes. According to such a scenario, syntactic recursion is actually part of the general-purpose part of the language faculty (FLB), permitting a more graceful integration of language with the rest of the brain. To be sure, language is different, in that it has put recursion to use in mapping between signals and thoughts; that sort of recursion is indeed unique in the animal kingdom. But if recursion occurs in the visual system as well, it DOES tell us something new about language, pace Chomsky, namely that recursion per se is not what makes language distinctive.

Part of the adaptation of recursion to syntax is HEADEDNESS: the fact that each constituent has a designated element as head, from which the label of the constituent as a whole is derived. The hierarchy in Fig. 1 is not headed, so it might be possible that headedness is the crucial innovation. To consider this possibility, and to show that Fig. 1 is not an isolated example of nonlinguistic recursion, I wish to enumerate some other plausibly recursive structures in cognition.

Figure 2 illustrates David Marr's (1982) theory of the 3D model in visual cognition, in which an object is understood in terms of recursive decomposition into parts and parts of parts. In this structure, meant to represent our spatial understanding of the human figure, there is a notion of headedness: the torso is the head of the body as a whole; the palm is the head of the hand.

Marr's representational approach to vision—the closest a theory of vision has come to being comparable with linguistic theory—lost favor in the vision community shortly after his death in 1982, but I know of no other proposals that aspire to account for the same range of insights.[7] Moreover, the sort of organization Marr proposes extends to the composition of the visual field as a whole, recursively structuring the understanding of the entire seen and unseen spatial environment: the room one is in is part of some building, the building is part of some campus, the campus is part of some city, and so on—though this part of the hierarchy does not seem to involve headedness (Landau & Jackendoff 1993).

[6] The 'general assumption' that syntactic structures are binary goes back to widely cited arguments by Kayne (1983) and Haegeman (1992). Culicover & Jackendoff 2005 demonstrates that these arguments are of questionable force, and that therefore there is no valid argument for purely binary branching, aside from a preference for this particular sort of simplicity.

Lasnik (2011) reviews Chomsky and Miller's (1963) argument that the syntactic structure of phrases like *the beautiful old brown dog* must be flat, showing that there is no motivation for constituent structure within the string of adjectives. Having pointed out that such shopping-list-like sequences pose a puzzle for the theory of Merge, Lasnik immediately drops the subject without attempting to reconcile them with binary Merge (2011:361).

[7] My colleagues in the vision community assure me that this is indeed the case. Unfortunately, it is difficult to give citations when the point being made is the ABSENCE of citations.

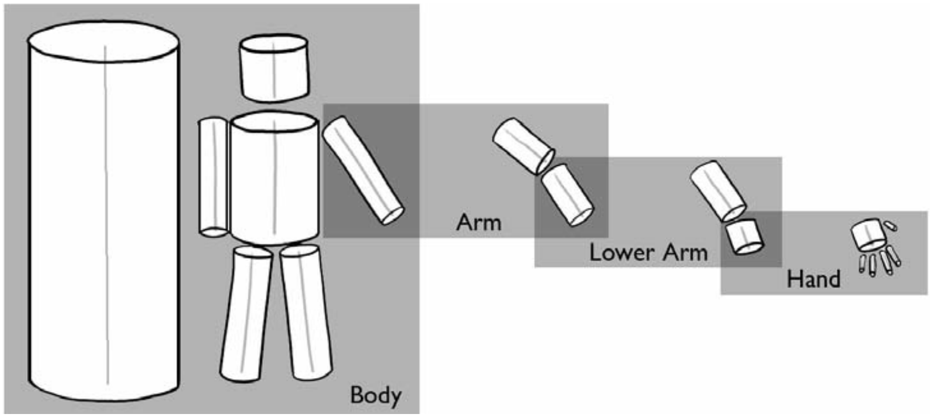

FIGURE 2. Marr's 3D model structure for the human body.

Music too has recursive hierarchical structure. Figure 3 illustrates the representations of three components of music proposed in Lerdahl & Jackendoff 1983, based in part on work by Heinrich Schenker (1935). (The tune is the Beatles' 'Norwegian wood'; see Lerdahl & Jackendoff 2006 for a fuller analysis. The words are not essential to the musical analysis; they are given here so that readers unfamiliar with musical notation can follow the exposition.)

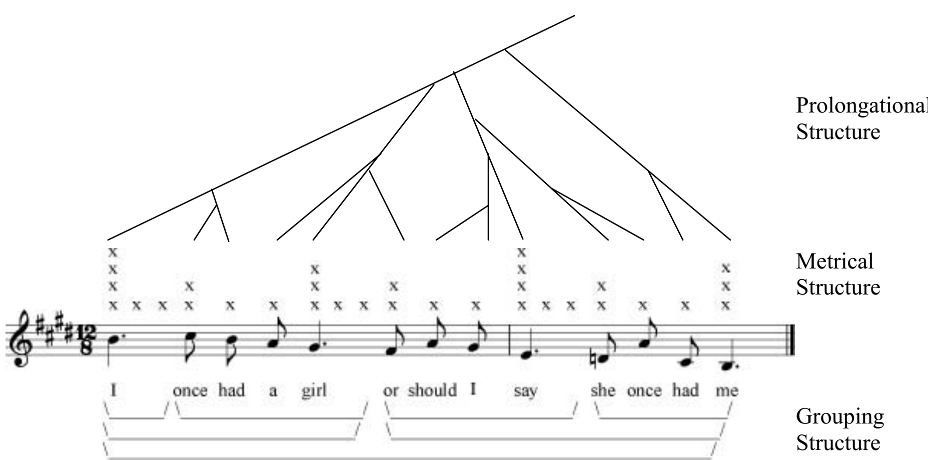

FIGURE 3. Grouping, metrical, and prolongational structures in music.

The GROUPING STRUCTURE below the notes encodes the phrasing of the music: the smallest groups consist of the music associated with the texts *I*; *once had a girl*; *or should I say*; and *she once had me*. These are grouped by pairs: *I once had a girl* and *or should I say she once had me*. These two groups form a larger group consisting of the whole fragment, which is paired with the second phrase of the song (*She showed me her room* ...), and so on until the entire song forms a group. Groups can be embedded indefinitely deeply, all the way to fifteen-minute movements of symphonies; hence this domain is structurally recursive, though not headed.

The *x*s above the notes in Fig. 3 encode the music's METRICAL STRUCTURE, the patterns of strong and weak beats. This structure is analogous to the stress grid in phonology, and even more so to the metrical grid in poetry. The smallest level of beats (the lowest row of *x*s) is isochronous; every third beat is a strong beat. Of those strong beats, every other one is stronger (on *I*, *girl*, *say*, and *me*), and the first and third of these (*I* and *say*) are even stronger. This domain, though hierarchical and headed, is not structurally recursive; it usually goes to about four to six levels and then fades out. In particular, there is no strongest beat in the whole piece.[8]

The tree structure above the metrical grid in Fig. 3 represents the PROLONGATIONAL STRUCTURE of the music, which encodes a listener's recursive sense of patterns of tension and release. Each note corresponds to a branch of the tree, and it is understood as subordinate to the note on the branch that its own branch attaches to. The longest branch, attached to the note sung to *I*, attaches eventually to the end of the piece. The next most important element in the phrase is the last note (sung to *me*). These two branches delineate the skeleton of the phrase, an octave descent from high *b* to low *b*. The next most important branches are connected to *g#* (*girl*) and *e* (*say*); thus a slightly more filled out skeleton is an arpeggiated chord *b-g#-e-b* (*I … girl … say … me*). Next most important are the *a* and *f#* surrounding the *g#*, and the *d* natural and *c#* toward the end. The resulting skeleton is a descending scale (*I … a girl or … say she … had me*). Finally, the most detailed elaboration adds the remaining notes as decoration. Again, this hierarchy can be extended indefinitely as the piece gets larger, so the domain is structurally recursive and headed.[9]

Neither of the two recursive structures—grouping and prolongational structure—has a direct analogue in language. In particular, despite recent proposals by Katz and Pesetsky (2009) and by Tsoulas (2010), prolongational structure bears little resemblance to syntax in language, beyond the very coarse observation that Merge can be used to create headed hierarchies of notes instead of words. (See Patel 2008, Jackendoff 2009 for detailed comparisons of music and language.)

For another example of nonlinguistic recursive structure, Jackendoff 2007a develops a representational system for the structure of complex actions, drawing on literature on robotics and planning (Badler et al. 2000, Bindiganavale et al. 2000, Kipper & Palmer 2000) and on brain damage to action planning (Humphries et al. 2001). The structure turns out to be richly hierarchical. Structures along similar lines, though less elaborate, have also been proposed for some sorts of ape actions (Whiten 2002). Figure 4 encodes the structure of my personal routine for the completely banal action of making coffee in an automatic drip coffee maker (others' routines may differ).

[8] Fred Lerdahl has suggested (p.c.) that metrical grids might be formally recursive but not structurally recursive—that the rules are recursive, but the restriction to a small number of levels is due to constraints on working memory for temporal regularities.

Berwick (2011a) suggests that metrical grids in poetry, along lines proposed by Fabb and Halle (2008), can be regarded as produced by Merge. Like musical metrical grids, however, their hierarchy is of limited depth, so the domain is not structurally recursive. In an effort to find animal analogues of linguistic structures, Berwick also attempts to draw a parallel between these metrical grids and the rhythmic patterns of birdsong measured and discussed by Coen (2006). He fails to establish, however, that there is any hierarchical temporal structure in birdsongs beyond a repeated alternation of high and low amplitude segments, that is, a single level of embedding.

[9] Kinsella (2009) and Fitch (2010) discuss recursion in music, but do not address the details of the structures adduced in Lerdahl & Jackendoff 1983. However, they both refer to Hofstadter's (1979) discussion of recursively embedded key regions in the course of a piece. This phenomenon actually is a special case of prolongational structure, in that a segment of a piece in a key other than the tonic generally forms a prolongational constituent.

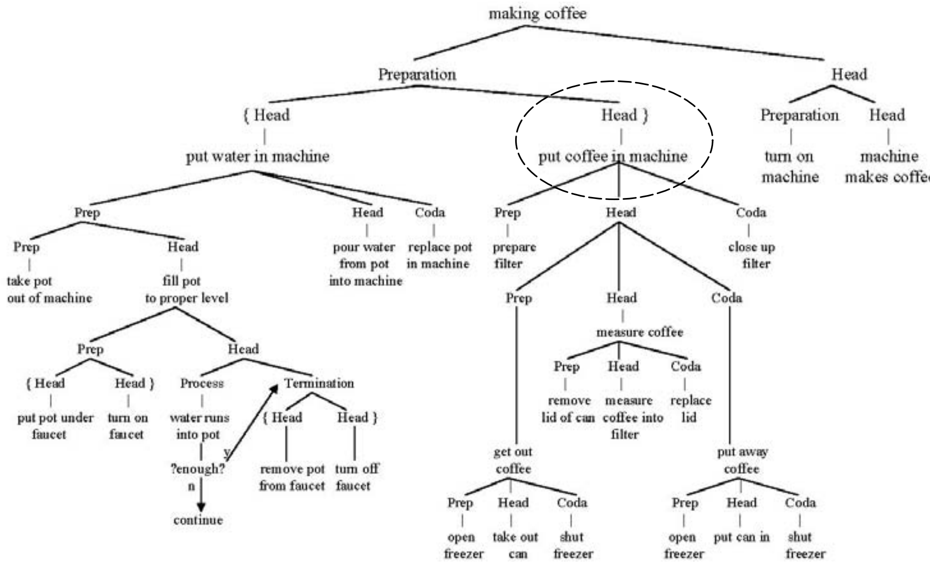

FIGURE 4. The structure of a complex action: Making coffee.

The basic structural unit is the elaboration of an action into a HEAD (performing the 'core' of the action), an optional PREPARATION (getting ready or setting the stage for the action), and an optional CODA (rounding off the action or reestablishing the status quo). Each of these constituents may itself be elaborated into preparation, head, and coda. For instance, the circled head labeled 'put coffee in machine' is elaborated into preparing the filter (preparation), putting the coffee in (head), and closing up the filter (coda). This subordinate head in turn requires getting the coffee can out of the freezer (preparation), measuring the coffee (head), and putting the coffee away (coda); and each of these has further elaboration.[10]

This domain of structures is clearly structurally recursive and headed, and there are even some possible elaborations that have the flavor of variable binding and long-distance dependencies. For instance, suppose you go to take the coffee out of the freezer and discover you're out of coffee. Then you may append to the structure a giant preparation of buying coffee, which involves finding your car keys, driving to the store, going into the store, taking coffee off the shelf—the head of the whole preparation—then paying, leaving the store, and driving home. The crucial thing is that in this deeply embedded head, what you take off the shelf had better be COFFEE—the same thing you need in the larger structure this action is embedded in. The resulting structure, therefore, has a kind of long-distance binding in it, rather analogous to the binding in a relative clause such as *the coffee$_i$ that I realized I needed to go to the store to buy t$_i$*—contrasting with **the coffee that I realized I needed to go to the store to buy cheese*, which is analogous to bringing the wrong thing home from the store.

[10] Figure 4 contains further structural elements, in particular a head elaborated as process plus termination, reminiscent of the test-operate-test-exit units in the action patterns of Miller et al. 1960. The structure would become still more elaborate if one were to try to take into account multitasking and interruption of one action by another.

For a final example of recursion, Neil Cohn has been developing a grammar for wordless comic strips (Cohn 2010, 2011, Cohn et al. 2010). The grammar is intended to explain how readers combine a series of panels into a coherent narrative. Figure 5 illustrates two of his structures for novel *Peanuts* strips, arranged by Cohn from published panels.

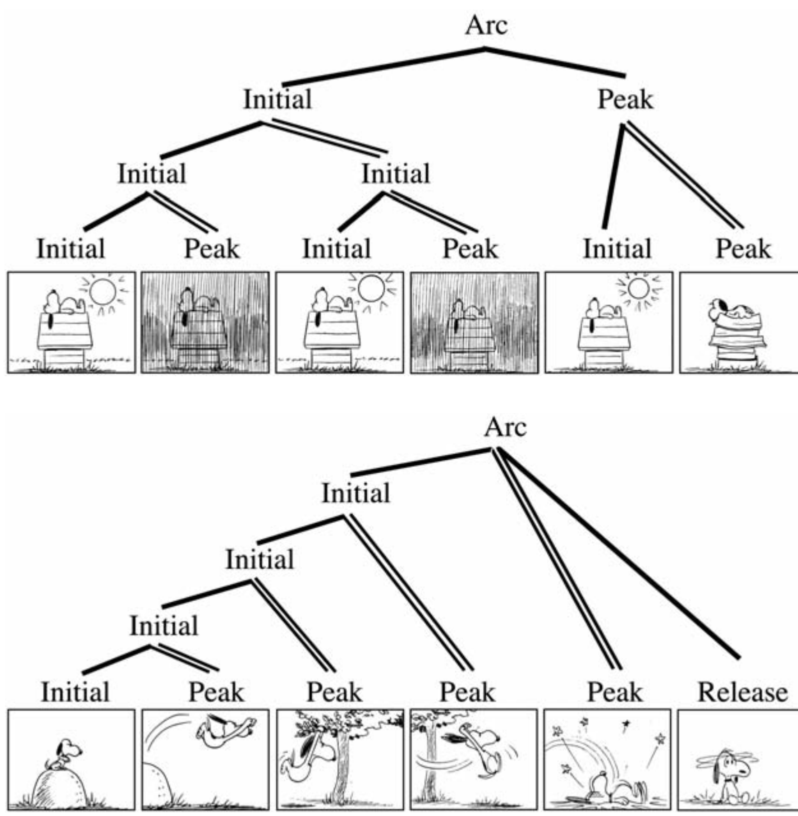

FIGURE 5. The narrative structure of two *Peanuts* comic strips.

In these structures, a narrative ARC decomposes canonically into an optional ESTABLISHER, which sets the scene (not shown in these examples), an optional INITIAL (setting an event in motion), a PEAK (the central aspect of the event), and an optional RELEASE (the aftermath of the event). (The initial, peak, and release are roughly analogous to preparation, head, and coda respectively in Fig. 4, but diverge in some respects.) In turn, any of these parts can be elaborated into a full event consisting of establisher, initial, peak, and release, potentially building up structures of indefinite depth. Hence the domain is structurally recursive, with the peak as head of a constituent.

Using both theoretical arguments and experimental evidence, Cohn shows that certain roles in a narrative structure are more prone to processing difficulty, that there are phenomena parallel to linguistic ellipsis and coercion, that there are long-distance dependencies, and that the coherence of the structures is based both on structural principles of narrative and on associative semantic coherence. Cohn also shows—and it is not too surprising—that these structures can readily be extended to verbal narratives; they are the structures that make a sequence of event descriptions come to be understood as a coherent story (Gee & Grosjean 1984, Mandler 1987). Of course, verbal narratives typically contain multiple sentences that are not connected syntactically. This means that the structure of verbal narratives cannot arise through Merge in syntax, since syntax is limited to the structure of single sentences (a point made also by Bickerton 2010).

The upshot of all these examples is that structural recursion and headedness are ubiquitous in cognition, and that across cognitive domains, formal recursion must be considerably richer than binary Merge. Therefore recursion does not belong to the narrow faculty of language, much less constitute it, as Hauser, Chomsky, and Fitch (2002) propose. (I hasten to point out that the ubiquity of recursion does not constitute grounds to conclude that there is NO narrow faculty of language. It is only possible to conclude that THIS particular claim about the narrow faculty of language is wrong.)

Of course, recursive structures cannot exist without units to combine. Each of the structures illustrated here uses a different collection of units, and the combinatorial details differ in various ways. Pinker & Jackendoff 2005 suggests that what makes language LANGUAGE is that recursion combines WORDS (and/or morphemes, depending on one's view of the lexicon), where words are long-term memory linkings of structured sound, syntactic features, and structured meaning. That is, FLN includes the capacity to learn WORDS in profusion and to apply recursion to THEM. No other species can do this. But the basics of recursion itself come off the shelf, from FLB.

4. STEP-BY-STEP VERSUS CONSTRAINT-BASED COMPUTATIONAL PRINCIPLES. Another tacit assumption behind Merge is that the grammar operates step by step: A is combined with B, then {A, B} is combined with C, and so on. This leads to an inherent bottom-up order to building structure. In older versions of generative grammar, of course, the procedure was top-down instead, starting with S and expanding it. In either case, it has always made sense to say one operation happens 'after' another, for example, that morphology takes place 'after' syntax, or that material is 'sent' to PF (phonological form) at the end of a phase 'before' syntax constructs the next phase. When I was a graduate student in the 1960s, we were always taught that this has nothing to do with processing: a speaker does not start with S, expand it into a syntactic structure, insert the words, move things around, then send the result off to semantics and phonology to finally find out what the sentence means and how to say it. Similarly in the minimalist program: a speaker does not choose a bag of words (the numeration) and then build them into a structure starting from the most deeply embedded point (generally near the right-hand edge), and gradually, phase by phase, figure out what it means and how to say it.

If one asks what this ordering of operations means in terms of the brain, however, the answer in the lore is something to the effect that it is 'metaphorical', or that it is merely a way of defining a set of sentences inductively, the way Peano's postulates define the set of natural numbers. It is meant as a characterization of competence, not of performance. But then the question is: What is this notion of ordering metaphoric FOR? What does this inductive definition have to do with how the brain actually processes language? My sense is that these locutions are just ways to evade the difficulty this approach has in bridging the competence-performance gap.[11]

[11] Berwick (2011b) proposes a parser incorporating Merge, but illustrates it only for the simple sentence *The guy drank the wine*. Moreover, his illustration derives only the superficial syntactic structure $[_{DP}$ D N$]$ $[_{VP}$ V $[_{DP}$ D N$]]$, omitting the decomposition of *drank* into 'drink' + 'past', not to mention the plethora of functional categories and movements required by minimalist technology. So it is hard to see how this constitutes a minimalist account of performance.

The notion of step-by-step derivation comes from two older traditions. The first is mathematical logic à la Bertrand Russell and Alfred North Whitehead, which constructs proofs algorithmically from a set of axioms. It came into linguistics via Emil Post's 'production systems', which Chomsky adapted for linguistic purposes. The second source of the notion of algorithmic derivation is from traditional grammar, which speaks of morphology DERIVING a complex form from a stem, as though the stem has some temporal or ontological priority. This approach has been variously termed 'derivational', 'procedural', or 'proof-theoretic'.

However, the past twenty-five years have seen the rise of a widespread alternative computational formalism, CONSTRAINT-BASED GRAMMAR, which has been adopted by lexical-functional grammar (LFG), head-driven phrase structure grammar (HPSG), optimality theory, and my own parallel architecture (Jackendoff 2002, Culicover & Jackendoff 2005), among others. Here the rules of grammar simply license parts of a structure—each piece of a well-formed structure has to be checked off or licensed by some rule of grammar. This approach is sometimes called 'representational', 'declarative', or 'model-theoretic'.

For a first approximation, it is possible to convert back and forth between algorithmic and constraint-based formalisms, as was pointed out decades ago by McCawley (1968) and Koster (1978) and more recently by Brody (1995), so the choice might seem like a matter of convenience. For instance, Merge can be construed either way, and if one believes in traces of movement, movement can be construed in constraint-based fashion (as pointed out by Koster and Brody). But at finer degrees of detail the two formalisms diverge (see e.g. Pullum 2007), and in every case I have ever looked at, the constraint-based formalism is superior (see Culicover & Jackendoff 2005 for examples).

I want to raise only one of these differences here: unlike a derivational formalism, a constraint-based formalism lends itself to a direct relation to theories of processing. The current consensus on sentence processing is that it is deeply 'incremental' or 'opportunistic': the processor uses whatever information is available at the moment, whether from phonology, syntax, semantics, or discourse or visual context, to build hypotheses about what is to come in the sentence and the sentence's likely interpretation (Marslen-Wilson & Tyler 1987, MacDonald et al. 1994, Trueswell et al. 1994, Tanenhaus et al. 1995, Gibson 1998, Cutler & Clifton 1999). As observed in Sag 1992 and Sag & Wasow 2011, a constraint-based grammar is ideally suited to this sort of processing, because constraints have no inherent directionality and no inherent ordering. They can be applied starting in phonology, moving through syntax to semantics, as in language perception, or the other way around, as in language production; they can be applied left to right, top down, or bottom up, depending on circumstance.

The competence theory then is a theory of what structures are possible in principle, and the performance theory is a theory of how those structures are built in real time, using the rules directly. Such an approach has been central to LFG parsers since Ford et al. 1982; Jackendoff 2002:Ch. 7 and Jackendoff 2007b show how a parallel architecture grammar (see §8) can be directly implemented in processing. Thus, to the degree that a biolinguistically inspired theory of language should integrate gracefully into a theory of

---

Phillips and Lewis (2012) discuss in some detail the stances that could be taken by advocates of bottom-up derivations. They argue on strong psycholinguistic grounds that a grammatical theory couched in terms of left-to-right derivations is superior, in that it mirrors actual human performance. They also acknowledge, however, that nonderivational theories of grammar of the sort advocated here meet their criteria just as well.

how the brain processes language, the constraint-based approach is not a notational variant of the derivational approach, and it is preferable.

5. Concatenation versus unification. A further problem with Merge is that it combines units by concatenating them into a set. All of the constraint-based frameworks adopt a different basic computational operation, Unification (Shieber 1986), which roughly speaking creates the union of the features of two units, including their hierarchical structure, if any. From a derivational perspective this may seem a bit odd, because two constituents that have been unified cannot always be clearly separated in the output. Thus Unification violates what Chomsky (2005) calls the 'no-tampering condition' on derivations. But from a constraint-based perspective Unification is quite natural: it says that parts of a complex structure can be multiply licensed—the licensing principles do not have to be mutually exclusive. Let me briefly list five cases where this makes a difference.

First is the licensing of an idiom such as *kick the bucket*. Merge (or any previous version of lexical insertion) has to insert the words of the idiom one at a time, and the VP above them has to come from the autonomous phrase structure rules. This loses the generalization that the idiom IS a whole VP; furthermore, the technical details of how the individual words add up to the meaning 'die' inevitably turn out to be rather peculiar.[12] A Unification-based grammar simply lists the idiom as a whole VP that means 'die' (plus perhaps some aspectual information); this VP unifies with the autonomous VP and NP structures. Hence the VP and NP nodes are licensed redundantly, once by the general phrase structure rule, and once by the idiom itself.

(3) Unification of $[_{\mathrm{VP}}$ V NP$]$, $[_{\mathrm{NP}}$ Det N$]$, and $[_{\mathrm{VP}}\ {}_{\mathrm{V}}$kick $[_{\mathrm{NP}}\ {}_{\mathrm{Det}}$the ${}_{\mathrm{N}}$bucket$]]$ =
$[_{\mathrm{VP}}\ {}_{\mathrm{V}}$kick $[_{\mathrm{NP}}\ {}_{\mathrm{Det}}$the ${}_{\mathrm{N}}$bucket$]]$

The ability of Unification-based grammars to account for idioms is not a peripheral issue: English contains tens of thousands of idioms.

A second advantage of Unification-based grammar is that it can check selectional restrictions on a verb's arguments by unifying them with the meanings of the arguments. For instance, in **John drank the apple*, the selectional restriction of *drink* conflicts with the semantic structure of *apple*, so Unification fails and the sentence is anomalous. Now consider the sentence *John drank it*. There is no semantic conflict between *drink* and *it*, so the sentence is acceptable. But in addition, as pointed out by Weinreich (1966), *it* is understood as a liquid, despite having no such features in its lexical representation. How does it get this interpretation? Just checking feature compatibility is not enough: that would merely say that *John drank it* is grammatical. But if the selectional restriction UNIFIES with the semantic structure of the argument, then the underspecified *it* is unified with the semantic feature 'liquid' and the interpretation follows automatically.

[12] For instance, it is sometimes proposed that the meaning of an idiom like *kick the bucket* can be associated with the verb alone, as a special meaning of *kick* when in the context *the bucket*. But it is rarely noticed that this proposal also requires special entries for *the* and *bucket* in the context of this special meaning of *kick*, namely that they bear no meaning. In other words, this approach actually requires the entire structure of the idiom to be encoded, redundantly, in the lexical entries of all three words.

Such an approach at least looks plausible when applied to a VP idiom, where it might make sense to invest the meaning in the syntactic head. It is less attractive for idioms like *day in day out*, *fit as a fiddle*, *down and dirty*, *cut and dried*, and *off the top of my head*, where there is no way to choose which single word carries the content and which words have null meanings.

(4) Unification of $[_{VP}$ V NP], $[_{VP}$ $_V$drink $[_{NP}$ +*liquid*]], and $[_{NP}$ it] =
$[_{VP}$ $_V$drink $[_{NP}$ it, +*liquid*]]

So Unification helps explain a basic aspect of the grammar, where Merge is of no use.[13]

Third, moving beyond syntax, Semitic morphology is easily framed in a Unification-based formalism: the sequence of consonantal roots unifies with the underspecified consonants in the inflectional pattern, and the underspecified vowels in the root unify with the specified vowels in the inflection, to yield a fully specified word, as shown in 5. I am unclear how Merge would apply to this case.

(5) Unification of kVtVv and CaCaCti = katavti

Similarly, in signed languages, a great deal of morphology is simultaneous: a root sign can be altered in its direction, path, or speed of motion to indicate subject and object agreement, intensity, aspect, and so on (Aronoff et al. 2005). The resulting sign is not a set consisting of the root sign plus its alterations; it is an altered sign.[14]

Fourth, consider poetic forms such as limericks or sonnets, which are abstract patterns of stress and rhyme. A text constitutes a limerick or sonnet if it can be unified with this pattern—or alternatively, if it is licensed both by the linguistic rules and by the abstract pattern. Similarly, setting a text to music involves unifying the metrical pattern of the music with that of the text. This is a slightly more complex case, in that the output rhythm of the performance is determined more by the melody than by the speech rhythm (*I …. once had a girl …. or should I say…. she once had me*). So this may require a notion of 'soft Unification', somewhat akin to prioritized constraint satisfaction in optimality theory. In any event, Merge gives us no purchase at all on these phenomena: combining texts with music is not simply creating sets, each of which consists of a note and a word.

Finally, the visual system involves many distinct brain areas that specialize in shape, color, motion, and so on (Crick 1994, Koch 2004). Somehow the computations performed by these areas are combined into a unified visual percept. Without an explicit theory of visual representation, one cannot make a very strong argument about how this takes place. But it seems more plausible that the features are collected into a single perceived entity, in the manner of Unification, than that a single visual percept is a binary tree structure created by Merge. That is, Unification again scales up gracefully to phenomena outside of language.

In the face of the last three examples, one may be tempted to say, 'Well, Merge is only for syntax'.[15] But such a reply flouts the biolinguistic desideratum of integrating linguistic theory into the larger architecture of the mind/brain and of reducing the scope of the narrow language faculty.

I should make clear that Unification alone cannot create constituent structure: it only creates a Boolean combination of preexisting features and structures. In order to build structure, one needs a skeletal constituent structure that can be unified with two or more

[13] This argument comes from Jackendoff 1990, where Unification is called Argument Fusion.

[14] These phenomena might be treated as a function of Spell-Out, for example, using mechanisms of distributed morphology (Halle & Marantz 1993). But such a treatment would be a completely different mechanism from Merge, which can only create the unit {*root*, *inflection*}.

[15] Hornstein and Boeckx (2009:89) make such an argument: 'it is reasonable to suppose that an operation like Merge, one that "puts two elements together" (by joining them or concatenating them or comprehending them in a common set), is not an operation unique to FL. It is a general cognitive operation, which when applied to linguistic objects, we dub "Merge".' In other words, they acknowledge that recursion may be found elsewhere in cognition, but they do not call it Merge. Thus they reduce the notion that Merge is unique to language from an empirical hypothesis to a tautology.

items. Such a skeleton is of course already richly present in cognition: the part-whole schema. One formal realization of this schema is a set $\{x, y\}$ with variable elements x and y as parts. This can be unified with specific elements A and B to form the set {A, B}—in effect the output of Merge. Similarly, a linearly ordered constituent [A^B] can be licensed by the unification of A and B with a linearly ordered schema [x^y], which is also ubiquitous in nonlinguistic cognition. Thus the effects of Merge can be constructed from Unification and one of these schemas.

One might say then that these schemas are nothing but constraint-based counterparts of Merge, and this would be partly correct. However, Merge per se is too limited, in that it countenances only the substitution of whole elements for atomic variables. It fails in the examples of Unification above, because each of them requires combining STRUCTURED elements (such as *kick the bucket*) with STRUCTURED variables (such as [$_{VP}$ V NP]) in such a way as to preserve the internal structure of both. So perhaps Unification ought to be viewed as a richer version of variable instantiation—but one that appears to be characteristic of cognitive computation.

Let me sum up so far.

- Recursion is not the defining characteristic of language; it is found everywhere in higher cognition.
- Binary Merge is not rich enough to capture the varieties of recursion found elsewhere in mental structures.
- A grammar stated in terms of constraint satisfaction is preferable to one stated in terms of algorithmic generation.
- The brain's characteristic combinatorial operation is Unification rather than Merge.

Although these points all violate the preconceptions of 'efficient computation' that gave rise to the notion of Merge, each of them is supported both within the theory of language and elsewhere in the mind/brain. Thus, they should be seen as welcome advances in pursuing a biolinguistic agenda, at once accounting for a wider range of phenomena in language and bringing the language faculty closer to graceful integration with the rest of the mind/brain.

6. THE ROLE OF PHONOLOGY. Let us return to Hauser, Chomsky, and Fitch's hypothesis that FLN comprises 'only the core computational mechanisms of recursion as they appear in narrow syntax [i.e. Merge] and the mappings to the interfaces' (2002:1573). So far we have been examining the status of Merge as the 'core computational mechanism' of language. Next let us consider the 'mappings to the interfaces'. One of these interfaces is the 'sensory-motor interface', which is identified with the auditory system and the part of the motor system concerned with speech production. Where does phonology fit into this picture?

Chomsky (1965:16, 17, 75, 198) explicitly assumes—without argument—that syntax is the sole source of combinatorial structure in language, and that phonology and semantics are 'interpretive'. This assumption persists through all of his subsequent frameworks, including the minimalist program. Hauser, Chomsky, and Fitch put it this way:

> FLN takes a finite set of elements and yields a potentially infinite array of discrete expressions … . Each of these discrete expressions is then passed to the sensory-motor and conceptual-intentional systems, which process and elaborate this information in the use of language. Each expression is, in this sense, a pairing of sound and meaning.[16] (2002:1571)

[16] Note the locution 'these expressions are passed', which again implies step-by-step derivation.

This seems to imply that phonology belongs to the sensory-motor interface. But phonology definitely is not sensory-motor, an encoding of sound per se. It is deeply cognitive, and quite abstract with respect to both raw auditory input and motor output (Liberman 1993).

Neither is phonology part of narrow syntax. Phonological units are not the same as syntactic units, and they follow their own principles of organization. Syllable boundaries do not always coincide with morphosyntactic boundaries; prosodic constituents cannot be identified with syntactic constituents; stress follows rhythmic principles that have nothing to do with syntax (see Jackendoff 2002:§§5.3 and 5.4 for examples). Common practice in phonology since the 1970s treats the tiers of phonological structure as generative systems independent of syntax. They may not be as elaborate as syntax, but they are combinatorial systems nevertheless, with their own characteristic units.

Mainstream syntax since the 1970s has rarely addressed in any detail the place of phonological structure in the grammar. Most recently, Chomsky (2010; see also Berwick & Chomsky 2011) speaks of morphology and phonology—including the establishment of linear order—as 'externalization', the 'imperfect' mechanisms necessary to make syntax compatible with the sensory-motor interface, and a part we should pare away in order to get at the real essence of language. He tentatively assigns externalization to FLB, and says, 'Solving the externalization problem may not have involved an evolutionary change—that is, genomic change. It might simply be a problem addressed by existing cognitive processes, in different ways, and at different times' (Chomsky 2010:61). Boeckx (2011:212), following Chomsky's line, speaks of externalization as 'akin to an afterthought, or appendix'.

Yet, as phonologists will attest, phonology is deeply linguistic. And although other species have combinatorial systems of vocal signals, none has the fine-scale articulation of phonology into syllables composed of segments, which in turn are composed of distinctive features—all of this overlaid with a prosodic pattern plus nonlinguistic signals such as tone of voice. So on the face of it, an important part of FLN is the ability to interpret a continuously varying acoustic signal in terms of this discrete or digitized representation; and an important part of language acquisition is the ability to induce the local digitized system from the acoustic input. None of this follows automatically from Merge as the basic element of FLN, nor from audition or motor control.

The discreteness of phonological units leads to a further point. The biolinguistic approach seeks to derive properties of language from what is called 'natural law' or 'third factor considerations', so that they are somehow not a burden on natural selection. Consider phonology in this light. As various people have argued (e.g. Hockett 1960, Nowak et al. 1999, Jackendoff 2002), the only way to develop a large vocabulary and still keep vocal signals distinguishable is to digitize them. This is a matter of signal detection theory, not of biology, and it is also responsible for the fact that we all have digital rather than analog computers. So we might want to say that the digital property of phonology comes by virtue of 'natural law', in (I think) precisely the sense Chomsky intends.

But notice that this does not take the burden off of evolution. The brain still requires the capability to digitize the sounds of language, and this requires computational power. This computational power in the brain has to have arisen somehow in our history. Other species have not discovered it in the course of THEIR evolution.[17] It does not matter

[17] Notice that this argument goes beyond asserting that humans are capable of categorical auditory perception, which has turned out not to be a distinctive human property. Here I am concerned with the ability to learn to use a phonological system, on the basis of hearing continuous input that is acoustically different for every individual with whom one interacts. The computational difficulty of this task is evidenced by the lim-

whether this ability arose in our species because of its affordance of better communication, or for some totally unrelated reason (as Chomsky puts it, as 'a problem addressed by existing cognitive processes'). The point is that the digitization of sound, based on the acoustic properties of the vocal tract, is one of those 'third factor' things, like efficient computation, that the biolinguistic approach claims to be seeking. And yet it is still a biological capability that had to evolve.

The situation parallels streamlining in dolphins. The advantage of streamlining is a consequence of natural law. But dolphins still had to evolve this shape, presumably through natural selection. In other words, in these cases, natural law does not eliminate the role of natural selection, as (I think) Chomsky is suggesting; rather, it shapes the adaptive landscape within which natural selection operates.

7. The role of semantics. Examining further the minimalist program's 'mapping to the interfaces', let us turn to the conceptual-intentional interface. From its earliest days, mainstream generative grammar has presumed that semantic combinatoriality is read off of syntactic structure (and this assumption goes back even further, at least to Tarski 1956). Of course, semantics cannot be read directly off the linguistic surface. The genius of generative grammar was to posit covert levels of syntax at which this direct linking is possible: first deep structure, later logical form (LF). But to my knowledge, there has never been any significant exploration within mainstream generative grammar of the other side of the conceptual-intentional interface—the nature of the human concepts that language expresses.[18] Hinzen as much as admits this when he says of semantic theory within biolinguistics: 'Truth be told, there is no field here with a state of the art to report. Do we know even the rough outlines of what an explanatory research program on the origins of human semantics looks like?' (2008:348).

Well, actually we do. The human conceptual system, in addition to being the endpoint of language comprehension and the starting point of language production, also has to serve nonlinguistic purposes. It has to encode our nonlinguistic understanding of the world, making it possible for us to perceive the world, draw inferences from what we perceive, and formulate actions.

We share many of these functions with nonlinguistic primates. Their natural vocal communication systems are for the most part not at all combinatorial (with a few very rudimentary exceptions; Zuberbühler 2002, Arnold & Zuberbühler 2008). Nevertheless, like humans, other primates individuate, categorize, and manipulate objects; they too find their way around in space; and they too engage in elaborate social interactions with other individuals. They may not do all of these things with the same sophistication we do (although, from another perspective, swinging from trees is pretty nifty), but it makes sense to reiterate that evolution typically does not throw things away—it just builds on them. So the abilities of other primates set a lower bound on what the human conceptual-intentional system can do. Here, unlike syntax and phonology, we have strong and revealing comparative evidence for the evolutionary roots of the human system (Köhler 1927, Byrne & Whiten 1988, Cheney & Seyfarth 1990, 2007, Wrangham & Peterson 1996, Povinelli 2000, Tomasello 2000, among many others).

My reading of the literature is that primate cognition (unlike primate communication) is richly combinatorial. Among the combinatorial concepts demonstrated by social

ited success of automatic speech recognition after over half a century of research. See Pinker & Jackendoff 2005 for more discussion.

[18] Approaches from formal semantics do not have the goal of explicating human concepts, so they cannot be taken as accounts of the conceptual-intentional interface. See Jackendoff 2002:Chs. 9 and 10 for elaboration.

primates are kinship relations ('X is kin of Y') and dominance relations ('X is dominant to Y'). As Cheney and Seyfarth (1990) demonstrate experimentally, vervet monkeys do not understand these relations just with respect to themselves ('X is my kin', 'I am dominant to Y'), but also with respect to all pairs of individuals in their group. So these relations are two-place predicates whose variables are instantiated differently for each pair of individuals the monkey knows. Moreover, kinship relations can play a role in more complex inferences. One striking case is 'redirected aggression': after individual A has a fight with individual B, it is more likely that C, who is kin of B, will either attack or show affiliative behavior toward D, who is kin of A. In order to be able to generate such behavior, C must take into account the history of A's and B's actions, plus the two relevant kinship relations that let C and D count as 'stand-ins' for A and B, either in retaliation or in reconciliation. Such behavior requires a combinatorial system.

To be sure, observing and categorizing primate behavior is not the same as developing a formal theory of their mental representations (see Flack et al. 2011 for an attempt). In particular, it is hard to determine whether primate cognition permits unlimited hierarchical embedding, that is, whether it is structurally recursive—and if not, whether this is because of limitations of the representations themselves or because of limitations on, say, working memory. In any event, whatever combinatoriality there is cannot be a result of syntactic Merge: monkeys do not have language.

There is also mounting evidence that prelinguistic babies have combinatorial concepts. For example, Gergely and Csibra (2003) show that five-month-olds recognize some sorts of goal-directed action, noticing that an actor could take a shortcut when an obstacle to its movement is removed, and being surprised when the actor fails to take it. Gordon (2003) shows that ten-month-olds watching an event can tell whether it is a two-character event such as hugging or a three-character event such as giving, even when the two-character event includes an inert third character. Song, Baillargeon, and Fisher (2005) show that 13.5-month-old infants can attribute to an actor a disposition to perform an action on a novel object, based on its spatial context and its affordances for motion. Again, it is not easy to characterize these abilities formally, but they are suggestive of a fair degree of combinatoriality, well prior to the onset of combinatorial language.[19]

Combinatorial structure within concepts is also necessary to explain the semantic connections among sentences in discourse and narrative, which, as observed in §3, are expressed only by sequence, not by syntactic structure, yet have a rich inferential structure (see e.g. H. Clark 1996, Asher & Lascarides 2003 for the formal complexity of discourse and narrative connections).

These considerations lead to the hypothesis that human COGNITION is combinatorial, but not by virtue of being derived from syntactic combinatoriality. Rather, an autonomous and evolutionarily prior system is responsible for the form of concepts. In humans this system is hugely enhanced over that of other primates, partly through developments internal to the system itself, and probably partly through its connection with language. But the language system gets the fundamentals of the conceptual system for free.

[19] One might protest that the babies have linguistic competence but just cannot achieve performance, or that they have language at the conceptual-intentional interface, but have not yet achieved externalization. But this begs the question, since at these ages there is no linguistic evidence, even in their language comprehension, for combinatoriality of such sophistication.

My own work on CONCEPTUAL STRUCTURE (Jackendoff 1983, 1987a, 1989, 1990, 2002; see also Pinker 1989, 2007) has been built on these premises. A major concern has been the language used to express spatial relations and spatial event structure (Jackendoff 1987b, 1996a, Landau & Jackendoff 1993). Following Jeffrey Gruber's pioneering study (Gruber 1965), Jackendoff 1972, 1976, and 1983 examine the extension of spatial language to nonspatial domains, a phenomenon also addressed by COGNITIVE GRAMMAR (e.g. Lakoff 1987, Langacker 1987, Talmy 2000). Jackendoff 1991a and 1996b study the relation between mereology of objects and event structure/aktionsarten. More recent work (Jackendoff 1999, 2007a) studies the language of social cognition. Many of these domains are aspects of our understanding of the world that we share to some degree with nonlinguistic primates.

Jackendoff 1983 placed a further criterion on the theory, called the GRAMMATICAL CONSTRAINT: one should expect a strong correspondence between syntactic and semantic form. Over the years, however, as the theory became more rigorous and fine-tuned, it became clear that optimal correspondence is only a default. There prove to be numerous mismatches between syntactic and semantic embedding, for instance, coercions and meaningful constructions, hence a less direct syntax-semantics interface (Jackendoff 1990, 1991b, 1996b, 1997a; see also Goldberg 1995, Pustejovsky 1995). Culicover & Jackendoff 2005 shows that many phenomena that are central to syntactic theory, such as control, sluicing, and certain aspects of long-distance dependencies, cannot be fully accounted for in terms of covert syntax such as LF; we formulate detailed alternatives that depend in part on noncanonical correspondence between syntax and conceptual structure. So far as I know, there are no detailed responses to these arguments in the mainstream literature, nor to related arguments such as those of Fillmore, Kay, and O'Connor (1988), Goldberg (1995), Pustejovsky (1995), and many others.

Berwick and Chomsky question the justification for an independently generative conceptual structure: 'Positing an independent, recursive, "language of thought" as a means to account for recursion in syntax leads to an explanatory regress as well as being unnecessary and quite obscure' (2011:38, n. 6). Taking their points in reverse order:

- 'Obscure': Conceptual structure falls under their characterization of 'language of thought', but it is far from obscure. It has been explicated in linguistic, philosophical, and psychological terms, and it has been worked out in considerable formal detail. In particular, it has been distinguished quite clearly from Fodor's (1975) conception of 'language of thought' (Jackendoff 1983, 1987a, 1989, 1990, 1991a,b, 1996a,b,c, 1997a, 2002, 2007a, 2010, Culicover & Jackendoff 2005, Pinker 2007). Berwick and Chomsky cite none of these sources.
- 'Unnecessary': As just argued, conceptual structure (or something like it) is far from unnecessary: it is necessary not only for the theory of language, but also in order to explain nonlinguistic cognition by adult humans, children, and nonhuman primates.
- 'Explanatory regress': Here Berwick and Chomsky are alluding to the question of where recursion in thought could have come from, if not from syntax. But one can ask the very same question about the numerous examples of recursion illustrated in §3, none of which have anything to do with language. Did recursion evolve in each of them independently, and if so how? Or did one of them have some sort of evolutionary primacy? I do not think we have any idea, but the question is not devoid of interest. The same goes for the origin of recursion in conceptual structure. The 'explanatory regress' should be engaged, not brushed aside.

- ‘As a means to account for recursion in syntax’: Conceptual structure actually does not account for recursion in syntax, since nonlinguistic organisms can have conceptual structure without syntax. However, recursive conceptual structure accounts for why recursive syntax is useful, namely for EXPRESSING recursive thoughts.

Within the biolinguistic community, there have been some attempts to develop a theory of the conceptual-intentional interface. For instance, Larson (2011) proposes to correlate subordinate clauses in syntax with domains of intentionality in semantics. He manages to implement such a correlation only by positing null elements in syntax that are reminiscent of derivations in the long-abandoned theory of generative semantics (e.g. deriving *want a bicycle* from *want ~~to have~~ a bicycle*). Furthermore, he does not address the question of why the clausal complements of predicates like *be lucky to VP* and *manage to VP* happen NOT to have intentional semantics. The reason, obviously, is that they are not verbs that express a state of mind—a semantic, not a syntactic, fact.

Berwick (2011a) suggests (following Chomsky) that other primates seem to have a conceptual-intentional system, perhaps even ‘lexical items’, but seemingly lack Merge. This seems to me a misconstrual of the situation. As discussed above, other primates clearly have concepts, many of which are combinatorial. What they lack is ‘externalization’: they cannot express their thoughts. Hinzen (2011) proposes that in fact combinatorial syntax directly generates conceptual structure, so there is no need to ‘map’ to the semantic interface; he has in effect reinvented generative semantics. Noticing that nonhuman primates may well have some combinatorial concepts, he suggests (as far as I can follow) that what makes human combinatoriality special is that it is ‘truth-evaluable’. But Hinzen couches this proposal in terms of ‘expressions’ being truth-evaluable, which presumes there are ‘expressions’, that is, language. So of course human combinatoriality is special in this respect: humans evaluate SENTENCES for truth, and other primates do not have sentences.

Chomsky’s own discussions of the conceptual-intentional interface (e.g. 2005, 2010, Berwick & Chomsky 2011) seem to be confined to two themes. The first theme, present in his work for decades (e.g. Chomsky 1975), is that a theory of meaning cannot be based on the notion of reference to the world, as assumed by standard philosophy of language. Rather, it must be based on the character of human conceptualization—on the structure humans impose on their perceptions of the world. This theme is central to CONCEPTUAL SEMANTICS as well (Jackendoff 1983:Ch. 2, 1987a:Ch. 7, 2002:Ch. 10). Chomsky’s second theme, more recent, is that the structure of syntax is molded more closely to the conceptual-intentional interface—to thought—than it is to the ‘externalization’ of the sensory-motor interface. Like Berwick (2011a), he speculates that although other primates may have simplex concepts (or ‘conceptual atoms’), it took a mutation in one of our hominid ancestors to be able to apply Merge to such concepts to create human thought. This hypothesis seems just wrong: it does not explain the combinatoriality of primate thought, as evidenced by the richness of their behavior.

In addition, none of these biolinguistically inspired approaches makes any serious attempt to deal with the sort of syntactic and semantic detail that has been characteristic of work in conceptual semantics; nor is there any contact with the voluminous psychological literature on concepts. The semantic theory remains entirely programmatic.

8. THE PARALLEL ARCHITECTURE AND ITS LEXICON. The last two sections lead to the conclusion that phonology and semantic/conceptual structure are not derived from syntax, as taken for granted in mainstream generative grammar. Rather, the outcome is what I have called a ‘parallel architecture’ (Jackendoff 1997a, 2002), in which phonology, syntax, and semantics/conceptual structure are independent generative systems linked by interfaces—principles that license correlations between pieces of structure in two or more distinct domains; see Figure 6. A well-formed sentence has well-formed structures in all three domains, plus well-formed links between the domains.[20]

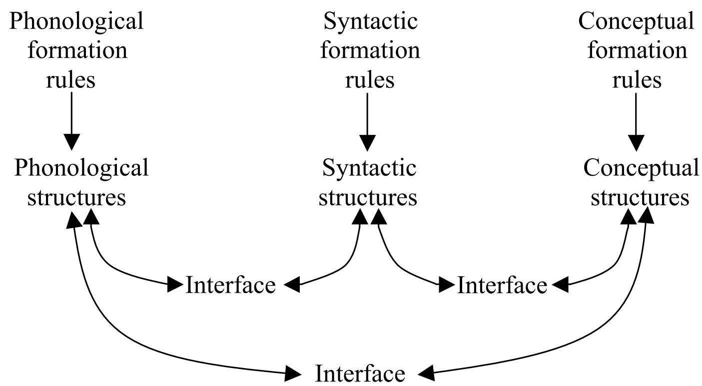

FIGURE 6. The parallel architecture.

In this framework, a word can be thought of as part of the interfaces linking the three structures: a stereotypical word is a little idiosyncratic rule that says THIS piece of phonology can be linked to THESE syntactic features and THIS piece of conceptual structure. This treatment of words integrates altogether naturally with experimental work on lexical access in perception and production (Jackendoff 2002, 2007b). By contrast, the minimalist program's treatment of words, in which they are selected from the lexicon to form a numeration, which in turn is used to build syntactic structure, bears no relation whatsoever to the demands of language processing.

Not all words correlate a full triplet of structures. Words like *hello*, *upsy-daisy*, *gee whiz*, *feh*, and *dammit* have phonology and meaning, but do not participate in syntactic combination (except in paratactic combinations like *Hello, Bill* and quotative contexts like *the word <u>hello</u>* and *she said 'hello'*), so there is no reason for them to have syntactic features. Grammatical words like the *do* of *do*-support, expletive *it*, and the *of* in *N of NP* have phonology and syntax but no semantics. If one believes in *PRO*, *pro*, and/or traces, these are items with syntax and semantics but no phonology. Finally, nonsense phrases such as *fiddle-de-dee*, *inka-dinka-doo*, and *brrr-raka-taka* have only phonological structure and are used just to fill up metrical space in songs.

Of course, words are not the only interface rules. The interface between syntax and conceptual structure also includes canonical principles for argument realization such as the thematic hierarchy; the interface between syntax and phonology includes the principles that establish possible correspondences between syntactic structure and prosodic contours. In addition, there may be a direct interface between phonology and semantics, which establishes the relation between focal stress and prosodic contour on the one hand and information structure on the other.

[20] LFG, autolexical syntax (Sadock 1991), and role and reference grammar (Van Valin & LaPolla 1997) are also organized as parallel architectures, with somewhat different configurations of components. It is possible to conceptualize HPSG as a parallel architecture, though this is not the way it is generally construed.

In addition, the interface rules include idioms (*kick the bucket*, *day in day out*), which link syntactic complexes to noncompositional semantics. They also include constructional idioms such as those in 6, which distort the normal linking between syntax and semantics, adding meaning that is not expressed by any of the words.

(6) a. Bill belched his way down the street. (Jackendoff 1990, Goldberg 1995)
   $[_{VP}$ V [*pro's* way] PP] = 'go PP while/by V-ing'
 b. Joe knitted the morning away. (Jackendoff 1997b)
   $[_{VP}$ V NP away] = 'spend NP[*time*] V-ing'
 c. The bus rumbled around the corner. (Levin & Rappaport Hovav 1990, Goldberg & Jackendoff 2004)
   $[_{VP}$ V PP] = 'go PP in such a way to make a V-ing sound'
 d. The more you eat, the fatter you get. (Culicover & Jackendoff 2005)
   [*the more S, the more S*]
 e. One more beer and I'm leaving. (Culicover 1972)
   $[_{Utterance}$ NP and S][21]
 f. How about some lunch? (Culicover & Jackendoff 2005)
   $[_{Utterance}$ how about XP?]
 g. that travesty of a theory (Asaka 2002, Booij 2002)
   [Det N_1 of $[_{NP}$ a N_2]] = 'Det N_2, which is a N_1')
 h. student after student (Jackendoff 2008)
   $[_{NP/AdvP}$ N-P-N]

Within the parallel architecture, such constructional idioms can be encoded with a formalism fully parallel to that for words and idioms (Culicover & Jackendoff 2005, Jackendoff 2010) (this is also true in construction grammar). Like idioms, their syntax is listed in the lexicon as a phrasal structure rather than an X^0 category. Unlike idioms, only some parts of this syntax are linked to phonology (e.g. *way* in 6a, *away* in 6b); other parts are variables that must be satisfied (e.g. V and PP in 6a). The syntactic structure is also linked to a conceptual structure, but in noncanonical fashion. In particular, parts of the conceptual structure, such as the sense of motion in 6a and 6c, receive no syntactic or phonological realization at all. These constructions are part of what one has to learn in learning English.

In addition, phrase structure rules and more general principles such as the head parameter can be stated in the very same formalism: they are phrasal structures whose parts are all syntactic variables, unlinked to phonology and semantics (Culicover & Jackendoff 2005). Thus they are a sort of syntactic counterpart of the words that have only phonology.

The upshot is that the parallel architecture's grammar has no strict formal division between words and rules.[22] Rather, they are at opposite corners of a multidimensional array of lexical items, with many intermediate cases in between. Within this array, lexical entries differ in whether they deal with words or phrases, whether they link multiple levels, how phonologically specific they are, and how many variables they contain (see Jackendoff 2002, Culicover & Jackendoff 2005 for much further discussion, and Jackendoff

[21] The constructions in 6e,f are characterized as Utterance rather than S because they do not embed. Other analyses are possible.

[22] This conclusion has been reached independently in Williams 1994, HPSG (Pollard & Sag 1994), cognitive grammar (Langacker 1998), and radical construction grammar (Croft 2001). In a curious way, one might see Marantz 1997 as arriving at the same conclusion, though with quite a different theoretical twist.

2007b for consequences of this conclusion for the theory of sentence processing). In a sense, then, the parallel architecture realizes Chomsky's occasional speculation that a language might consist of just a lexicon plus Merge, except that the lexicon is far richer than his 'conceptual atoms', and the basic combinatorial principle is Unification rather than Merge.

Some practitioners of the minimalist program (e.g. Phillips & Lau 2004, Marantz 2005) have commented that a theory with only one 'generative engine' is preferable on first principles. This is taken to be sufficient to dismiss the parallel architecture, which has three generative engines. But the minimalist program addresses only syntax: it has no theory of phonology, no (generally accepted) theory of semantics, and no (generally accepted) theory of the lexicon. So it should be no surprise that it needs only one 'generative engine'. My view is that first principles of simplicity have to give way in this case to empirical adequacy.[23]

Language is not the only faculty that requires a parallel architecture. When Fred Lerdahl and I began our research on a generative grammar for tonal music (Lerdahl & Jackendoff 1983), we attempted to develop a 'single-engine' derivational type of grammar, along the lines of the linguistic grammars of the 1970s. But we quickly found we could not make it work, and we ended up improvising what amounts in present-day terms to a parallel constraint-based grammar with a lot of default principles: four independent structures that have to align with each other in various ways (three of these structures were illustrated in Fig. 3 above).

More generally, a parallel architecture for language integrates gracefully into the rest of the mind/brain, which is full of independent faculties that have to interface with each other. We know this from the structure of the brain, as has already been pointed out. But we also can see this in terms of a sort of first-principles analysis of the tasks humans do. Figure 7 lays out some of the general architecture of human capacities.

Crucial to this architecture is the level of SPATIAL STRUCTURE, which encodes our understanding of the geometrical organization of the physical world (Jackendoff 1987b, 1996a, Landau & Jackendoff 1993). Its structures may be thought of as suitably amplified versions of the Marr 3D model illustrated in Fig. 2. Spatial structure must integrate information derived from vision, touch, and proprioception (and probably auditory localization as well). Again, a theory worked out in terms of constraint satisfaction and Unification seems appropriate here, as opposed to one based on step-by-step derivation via Merge. Furthermore, each of the perceptual processes requires one or more of its own 'generative engines', leading inevitably to a parallel architecture mediated by interface components. The formulation of action in the world, involving structured action plans along the lines of Fig. 4 above, also calls upon spatial structure: action requires moving one's body and acting on objects in the environment in cognizance of their spatial configurations. And muscle activations cannot be derived 'interpretively' from visual representations.

[23] As Newmeyer (2003) observes, we have been here before. Postal (1972) asserted that, on first principles, generative semantics was superior to interpretive semantics, because it required only one component, namely syntax. Chomsky (1972) responded (without, however, explicitly citing Postal), observing that the number of components is not the only consideration: theoretical elegance, although important, cannot trump empirical adequacy of the theory. I would add now that in the version of the biolinguistic perspective advocated here, the theorist is also responsible for the integration of the theory into a larger theory of the mind/brain. (But it seems that among the biolinguists, the tune has changed: the 'Galilean perspective' licenses disregarding adequacy of description in favor of 'first principles'.)

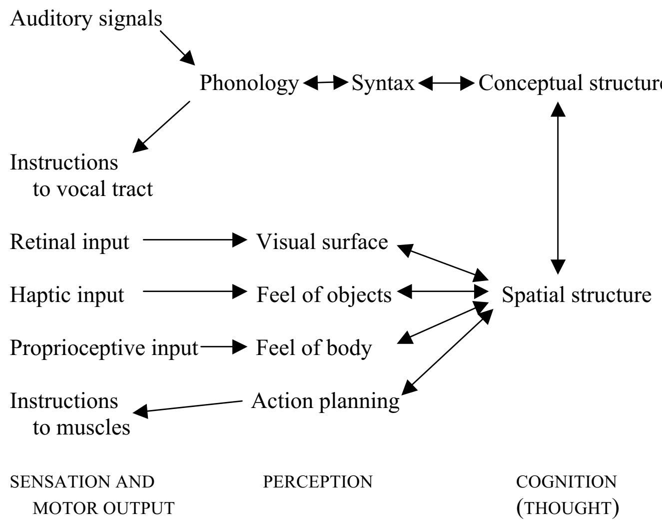

FIGURE 7. Overall architecture of the mind/brain.

Language integrates with these other faculties in much the same way as they interface with each other. The interface between conceptual structure and spatial structure in Fig. 7 enables the processes that construct utterance meaning, encoded primarily in conceptual structure, to combine information coming from phonology and syntax with so-called 'pragmatic' information derived from perception (for instance pointing, as in deictic anaphora), again by Unification. It is through this connection between conceptual structure and spatial structure that we can talk about what we see. Moreover, visual input has been demonstrated to play a role in sentence processing (Tanenhaus et al. 1995), presumably also via the conceptual-spatial interface. In other words, the means by which we integrate perceptual information with language is no longer a mystery. It is only, as Chomsky (1975) would put it, a problem.

To sum up the last three sections: a parallel architecture stated in terms of constraint-based Unification does not satisfy the criterion of 'perfection' in the biolinguistic vision of the language faculty. However, it allows a better fit to a large range of linguistic phenomena that remain unaddressed in the minimalist program. At the same time, it better satisfies the criterion of graceful integration, in that (i) other faculties of mind also appear to have a parallel organization, and (ii) a language faculty organized in terms of the parallel architecture itself integrates naturally with the conceptual and perceptual systems.

9. EVOLUTIONARY CONSIDERATIONS: THE ROLE OF LANGUAGE IN THOUGHT, VIA INNER SPEECH. From this perspective on the language faculty and its place in the mind/brain, let us consider a speculation of Chomsky's about the evolution of the language faculty. For decades he has asserted (e.g. Chomsky 1975) that the basic function of the language faculty is not communication, and that language is in fact poorly designed for communication. More recently (e.g. Chomsky 2000, Berwick & Chomsky 2011), he has added an argument to the effect that we use language mostly to talk to ourselves, in an inner monologue or Joycean stream of consciousness, and he suggests that if language is 'designed' for anything, it is for thought. The 'externalization' of language as sound, he suggests, was a later stage in the evolution of the language faculty. As we saw in §6, he further suggests that 'externalization' may not even have required a cognitive innovation specific to language.

This line of reasoning ignores two basic observations. First, we have already seen that nonlinguistic organisms also have forms of combinatorial thought. They do not, however, have language, so they cannot be using inner speech to think. Second, consider the character of the inner monologue. It comes as a string of words, with language-particular phonology, stress, word order, agreement morphology, and at least some semblance of syntax. It is 'in a language': we can ask a bilingual, 'Do you think in French or in English?'. This is rather odd, because thought CANNOT be language-particular: the thought being expressed is what is supposed to remain (relatively) constant as we translate from one language to another. In other words, on the one hand thought exists without language, and on the other hand the inner monologue displays all the symptoms of what Chomsky calls 'externalization', except for the fact that one does not actually speak it.

My analysis of this situation (Jackendoff 1987a, 1996c, 2007a, 2012) is that conceptual structure, that is, the formal structure of the thought conveyed by a sentence, is almost completely unconscious, and that what we EXPERIENCE as our inner monologue is actually the phonological structure linked to the thought. We are aware of our thinking because we hear the associated sounds in our head. This analysis immediately explains the fact that we seem to be 'thinking in a language'. More subtly, it is responsible for the widespread preconception that language is necessary for thought, and therefore, since animals do not speak, they cannot think (and on some views are not conscious). It also leads to the widespread philosophical presumption that propositional thought must be linguistically structured.[24]

This analysis of the way in which we are conscious of our thought is profoundly at odds with the usual view that consciousness is a high-order human cognitive capacity that is deeply linked with thought. Nevertheless, I think it is an accurate observation of the experience. Consciousness is linked not with THOUGHT, but with PHONOLOGY and the other forms of perception and perceptual imagery shown in Fig. 7. That is, one has the experience of thought through phonological, visual, haptic, proprioceptive, and perhaps auditory imagery. As far as I know, this observation has not been made by any of the numerous philosophers and neuroscientists investigating consciousness. This is in part because they treat 'language' as a single cognitive function, failing to recognize the crucial distinction between phonology, syntax, and semantics. As a result, they think of phonological imagery as simply 'linguistic imagery'—a characterization that is too coarse to distinguish between a 'mere' sound pattern and the thought it expresses.

If this account is on the right track, our inner monologue is possible only by virtue of having WORDS—and we have to learn words in the context of a communicative speech community.[25] We can only talk to OURSELVES by virtue of having learned to talk to oth-

[24] For example, Hinzen says (2011:424) that Pinker and Jackendoff (2005) 'presuppose … that we think propositionally, in the format of language'. He assumes that Pinker and I identify a 'proposition' with a 'sentence'. We do not. When we use the term 'proposition', we mean a certain type of conceptual structure. A sentence is a linkage of this type of conceptual structure with language-specific syntax and phonology. But, as laid out in the previous two sections, conceptual structure can function in reasoning without being linked to language.

[25] This statement should be qualified by the fact that deaf children who are not exposed to a signed language develop systems of home sign (Goldin-Meadow 2003), which involves inventing words. However, they are doing so in the context of communicating with others; I know of no evidence in the literature that they first use it to talk to themselves. Moreover, although the grammar of home signs may involve some constituent structure, there is no evidence of structural recursion. One might imagine Chomsky responding that

ers. To be sure, speech and inner speech can ENHANCE thought, by making it possible for us to attend to the combinatorial structure of thought, through its phonological proxies (Jackendoff 1996c, 2007a, 2012). But inner speech is not the thought itself (or the thought alone), as Chomsky appears to be claiming.

Going back to evolution, I conclude that hominids could not have had an inner monologue until they had a language with phonology in it. I further conclude that the advantages of language for thought are a byproduct—albeit a substantial byproduct—of its advantages for communication, not the other way around, as Chomsky claims.

10. FURTHER EVOLUTIONARY CONSIDERATIONS: PROTOLANGUAGE. Let us now go back to the major themes of generative grammar stated at the outset.

- KNOWLEDGE OF A LANGUAGE consists of (a) a set of stored structures in the phonological, syntactic, and conceptual domains, plus (b) stored links among pieces of those structures, plus (c) the capacity to combine them by the operation of Unification.
- ACQUISITION OF A LANGUAGE requires at the very least a predisposition (a) to treat vocal signals (or gestures) as linked to concepts (i.e. to treat them as symbolic), (b) to organize vocal signals in terms of phonological structure, and (c) to learn to organize sequences of symbols into a syntactic structure that helps signal the relations among the associated concepts.
- The BROAD FACULTY OF LANGUAGE includes (among other things) the basic combinatorial structure of thought, the use of recursion in syntax, and the use of Unification as the fundamental combinatorial operation.
- The NARROW FACULTY OF LANGUAGE includes the essential architecture of syntax and phonology, the interface between them, and their interfaces to auditory and motor representations and to conceptual structure. And since words are interface rules, the ability to learn words is also part of the narrow faculty.

This architecture puts considerably more machinery between the sensory-motor capacities and the conceptual capacity than does the minimalist program; I have argued here that it is all necessary for good empirical reasons. But it does raise the problem of how the human brain evolved to be able to instantiate these structures; it is unlikely that such a major structural difference is the product of a single mutation. Jackendoff 2002:Ch. 8 suggests a sequence of innovations, each of which would have improved the expressive power of a communicative system over the previous step, culminating in full human language. This sequence of steps provides a route that evolution COULD have taken, in principle answering the 'What good is half a wing?' argument. Some of these steps might have been cultural innovations, but some—at least the advent of phonological and syntactic structure—appear to require innovations in mental representation that have the hallmarks of genetically guided rewiring of the brain.

A prominent step in this sequence is what Bickerton (1990) calls PROTOLANGUAGE—a system in which words can be concatenated indefinitely, but in the manner of a shopping list (i.e. a '*W**' grammar). A protolanguage has phonology, but it has little or no constituent structure, no recursion, no subordination, and no functional categories. The semantic relations among words are signaled by word order alone, using default princi-

---

the absence of recursion is a matter just of externalization, not of the childrens' internal language, which is fully recursive. But this would beg the question, by defining thought as requiring syntactic rather than conceptual combinatoriality, and therefore concluding that SYNTAX is present when there is evidence only for the presence of THOUGHT.

ples such as agent first, focus last, and modifier adjacent to modified (see also Klein & Perdue 1997). Since all combinatoriality is driven by phonology and semantics, there is no need for words to be categorized by parts of speech.

Bickerton argues that the symptoms of protolanguage turn up in modern humans in situations where for various reasons full language is not present, such as early stages of child acquisition, conditions of late acquisition such as pidgins, and perhaps in some cases of aphasia. To this list we can now add other aspects of late second language acquisition (Klein & Perdue 1997), the character of home signs (Goldin-Meadow 2003), and the performance of individuals under time pressure, where the symptoms of protolanguage have been termed 'perceptual strategies' and 'good-enough parsing' (Dick et al. 2001, Townsend & Bever 2001, Ferreira & Patson 2007). In addition, there is now some evidence that protolinguistic principles are lurking in the background during normal sentence processing: ERP signals reveal differences in processing between a canonical and a noncanonical protolinguistic order (Bornkessel & Schlesewsky 2006, Paczynski & Kuperberg 2011). For instance, Paczynski and Kuperberg show that, following an animate (and therefore typical) agent, it is easier to process an inanimate (more typical) patient than an animate (less typical) patient. Jackendoff 2002 suggests that protolinguistic principles leave their traces in full grammars as well, for instance, in animacy hierarchies, in the grammar of compounding, and in the loose connection between the position of adjuncts and their semantic roles. Progovac (2006, 2007) explores this position further, arguing that protogrammar survives INSIDE or UNDERNEATH full modern grammars.

The formal properties of protolinguistic systems can be modeled by an architecture that consists only of phonology, conceptual structure, and an interface between them, as in Figure 8: there are no syntactic categories, and the only constituent structure is provided by prosody, which does not extend to an indefinite depth.

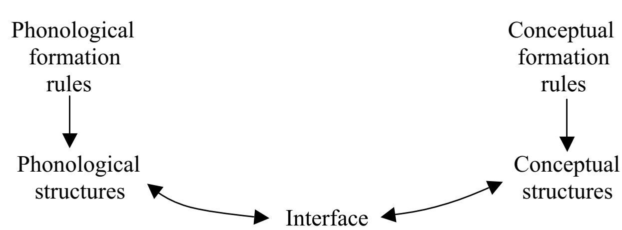

FIGURE 8. Architecture of protolanguage.

Even without syntax, phonology still has the means to convey some aspects of semantic combinatoriality by virtue of linear order—an item's being first, being last, or being adjacent to something else—and by virtue of pauses, intonation, and stress. The evidence from modern grammars and from psycholinguistics suggests that this system is part of the active grammar of full languages. It often reinforces syntactic connections, but sometimes conflicts with them, for instance in passives and object relatives, which are known to be more difficult to process.

I want to stress that even if it were proven that protolanguage is a component of modern language, this would not show that it was a stage in the evolution of the modern language capacity. All that could be argued is that it is PLAUSIBLE (though I can imagine that someday genetic evidence might be telling). Under such a story, conceptual structure evolved first, long before language, in fact long before hominids. Then came phonology and its links to meaning, yielding protolanguage, and last came syntax and its links to both. On this view, the protolinguistic phenomena persisting in modern-day grammar are a sort of 'living fossil' of an earlier evolutionary stage of the language faculty.

In a curious way this story is compatible with Chomsky's speculations. His notion of the inception of recursion is that it applied to 'conceptual atoms', which on the present story fall under conceptual structure. The main difference is in the timing: Chomsky sees the advent of combinatoriality/recursion in thought as occurring at some point in the hominid line, whereas I see it as earlier, in order to account for primate (and other mammalian) cognition. He sees 'externalization' as a second step in the evolution of language. But for him, externalization includes all of phonology and all of morphology, plus most of the aspects of syntax that differentiate one language from the next: word order, agreement, overt case marking, the distinction between WH-movement and WH-in-situ, and so on—in short, most of the things that most linguists think of as 'language'. So perhaps beyond timing, the difference is predominantly one of terminology.

On the present story, there still remains the question of how phonology and its links to semantics evolved, and what was involved in the evolution of syntax, for instance, recursive phrase structure and parts of speech, not to mention long-distance dependencies and the like. Some people (e.g. A. Clark 1996, Deacon 1997, Croft 2001, Arbib 2003, Bybee 2003, Tomasello 2003, 2005, Bybee & McClelland 2005, Heine & Kuteva 2007, Chater et al. 2009, O'Grady 2010) think all of these properties of language can be explained through cultural evolution and/or grammaticalization. I agree that MANY of them can. But it seems to me that these accounts do not ask a number of crucial questions: Where did the cognitive playing field come from, on which cultural innovation can result in language change? How did the human language faculty expand from merely words to the possibility of syntactic categories, of morphology, of long-distance dependencies, and so on? How did the language faculty acquire the predisposition to organize the child's experience along lines made possible by these mechanisms? My sense is that these foundational issues tend to be overlooked because they seem so transparent. But in fact they pose difficult challenges.

In order to explore these issues further, it is of interest to work out the formal steps involved in getting from something like primate calls to protolanguage and from there to full modern language (Jackendoff & Wittenberg 2011). What is called for is a hierarchy (or lattice) of grammars—not the familiar Chomsky hierarchy, which involves uninterpreted formal languages, but rather a hierarchy of formal systems that map between sound and meaning. Each of these systems is more elaborate than the one below it, and each one contains the one below it as a special case. Factors in this hierarchy might include:

- whether utterances contain only one constituent, two concatenated constituents, or indefinitely many constituents
- whether the constituents are single words or whether they can consist of phrases
- whether phrases contain only concatenated words or whether they can consist of concatenated phrases (the latter case is what makes recursion possible)
- whether words and phrases have syntactic categories—creating the possibility of true syntax
- whether words can themselves be decomposed into smaller parts, creating the possibility of morphology
- how much the interface between thoughts and utterances relies on pragmatics and coercion, and how much it is governed by strict compositional principles (fuller languages have more of the latter, but the former never is eliminated).

The work on protolanguage suggests that systems at different levels in the hierarchy may manifest themselves in fully developed grammars, in acquisition, in processing, in aphasia and other language impairments, and also in the communicative abilities of other primates, language-trained or not. If this speculation bears fruit, it will make the language faculty out to be a palimpsest, a layered composite of grammars from many different levels of the hierarchy. But this is an issue well beyond the scope of the present discussion note.

11. Conclusions. I have proposed here that a theory of language that aspires to biological plausibility should take seriously the need for the language faculty to integrate gracefully into the rest of the mind/brain. Even in the absence of detailed theories of mental representation for other cognitive faculties, it is possible to bring this criterion to bear on many fundamental aspects of linguistic theory: the presence of redundancy, the use of recursion, the character of rules and of fundamental computational operations, the nature of the conceptual-intentional system, the interaction between distinct linguistic representations such as syntax and phonology, the interaction between linguistic and nonlinguistic representations, and the possible course of the evolution of the language faculty.

In each case the proposed answer is also motivated on grounds internal to language, and in each case it differs from the proposals of the minimalist program and biolinguistics, which are based on criteria of ‘perfection’, optimal design, and efficient computation. I conclude that a constraint-based and Unification-based parallel architecture leads to more satisfactory accounts of the linguistic phenomena in question, incorporating the insights of many other constraint-based frameworks. At the same time it provides a far more promising approach to the criterion of graceful integration, offering a quite different direction for biolinguistic research.

REFERENCES

Arbib, Michael A. 2003. brain, Meaning, Grammar, evolution. [sic] *Behavioral and Brain Sciences* 26.668–69.

Arnold, Kate, and Klaus Zuberbühler. 2008. Meaningful call combinations in a nonhuman primate. *Current Biology* 18.R202–R203.

Aronoff, Mark; Irit Meir; and Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81.301–44.

Asaka, Toshihiko. 2002. A lexical licensing analysis of the adjectival noun construction. *English Linguistics* 19.113–41.

Asher, Nicholas, and Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.

Baayen, Harald; Robert Schreuder; Nivja de Jong; and Andrea Krott. 2002. Dutch inflection: The rules that prove the exception. *Storage and computation in the language faculty*, ed. by Sieb Nooteboom, Fred Weerman, and Frank Wijnen, 61–92. Dordrecht: Kluwer.

Badler, Norman I.; Rama Bindiganavale; Jan Allbeck; William Schuler; Liwei Zhao; and Martha Palmer. 2000. A parameterized action representation for virtual human agents. *Embodied conversational agents*, ed. by Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth F. Churchill, 256–84. Cambridge, MA: MIT Press.

Berwick, Robert. 2011a. All you need is Merge: Biology, computation, and language from the bottom up. In Di Sciullo & Boeckx, 461–91.

Berwick, Robert. 2011b. Syntax facit saltum redux: Biolinguistics and the leap to syntax. In Di Sciullo & Boeckx, 65–99.

Berwick, Robert, and Noam Chomsky. 2011. The biolinguistic program: The current state of its development. In Di Sciullo & Boeckx, 19–41.

Bickerton, Derek. 1990. *Language and species*. Chicago: University of Chicago Press.

BICKERTON, DEREK. 2010. On two incompatible theories of language evolution. In Larson et al., 199–210.

BINDIGANAVALE, RAMA; WILLIAM SCHULER; JAN M. ALLBECK; NORMAN I. BADLER; ARAVIND K. JOSHI; and MARTHA PALMER. 2000. Dynamically altering agent behaviors using natural language instructions. *AGENTS '00: Proceedings of the fourth international conference on Autonomous Agents*, 293–300.

BOECKX, CEDRIC. 2011. Approaching parameters from below. In Di Sciullo & Boeckx, 205–21.

BOOIJ, GEERT. 2002. Constructional idioms, morphology and the Dutch lexicon. *Journal of Germanic Linguistics* 14.301–29.

BORNKESSEL, INA, and MATTHIAS SCHLESEWSKY. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113.787–821.

BRODY, MICHAEL. 1995. *Lexico-logical form*. Cambridge, MA: MIT Press.

BYBEE, JOAN. 2003. Cognitive processes in grammaticalization. *The new psychology of language: Cognitive and functional approaches to language structure, vol. 2*, ed. by Michael Tomasello, 145–67. Mahwah, NJ: Lawrence Erlbaum.

BYBEE, JOAN, and JAMES MCCLELLAND. 2005. Alternatives to the combinatorial paradigm. *The Linguistic Review* 22.381–410.

BYRNE, RICHARD, and ANDREW WHITEN (eds.) 1988. *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Oxford University Press.

CHATER, NICK; FLORENCIA REALI; and MORTEN H. CHRISTIANSEN. 2009. Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences* 106.1015–20.

CHENEY, DOROTHY, and ROBERT SEYFARTH. 1990. *How monkeys see the world*. Chicago: University of Chicago Press.

CHENEY, DOROTHY, and ROBERT SEYFARTH. 2005. Constraints and preadaptations in the earliest stages of language. *The Linguistic Review* 22.135–60.

CHENEY, DOROTHY, and ROBERT SEYFARTH. 2007. *Baboon metaphysics*. Chicago: University of Chicago Press.

CHOMSKY, NOAM. 1959. Review of *Verbal behavior*, by B. F. Skinner. *Language* 35.26–58.

CHOMSKY, NOAM. 1963. Formal properties of grammars. *Handbook of mathematical psychology, vol. 2*, ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter, 323–418. New York: Wiley.

CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

CHOMSKY, NOAM. 1968. *Language and mind*. New York: Harcourt, Brace & World.

CHOMSKY, NOAM. 1972. Some empirical issues in the theory of transformational grammar. *Studies on semantics in generative grammar*, 120–202. The Hague: Mouton.

CHOMSKY, NOAM. 1975. *Reflections on language*. New York: Pantheon.

CHOMSKY, NOAM. 1981. *Lectures on government and binding*. Dordrecht: Foris.

CHOMSKY, NOAM. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

CHOMSKY, NOAM. 2000. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

CHOMSKY, NOAM. 2005. Three factors in language design. *Linguistic Inquiry* 36.1–22.

CHOMSKY, NOAM. 2010. Some simple evo devo theses: How true might they be for language? In Larson et al., 45–62.

CHOMSKY, NOAM, and GEORGE A. MILLER. 1963. Introduction to the formal analysis of natural languages. *Handbook of mathematical psychology, vol. 2*, ed. by R. Duncan Luce, Robert R. Bush, and Eugene Galanter, 269–321. New York: Wiley.

CLARK, ANDY. 1996. Linguistic anchors in a sea of thought? *Pragmatics and Cognition* 4.93–103.

CLARK, HERBERT H. 1996. *Using language*. Cambridge: Cambridge University Press.

COEN, MICHAEL HARLAN. 2006. *Multimodal dynamics: Self-supervised learning in perceptual and motor systems*. Cambridge, MA: MIT dissertation.

COHN, NEIL. 2010. The limits of time and transitions: Challenges to theories of sequential image comprehension. *Studies in Comics* 1.127–47.

COHN, NEIL. 2011. You're a good experiment, Charlie Brown: Narrative structure in comic strips. Medford, MA: Tufts University, work in preparation.

Cohn, Neil; Martin Paczynski; Phillip Holcomb; Ray Jackendoff; and Gina Kuperberg. 2010. Impact of structure and meaning on sequential image comprehension. Paper presented at the 23rd annual CUNY Conference on Human Sentence Processing, New York.

Crick, Francis. 1994. *The astonishing hypothesis: The scientific search for the soul*. New York: Charles Scribner's Sons.

Croft, William. 2001. *Radical construction grammar*. Oxford: Oxford University Press.

Culicover, Peter W. 1972. OM-sentences: On the derivation of sentences with systematically unspecifiable interpretations. *Foundations of Language* 8.199–236.

Culicover, Peter W., and Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.

Cutler, Anne, and Charles Clifton, Jr. 1999. Comprehending spoken language: A blueprint of the listener. *The neurocognition of language*, ed. by Colin Brown and Peter Hagoort, 123–66. Oxford: Oxford University Press.

Dawkins, Richard. 1976. *The selfish gene*. Oxford: Oxford University Press.

Deacon, Terrence. 1997. *The symbolic species*. New York: Norton.

Di Sciullo, Anna Maria, and Cedric Boeckx (eds.) 2011. *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*. Oxford: Oxford University Press.

Dick, Frederic; Elizabeth Bates; Beverly Wulfeck; Jennifer Aydelott Utman; Nina Dronkers; and Morton Ann Gernsbacher. 2001. Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* 108.759–88.

Evans, Nicholas, and Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–48.

Fabb, Nigel, and Morris Halle. 2008. *Meter in poetry: A new theory*. Cambridge: Cambridge University Press.

Ferreira, Fernanda, and Nikole D. Patson. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1.71–83.

Fillmore, Charles; Paul Kay; and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64.501–38.

Fitch, W. Tecumseh. 2010. Three meanings of 'recursion': Key distinctions for biolinguistics. In Larson et al., 73–90.

Fitch, W. Tecumseh. 2011. 'Deep homology' in the biology and evolution of language. In Di Sciullo & Boeckx, 135–66.

Flack, Jessica; David Krakauer; Simon DeDeo; and Ray Jackendoff. 2011. A 'grammar' for primate social behavior. Santa Fe: Santa Fe Institute, and Medford, MA: Tufts University, work in preparation.

Fodor, Jerry A. 1975. *The language of thought*. Cambridge, MA: Harvard University Press.

Ford, Marilyn; Joan Bresnan; and Ronald M. Kaplan. 1982. A competence-based theory of syntactic closure. *The mental representation of grammatical relations*, ed. by Joan Bresnan, 727–96. Cambridge, MA: MIT Press.

Gallistel, C. R., and Adam Philip King. 2009. *Memory and the computational brain: Why cognitive science will transform neuroscience*. Malden, MA: Wiley-Blackwell.

Gazzaniga, Michael. 2010. Neuroscience and the correct level of explanation for understanding mind. *Trends in Cognitive Sciences* 14.291–92.

Gee, James Paul, and François Grosjean. 1984. Empirical evidence for narrative structure. *Cognitive Science* 8.59–85.

Gentner, Timothy Q.; Kimberly M. Fenn; Daniel Margoliash; and Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440.1204–7.

Gergely, György, and Gergely Csibra. 2003. Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences* 7.287–92.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68.1–76.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7.219–24.

Goldberg, Adele, and Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80.532–68.

Goldin-Meadow, Susan. 2003. *The resilience of language*. New York: Psychology Press.

Gordon, Peter. 2003. The origin of argument structure in infant event representations. *Proceedings of the Boston University Conference on Language Development* 26.189–98.

Gruber, Jeffrey. 1965. *Studies in lexical relations*. Cambridge, MA: MIT dissertation. [Published as part of *Lexical structures in syntax and semantics*, Amsterdam: North-Holland, 1976.]

Haegeman, Liliane. 1992. *Introduction to the theory of government and binding*. Oxford: Blackwell.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–76. Cambridge, MA: MIT Press.

Hauser, Marc D. 1996. *The evolution of communication*. Cambridge, MA: MIT Press.

Hauser, Marc D.; Noam Chomsky; and W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298.1569–79.

Heine, Bernd, and Tania Kuteva. 2007. *The genesis of grammar: A reconstruction.* Oxford: Oxford University Press.

Heinrich, Bernd. 1989. *Ravens in winter*. New York: Summit Books.

Hinzen, Wolfram. 2008. Forum: Prospects for an explanatory theory of semantics. *Biolinguistics* 2.348–63. Online: http://biolinguistics.eu/index.php/biolinguistics/article/view/77/88.

Hinzen, Wolfram. 2011. Emergence of a systemic semantics through minimal and underspecified codes. In Di Sciullo & Boeckx, 417–39.

Hochberg, Julian E. 1978. *Perception*. 2nd edn. Englewood Cliffs, NJ: Prentice-Hall.

Hockett, Charles. 1960. The origin of speech. *Scientific American* 203.88–111.

Hofstadter, Douglas. 1979. *Gödel, Escher, Bach: An eternal golden braid*. New York: Basic Books.

Hornstein, Norbert, and Cedric Boeckx. 2009. Approaching universals from the bottom up. *Language universals*, ed. by Morten Christiansen, Chris Collins, and Shimon Edelman, 79–98. Oxford: Oxford University Press.

Hultsch, Henrike; Roger Mundry; and Dietmar Todt. 1999. Learning, representation and retrieval of rule-related knowledge in the song system of birds. *Learning: Rule extraction and representation*, ed. by Angela D. Friederici and Randolf Menzel, 89–115. Berlin: Walter de Gruyter.

Humphries, Glyn W.; Emer M. E. Forde; and M. Jane Riddoch. 2001. The planning and execution of everyday actions. *The handbook of cognitive neuropsychology: What deficits reveal about the human mind*, ed. by Brenda Rapp, 565–89. Philadelphia: Psychology Press.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51.639–71. [Reprinted in Jackendoff 2010.]

Jackendoff, Ray. 1976. Toward an explanatory semantic representation. *Linguistic Inquiry* 7.89–150.

Jackendoff, Ray. 1983. *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1987a. *Consciousness and the computational mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1987b. On beyond zebra: The relation of linguistic and visual information. *Cognition* 26.89–114. [Reprinted in Jackendoff 2010.]

Jackendoff, Ray. 1989. What is a concept, that a person can grasp it? *Mind and Language* 4.68–102.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1991a. The problem of reality. *Noûs* 25.411–33.

Jackendoff, Ray. 1991b. Parts and boundaries. *Cognition* 41.9–45. [Reprinted in Jackendoff 2010.]

Jackendoff, Ray. 1996a. The architecture of the linguistic-spatial interface. *Language and space*, ed. by Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, 1–30. Cambridge, MA: MIT Press. [Reprinted in Jackendoff 2010.]

JACKENDOFF, RAY. 1996b. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory* 14.305–54. [Reprinted in Jackendoff 2010.]

JACKENDOFF, RAY. 1996c. How language helps us think. *Pragmatics and Cognition* 4.1–34.

JACKENDOFF, RAY. 1997a. *The architecture of the language faculty*. Cambridge, MA: MIT Press.

JACKENDOFF, RAY. 1997b. Twistin' the night away. *Language* 73.534–59.

JACKENDOFF, RAY. 1999. The natural logic of rights and obligations. *Language, logic, and concepts: Essays in memory of John Macnamara*, ed. by Ray Jackendoff, Paul Bloom, and Karen Wynn, 67–95. Cambridge, MA: MIT Press.

JACKENDOFF, RAY. 2002. *Foundations of language*. Oxford: Oxford University Press.

JACKENDOFF, RAY. 2007a. *Language, consciousness, culture: Essays on mental structure*. Cambridge, MA: MIT Press.

JACKENDOFF, RAY. 2007b. A parallel architecture perspective on language processing. *Brain Research* 1146.2–22.

JACKENDOFF, RAY. 2008. *Construction after construction* and its theoretical challenges. *Language* 84.8–28.

JACKENDOFF, RAY. 2009. Parallels and non-parallels between language and music. *Music Perception* 26.195–204.

JACKENDOFF, RAY. 2010. *Meaning and the lexicon: The parallel architecture 1975–2010*. Oxford: Oxford University Press.

JACKENDOFF, RAY. 2012. *A user's guide to thought and meaning*. Oxford: Oxford University Press, to appear.

JACKENDOFF, RAY, and EVA WITTENBERG. 2011. A new hierarchy of grammars. Medford, MA: Tufts University, work in preparation.

JENKINS, LYLE. 2000. *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.

JOHNSON, DAVID E., and SHALOM LAPPIN. 1999. *Local constraints vs. economy*. Stanford, CA: CSLI Publications.

KATZ, JONAH, and DAVID PESETSKY. 2009. The identity thesis for language and music. Cambridge, MA: MIT, MS. Online: http://ling.auf.net/lingBuzz/000959.

KAYNE, RICHARD. 1983. Connectedness. *Linguistic Inquiry* 24.223–49.

KINSELLA, ANNA. 2009. *Language evolution and syntactic theory*. Cambridge: Cambridge University Press.

KIPPER, KARIN, and MARTHA PALMER. 2000. Representation of actions as an interlingua. *Proceedings of the 2000 NAACL-ANLP Workshop on Applied Interlinguas: Practical applications of interlingual approaches to NLP*, 12–17. Stroudsburg, PA: Association for Computational Linguistics.

KLEIN, WOLFGANG, and CLIVE PERDUE. 1997. The basic variety, or: Couldn't language be much simpler? *Second Language Research* 13.301–47.

KOCH, CHRISTOF. 2004. *The quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts.

KÖHLER, WOLFGANG. 1927. *The mentality of apes*. London: Routledge & Kegan Paul.

KOSTER, JAN. 1978. *Locality principles in syntax.* Dordrecht: Foris.

LAKOFF, GEORGE. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

LANDAU, BARBARA, and RAY JACKENDOFF. 1993. 'What' and 'where' in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16.217–38.

LANGACKER, RONALD. 1987. *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.

LANGACKER, RONALD. 1998. Conceptualization, symbolization, and grammar. *The new psychology of language: Cognitive and functional approaches to language structure, vol. 1*, ed. by Michael Tomasello, 1–39. Hillsdale, NJ: Lawrence Erlbaum.

LARSON, RICHARD. 2011. Clauses, propositions, and phrases. In Di Sciullo & Boeckx, 366–91.

LARSON, RICHARD; VIVIANE DÉPREZ; and HIROKO YAMAKIDO (eds.) 2010. *The evolution of human language: Biolinguistic perspectives*. Cambridge: Cambridge University Press.

LASNIK, HOWARD. 2011. What kind of computing device is the human language faculty? In Di Sciullo & Boeckx, 354–65.

LERDAHL, FRED, and RAY JACKENDOFF. 1983. *A generative theory of tonal music*. Cambridge, MA: MIT Press.
LERDAHL, FRED, and RAY JACKENDOFF. 2006. The capacity for music: What's special about it? *Cognition* 100.33–72.
LEVIN, BETH, and MALKA RAPPAPORT HOVAV. 1990. Wiping the slate clean: A lexical semantic exploration. *Cognition* 41.123–55.
LIBERMAN, ALVIN. 1993. Some assumptions about speech and how they changed. *Haskins Laboratories Status Report on Speech Research* SR-113.1–32. Online: http://www.haskins.yale.edu/sr/sr113/SR113_01.pdf.
LORENZ, KONRAD. 1952. *King Solomon's ring*. New York: Thomas Y. Crowell.
MACDONALD, MARYELLEN C.; NEAL J. PEARLMUTTER; and MARK S. SEIDENBERG. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 101.676–703.
MANDLER, JEAN. 1987. On the psychological reality of story structure. *Discourse Processes* 10.1–29.
MARANTZ, ALEC. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics (Proceedings of the 21st annual Penn Linguistics Colloquium)* 4.2.201–25.
MARANTZ, ALEC. 2005. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* 22.429–46.
MARLER, PETER, and HANS SLABBEKOORN (eds.) 2004. *Nature's music: The science of birdsong*. San Diego: Elsevier.
MARR, DAVID. 1982. *Vision*. San Francisco: Freeman.
MARSLEN-WILSON, WILLIAM, and LORRAINE TYLER. 1987. Against modularity. *Modularity in knowledge representation and natural-language understanding*, ed. by Jay Garfield, 37–62. Cambridge, MA: MIT Press.
MCCAWLEY, JAMES D. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4.243–69.
MILLER, GEORGE A.; EUGENE GALANTER; and KARL H. PRIBRAM. 1960. *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston.
NEWMEYER, FREDERICK J. 2003. Review article on *On nature and language*, by Noam Chomsky; *The language organ*, by Stephen R. Anderson and David W. Lightfoot; and *Language in a Darwinian perspective*, ed. by Bernard H. Bichakjian. *Language* 79.583–99.
NEWMEYER, FREDERICK J. 2008. Review of *Linguistic minimalism*, by Cedric Boeckx. *Language* 84.387–95.
NOWAK, MARTIN; DAVID C. KRAKAUER; and ANDREAS DRESS. 1999. An error limit for the evolution of language. *Proceedings of the Royal Society of London B: Biological Sciences* 266.2131–36.
O'GRADY, WILLIAM. 2010. An emergentist approach to syntax. *The Oxford handbook of linguistic analysis*, ed. by Bernd Heine and Heiko Narrog, 257–83. Oxford: Oxford University Press.
PACZYNSKI, MARTIN, and GINA KUPERBERG. 2011. Electrophysiological evidence for use of the animacy hierarchy, but not thematic role assignment, during verb argument processing. *Language and Cognitive Processes*, to appear.
PATEL, ANIRUDDH. 2008. *Music, language, and the brain*. Oxford: Oxford University Press.
PHILLIPS, COLIN, and ELLEN LAU. 2004. Foundational issues (Review article on Jackendoff 2002). *Journal of Linguistics* 40.571–91.
PHILLIPS, COLIN, and SHEVAUN LEWIS. 2012. Derivational order in syntax: Evidence and architectural consequences. *Directions in derivations*, ed. by Cristiano Chesi. Amsterdam: Elsevier, to appear.
PINKER, STEVEN. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
PINKER, STEVEN. 2007. *The stuff of thought*. New York: Penguin.
PINKER, STEVEN, and RAY JACKENDOFF. 2005. The faculty of language: What's special about it? *Cognition* 95.201–36.
PODOS, JEFFREY; SARAH HUBER; and BENJAMIN TAFT. 2004. Bird song: The interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics* 35.55–87.

Pollard, Carl, and Ivan Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Postal, Paul. 1972. The best theory. *The goals of linguistic theory*, ed. by P. Stanley Peters, 131–70. Englewood Cliffs, NJ: Prentice-Hall.

Postal, Paul. 2004. *Skeptical linguistic essays*. Oxford: Oxford University Press.

Povinelli, Daniel J. 2000. *Folk physics for apes*. Oxford: Oxford University Press.

Price, J. Jordan, and Scott Lanyon. 2002. Reconstructing the evolution of complex bird song in the oropendolas. *Evolution* 56.7.1514–29.

Progovac, Ljiljana. 2006. The syntax of nonsententials: Small clauses and phrases at the root. *The syntax of nonsententials*, ed. by Ljiljana Progovac, Kate Paesani, Eugenia Casielles, and Ellen Barton, 33–71. Amsterdam: John Benjamins.

Progovac, Ljiljana. 2007. Grammars within without recursion: Implications for evolutionary studies. Paper presented at the Illinois State University Conference on Recursion in Human Languages, Normal, IL.

Pullum, Geoffrey K. 2007. The evolution of model-theoretic frameworks in linguistics. *Proceedings of Model-Theoretic Syntax at 10* (*part of ESSLLI 2007)*, ed. by James Rogers and Stephan Kepser, 1–10. Dublin: Trinity College Dublin.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Sadock, Jerrold. 1991. *Autolexical syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.

Sag, Ivan. 1992. Taking performance seriously. *Lenguajes naturales y lenguajes formales: Actas del VII congreso de lenguajes naturales y lenguajes formales (Vic-Barcelona, 23–27 de septiembre de 1991)*, ed. by Carlos Martín Vide, 61–74. Barcelona: Promociones y Publicaciones Universitarias.

Sag, Ivan, and Thomas Wasow. 2011. Performance-compatible competence grammar. *Non-transformational syntax*, ed. by Robert Borsley and Kersti Borjars, 359–77. London: Wiley-Blackwell.

Schenker, Heinrich. 1935. *Der freie Satz*. Vienna: Universal Edition. [Trans. and ed. by Ernst Oster, New York: Longman, 1979.]

Seuren, Peter. 2004. *Chomsky's minimalism*. New York: Oxford University Press.

Shieber, Stuart. 1986. *An introduction to unification-based approaches to grammar*. Stanford, CA: CSLI Publications.

Song, Hyun-joo; Renée Baillargeon; and Cynthia Fisher. 2005. Can infants attribute to an agent a disposition to perform a particular action? *Cognition* 98.B45–B55.

Stromswold, Karen. 2010. Genetics and the evolution of language: What genetic studies reveal about the evolution of language. In Larson et al., 176–90.

Talmy, Leonard. 2000. *Toward a cognitive semantics*. Cambridge, MA: MIT Press.

Tanenhaus, Michael K.; Michael J. Spivey-Knowlton; Kathleen M. Eberhard; and Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268.1632–34.

Tarski, Alfred. 1956. The concept of truth in formalized languages. *Logic, semantics, metamathematics: Papers from 1923 to 1938*, 2nd edn., trans. by Joseph H. Woodger, ed. by John Corcoran, 152–97. London: Oxford University Press.

Tinbergen, Niko. 1951. *The study of instinct*. Oxford: Clarendon.

Tomasello, Michael (ed.) 2000. *Primate cognition*. Special issue of *Cognitive Science* 24.3.

Tomasello, Michael. 2003. *Constructing a language*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2005. Beyond formalities: The case of language acquisition. *The Linguistic Review* 22.183–97.

Tomasello, Michael; Josep Call; and Brian Hare. 2003. Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in Cognitive Sciences* 7.153–56.

Townsend, David J., and Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Trueswell, John C.; Michael K. Tanenhaus; and Susan M. Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33.285–318.

TSOULAS, GEORGE. 2010. Computations and interfaces: Some notes on the relation between the language and the music faculties. *Musicae Scientiae, Discussion Forum* 5.11–41.
VAN VALIN, ROBERT, JR., and RANDY J. LAPOLLA. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.
WEINREICH, URIEL. 1966. Explorations in semantic theory. *Current trends in linguistics, vol. 3: Theoretical foundations*, ed. by Thomas A. Sebeok, 99–201. The Hague: Mouton. [Reprinted in *On semantics*, by Uriel Weinreich, ed. by William Labov and Beatrice S. Weinreich. Philadelphia: University of Pennsylvania Press, 1980.]
WHITEN, ANDREW. 2002. The imitator's representation of the imitated: Ape and child. *The imitative mind: Development, evolution, and brain bases*, ed. by Andrew N. Meltzoff and Wolfgang Prinz, 98–121. Cambridge: Cambridge University Press.
WILLIAMS, EDWIN. 1994. Remarks on lexical knowledge. *The acquisition of the lexicon*, ed. by Lila Gleitman and Barbara Landau, 7–34. Cambridge, MA: MIT Press.
WRANGHAM, RICHARD, and DALE PETERSON. 1996. *Demonic males: Apes and the origins of human violence*. New York: Houghton Mifflin.
ZUBERBÜHLER, KLAUS. 2002. A syntactic rule in forest monkey communication. *Animal Behaviour* 63.293–99.

Center for Cognitive Studies
11 Miner Hall
Tufts University
Medford, MA 02155
[Ray.Jackendoff@tufts.edu]

[Received 24 August 2010;
revision invited 29 October 2010;
revision received 19 February 2011;
accepted 9 June 2011]