CHAPTER 23

THE PARALLEL ARCHITECTURE AND ITS PLACE IN COGNITIVE SCIENCE

RAY JACKENDOFF

It has become fashionable recently to speak of linguistic inquiry as *biolinguistics*, an attempt to frame questions of linguistic theory in terms of the place of language in a biological context. The Minimalist Program (Chomsky 1995; 2001) is of course the most prominent stream of research in this paradigm. However, an alternative stream within the paradigm, the Parallel Architecture, has been developing in my own work over the past 30 years; it includes two important subcomponents, Conceptual Structure and Simpler Syntax (Jackendoff 2002; 2007*b*; Culicover and Jackendoff 2005). This chapter will show how the Parallel Architecture is in many ways a more promising realization of biolinguistic goals than the Minimalist Program and that, more than the Minimalist Program, it is conducive to integration with both the rest of linguistic theory and the rest of cognitive science.

978–0–19–954400–4 Heine-main-drv Heine-Narrog (Typeset by Spi, Chennai) 646 of 778 April 8, 2009 21:55

23.1 Parallel architectures, broadly conceived

The Parallel Architecture (PA) can be explored at two levels: First, what is a parallel architecture in general? Second, what distinguishes "the" Parallel Architecture from other theories within this genre? In both cases, the basic question is:

(1) What is the best way to allocate the generative capacity of language, so as to account for the observed relations between sound and meaning?

Traditional generative grammar, from *Syntactic Structures* (Chomsky 1957) through the Minimalist Program, has answered:

(2) (Syntactocentric architecture) The recursive rules of the syntactic component provide the generative capacity of language. The relation between sound and meaning arises through mapping syntactic structures into phonetic form (PF) (or the "sensorimotor interface") on one hand and logical form (LF) (or the "conceptual–intentional interface") on the other.

However, theoretical developments as early as the 1970s showed that phonological structures have their own primitives and principles of combination that cannot be reduced to syntactic terms. For instance, rules of syllabification, prosody, and morphophonology are stated in terms of units that are thoroughly comfortable in phonological terms but often quite uncomfortable in syntactic terms. The same is true of meanings: semantic notions like event, manner, quantification, and focus cut across syntactic categories but are characterizable in independent semantic terms. In particular, it has been a staple of linguistic theory and psycholinguistics to distinguish semantic ill-formedness (**Colorless green ideas sleep furiously*) from syntactic ill-formedness (**A good ideas am rare*), which depends on the distinction between semantic and syntactic combinatoriality. (Note: "syntactic" is sometimes used to mean "combinatorial in any sense", including music, phonology, and semantics. I am using the term here in the narrow sense of "combinatoriality whose units are things like Noun and Verb.")

Within the syntactocentric approach, mismatches between phonology and syntax have been either incorporated into syntax (e.g., "Spell-Out") or ignored, at least by syntacticians. More striking has been the constant attempt to build more and more aspects of semantics into syntactic structure—as is required by an architecture in which all combinatoriality is a consequence of syntax. The inevitable result is a syntactic component overstuffed with invisible structure, in which every constituent moves multiple times. Differences of opinion arise as to whether this is a good thing or not.

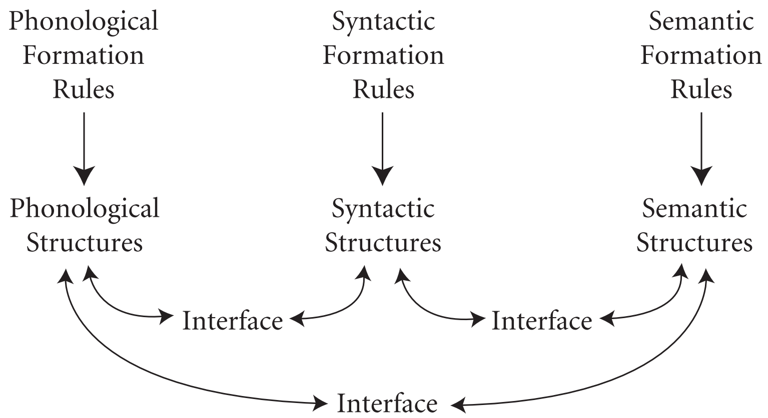
A parallel architecture answers question (1) like this:

(3) (Parallel architecture)

a. The generative capacity of language is invested in multiple components—at the very least, autonomous generative components for phonological, syntactic, and semantic structure. Each component has its own distinctive primitives and principles of combination, and generates its own structure.

b. The relation between sound and meaning is mediated by a set of interface components, which establish optimal linkings among the various structures and their parts. (Note: in this theory, an interface is not a level of structure but rather a *connection* between *two* levels of structure.)

c. The structure of a sentence is therefore an n-tuple of structures, one for each appropriate component, plus the linkages established among them by the interface components.



A priori, answer (2) seems simpler, since it has has only one "generative engine" and fewer components overall. But, to parallel Chomsky's (1972?) rejoinder to Postal's (1972) *The Best Theory*, architectures must be judged not only on their formal elegance but also on their affordance for describing the data of language in full detail (descriptive adequacy), in explaining language variation and the possibility of language acquisition (explanatory adequacy), and in explaining how the system can arise from more general cognitive and biological principles ("beyond explanatory adequacy", to use the term of Chomsky 2001). In particular, formal elegance must not be conflated with biological or cognitive elegance, which might be quite different.

Pursuing the goal of going beyond explanatory adequacy, consider which sort of architecture conforms more closely to what is known about the brain. The visual system is known to contain numerous areas specialized to different aspects of visual perception: detection of motion, detection of color, several independent mechanisms for constructing the perception of depth, possibly face perception, and many

many others. Each of them accounts for a relatively limited aspect of visual understanding; the totality of visual understanding arises from their combined efforts. In order for their efforts to combine, they must communicate, linking their respective representations in an optimal fashion (Koch 2004). At the moment, we don't know a lot about the formal details of representations computed by various visual areas, and there is still much dispute about what brain areas are responsible for different aspects of linguistic understanding. Nevertheless, the overall flavor of the visual system is far more compatible with a parallel architecture, with its multiple independent but linked components, than with a syntactocentric one. No one to my knowledge has ever proposed a visual counterpart of a syntactocentric grammar.

There is one cognitive capacity other than language for which formal details of the representations have been explored in some detail: music. Here it proves impossible to generate musical structures from a single component. Lerdahl and Jackendoff 1983 (see also Jackendoff and Lerdahl 2006) develop a parallel architecture for music containing four components linked by interface rules. One of these structures, grouping, is a general-purpose cognitive capacity that also plays an important role in vision. Another, metrical structure, bears strong similarities to the metrical systems that determine stress and prosody in language. The other two structures are, so far as we know at the moment, particular to music.

One of my original motivations for a parallel architecture in language (Jackendoff 1997, 2002) was the existence of multiple independent tiers in phonology, such as syllabic structure, metrical structure, prosody, and tone, also linked by correspondence or interface rules. Similarly, it is now fairly clear that semantic structure can be dissected into semi-independent structures—at least propositional structure (who did what to whom) and information structure (topic vs. focus vs. common ground). Finally, the relation of language to vision, such that we can talk about what we see, has to be mediated by a set of principles that link linguistic representations of some level to visual representations of some level—it cannot be accounted for through further derivation from syntactic structure (Jackendoff 1987). Thus the internal structure of some components of language, as well as the relation of language to other faculties, is consonant with a parallel architecture for language as a whole.

A parallel architecture for language and other cognitive faculties displays a version of modularity. This is not modularity in Fodor's (1983) sense, which seals off various capacities from each other, but what could be called *representational* or*structure-based modularity*. Each separate form of representation has its own particular autonomous (i.e., domain-specific) structure, and its own interfaces to other structures. One form of representation is *relatively* informationally encapsulated from another to the degree that one can influence the other only through a series of interfaces, or through a narrowly specialized interface. For example, phonological structure is relatively encapsulated from visual representations because, in order to speak about what one sees, one has to pass from high-level visual understanding through linguistic semantic structure and syntactic structure in order to influence phonology—i.e., through a series of interfaces. However, there is also a narrowly circumscribed vision-to-phonology interface that subserves reading, and this aspect of vision is rather tightly yoked to phonology. (For more detail, see Jackendoff 1987, chapter 12; 2002, section 7.5.)

In short, the spirit of parallel architectures is in overall accord with what is known about (a) the brain, (b) the structure of other cognitive capacities, (c) the interior structure of linguistic components, and (d) the interaction of language with other cognitive capacities. The syntactocentric architecture, including the Minimalist Program as one realization, is not. (An advocate of Minimalism might respond that this issue is one of performance or implementation, and so this sort of evidence is not pertinent to Minimalist inquiry. I would consider such a response simply a rhetorical avoidance of the evidence.)

Many different theories of grammar employ parallel architectures in this broad sense. As noted above, phonological theory since the mid-1970s has been thoroughly parallel in conception. Among syntactic theories, the most prominent parallel architecture is Lexical Functional Grammar (Bresnan 2001), where the work of syntax is divided between f-structure, c-structure, and the interface between them. Autolexical Syntax (Sadock 1991) has parallel components for morphosyntactic structure and phrasal syntactic structure, with the possibility of further subdivision. Role and Reference Grammar (Van Valin and LaPolla 1997) subdivides syntax into morphosyntax and phrasal syntax, and semantics into propositional and information structures, with interfaces running in all directions. Construction Grammar (Fillmore 1988, Goldberg 1995) is not formally laid out as a parallel architecture, but it acknowledges the independence of semantics from syntactic form, in that it emphasizes the many–many mapping between syntactic form and meaning, possible only if semantics is autonomous. And the granddaddy of them all is Stratificational Grammar (Lamb 1966), which decomposes the entire grammar into a long sequence of autonomous levels linked by interface components.

Another fundamental question in the architecture of grammar is this:

(4) What formal operations are employed in building linguistic structure?

The mainstream architecture (along with Tree-Adjoining Grammar, Joshi 1987) gives the following answer:

(5) (Derivation-based generation) Syntactic trees are built algorithmically, either from the top down (as in pre-Minimalist theories) or from the bottom up (as in MP and TAG), and they undergo a sequence of distortions (movements and deletions) to derive sound and meaning.

In parallel architectures, the interface relation between different components cannot be a sequenced *derivation*, since structures in different components often stand in a many-to-many relation. Rather, the interface components must be treated as *constraints* (possibly violable), which establish (or license) well-formed links among different kinds of structure. In principle, the rules responsible for each individual component of structure could be algorithmic. But in practice, almost all parallel architectures I have encountered have utilized a constraint-based formalism, in which each independent structure is licensed by simultaneously applied constraints. (An exception is Synchronous TAG, Shieber and Schabes 1991.) To sum up, the answer to question (4) is (6).

(6) (Constraint-based generation) The structures of each component are licensed by simultaneously applied component-internal constraints. The relationships among structures in different components are licensed by interface constraints.

Thus a parallel derivation has no notion of logical sequence, as is essential in a syntactocentric derivation. This has consequences for the relation of linguistic theory to theories of processing, as we will see in the next section.

23.2 The Parallel Architecture: The lexicon

Having settled on a parallel architecture, the more specific question is: What are the autonomous representational formats, and what are the interfaces among them? What I have been calling "the" Parallel Architecture (in capitals, or PA) incorporates specific proposals about semantics, phrasal syntax, and the interface between them, plus less specific proposals about morphology and phonology.

A leading question in the Parallel Architecture is the structure of the lexicon. The question is stated in essentially psycholinguistic terms:

(7) What linguistic material does a speaker have to store in memory—i.e., What is in the lexicon? What structures can be built online in the course of speaking and understanding?

Traditionally, the lexicon is thought of as consisting of words (or morphemes), a distinct component of the language from the rules of grammar. Thinking in terms of question (7) leads to quite a different conception, as we will now see.

A typical word—in any theory—is a triple of phonological, syntactic, and semantic information. In syntactocentric theories, a word is inserted into a syntactic derivation (by lexical insertion or Merge), and it is carried through the derivation to the points where its phonological and semantic properties are "read off". In the Parallel Architecture, the picture is quite different. The structure of a word suits it perfectly to function as a part of the interface components: it establishes a correspondence between small chunks of phonological, syntactic, and semantic structures. (Larger chunks are connected by other interface rules.)

There is no "point in the derivation" where a word is inserted. Rather, one can think of the word being "inserted" into all three structures at the same time, along with the links among them. Or one can think of the word as licensing the connection among preexisting structures. Alternatively, one can think in terms of processing. Given a perceived phonological structure, the word licenses the building of a connection to the corresponding pieces of syntactic and semantic structure; given a piece of meaning to be expressed, the word licenses connecting it to appropriate pieces of syntactic and phonological structures. This last view suits PA to serve directly as a component of a theory of sentence processing (Jackendoff 2002, chapter 7; 2007*a*). PA itself is nondirectional, but its constraints can be implemented in an order suited to particular processing tasks.

Among the information coded in a lexical item is its contextual restrictions. Syntactic contextual restrictions include subcategorization features on syntactic arguments; semantic contextual restrictions include selectional restrictions on semantic arguments. Often these two are partly redundant with each other, but not always (see Jackendoff 2002, section 5.9).

Not every word has to connect all three components. English contains a small collection of "defective" words such as (8a). These have phonology and meaning but no syntactic properties that allow them to combine into larger phrases (aside from within direct quotes, where anything at all is allowed). There are also a few words that have phonological and syntactic properties but no meaning, such as (8b).

(8) a. Phonology and meaning, no syntax
hello, ouch, upsy-daisy, allakazam, wow, shhh, ...

b. Phonology and syntax, no meaning
do (*do*-support), it (pleonastic), of (N *of* NP)

A lexicon conceived in terms of question (7) must contain more than single words. Most obviously, it must contain the thousands of idioms and other fixed expressions in the language such as (9), all of which are units known by native speakers.

(9) a. Idioms
kick the bucket, a breath of fresh air, right on the money, the jig is up, day in day out, clean as a whistle, pie in the sky, ...

b. Fixed expressions (clichés, etc.)
baby-blue eyes, home sweet home, take it from me, weapons of mass destruction, no money down, leave a message at the tone, ...

Including these items in the lexicon (as they must be—where else would they be in the language?) leads to two important conclusions.

First, lexical items cannot be conceived of as syntactic atoms, since many items in (9) have internal syntactic structure. *Kick the bucket* is a transitive VP, *clean as a whistle* is an NP with a comparative complement, *weapons of mass destruction* is a complex NP, and so on. Thus they cannot be inserted by a process like MP's Merge, which builds structure out of syntactic atoms. However, treated as interface constraints, they pose no problem: they simply link a complex syntactic structure with an idiosyncratic meaning. (This approach is shared with HPSG.)

Second, the lexicon cannot be conceived of as a nonredundant list of exceptions, as Chomsky has often asserted (citing Bloomfield). The lexical item *weapons of mass destruction* contains four independently attested words, meaning exactly what they ought to mean. It adds the information that these four form a known unit, and adds some extra meaning or connotation. It is impossible to extract the redundant information, leaving only the extra information, and end up with something that is formally coherent. The conclusion is that the lexicon is full of redundancy. In terms of formal elegance this is less than satisfactory, but it is where the facts urge us. In terms of "brain" elegance, though, it seems entirely in line with the rest of the brain, which seems to favor redundancy where possible, in the interests of more reliable memory and processing.

In addition to items such as (9) that are larger than a word, the PA's lexicon also contains items that are smaller than a word. For example, the regular plural suffix *-z*/*-s*/*-əz* in English establishes a correspondence between a piece of phonology, a syntactic feature, and a piece of meaning. Its contextual restrictions state that it is to be affixed to a noun (syntactic context) that is count (semantic context); the conditions for its allomorphy depend on its phonological context. It can be affixed to a noun of any phonological shape, including novel ones (as in the *wugs* test). Thus its manner of combining with its host is formally no different from the way a transitive verb combines with its object, except that it combines below the word level rather than at the phrasal level.

On the other hand, irregular plurals (*oxen*, *women*, *axes*, etc.) have to be learned individually and therefore have to be stored in the lexicon. Formally, they are semantically and syntactically composite, but phonologically unitary. They are therefore parallel in structure to idioms, which are phonologically and syntactically composite but semantically unitary. We can therefore think of these cases as "morphological idioms". (There may of course be subregularities among irregular forms, but we set this aside for purposes of the present chapters; see Jackendoff 2002, sections 6.2–6.4.)

The treatment of regular inflectional morphology as lexical items extends easily to other regular morphological phenomena, including unusual ones. For instance, English expletive infixation (*manu-fuckin-facturer*) is a stored morpheme with a distinct (non-truth-conditional) meaning, and can be affixed to any syntactic category. Its main contextual restriction is prosodic. Similarly, reduplicative morphemes have meanings and syntactic contextual restrictions just like any other affix, but their phonological shape is listed in the lexicon as a sort of binding: "Copy such-and-such a part of the word I'm attached to" (Ghomeshi et al. 2004).

PA's treatment of regular morphology parts company here with "lexicalist" theories such as LFG and HPSG, which derive morphologically complex words "in the lexicon", "prior to" inserting them into sentences. In PA, both phrasal grammar and morphology contain processes of free combination that can be used online, and both also include lexically listed "prefabs" (idioms and irregular morphological combinations respectively). The difference between phrasal grammar and morphology is only that units and principles of combination for phrases are in part different from those for words. In this framework, LFG's notion of Lexical Integrity amounts to the claim that the two sets of principles do not interact, except through inflectional morphology.

PA's lexicon also incorporates the insight of Construction Grammar that certain pieces of syntax can carry idiomatic meaning, with or without overt morphemes that mark the constructional meaning. Some of these constructional idioms have ordinary syntax, for instance the VP constructions in (10); others, such as (11), have unusual syntax ("syntactic nuts" in the sense of Culicover 1999).

(10) a. joke your way into the meeting (V Pro's way PP = 'go PP while/by V-ing')
  b. rumble around the corner (V PP = 'go PP in such a way to make a V-ing sound')
  c. knit the afternoon away (V NP away = 'spend NP[time] V-ing')
  d. paint me a picture (V NP_1 NP_2 = 'V NP_2 for the benefit of NP_1')

(11) a. The more you eat, the fatter you get (the more S, the more S)
  b. One more beer and I'm leaving (one more X and S)
  c. student after student (N P N)
  d. How about some lunch? (How about XP?)

Each of these constructions is listed in the lexicon as a linking between a syntactic complex and a meaning; some parts of the syntactic complex may be linked also to phonology (e.g., *way*). The syntactic variables in these constructions correspond to semantic variables in the usual way, and the constructions can therefore be combined with other items to form a sentence in exactly the same way as words and other idioms are. (However, notice that the verbs in (10), though they are syntactic heads, serve semantically as manner or means modifiers.)

Since the lexicon contains linked phonological, syntactic, and semantic complexes, nothing in principle prevents it from also containing phonological and syntactic complexes that are not inherently linked to anything. For example, a "generative" phrase structure rule such as (12a)—which, as part of one's knowledge of English, must be stored in memory somehow—can also be stated as a "treelet" (12b), a syntactic complex that constrains possible syntactic structures. PA treats it as a stored piece of structure; it can therefore be localized in the lexicon alongside semantically and phonologically linked VPs such as *kick the bucket*.

(12) a. VP 6 V — NP
 b. $[_{VP}$ V NP$]$

Thus, to the extent that there are autonomous principles of syntax such as fixed head position, the availability of ditransitive constructions, the means for forming relative clauses, and so on, these are stated in precisely the same format as constructional idioms, and they therefore belong in the lexicon as well. In phonology, one can view syllable structure rules as lexical entries that specify pieces of autonomous phonology.

The upshot is that there is no principled distinction between words and rules of grammar. Both are stored pieces of structure, lying at opposite ends of a multidimensional continuum of idiosyncrasy and regularity. This conclusion has been arrived at within HPSG (Pollard and Sag 1994), Cognitive Grammar (Langacker 1987), and Construction Grammar as well as PA, in each case through attention to a multitude of intermediate cases such as idioms and constructions. Mainstream generative grammar, partly because of its algorithmic formulation, has followed traditional grammar in making a strong lexicon/grammar distinction. This has made it difficult to assimilate idioms and constructions into the theory, resulting in loss of descriptive adequacy.

In pursuit of explanatory adequacy, the MP has arrived at the conjecture that there is actually only one rule of grammar, Merge, and that all differences among languages are localized in the lexicon (Chomsky 2001); this conjecture has not proven as simple in execution as in principle (particularly since MP has *no* theory of the organization of the lexicon!). Within PA, HPSG, and Construction Grammar, the counterpart of this conjecture is quite straight forward. All words and all rules of grammar are pieces of structure stored in the lexicon. The only "procedural" part of language is the fundamental operation of Unification (Shieber 1986), which assembles pieces of structure. Merge proves to be a special case of Unification: it combines two given elements with a piece of tree structure.

Unification can be generalized to combinatorial cognitive capacities other than language, thus better satisfying the goal of "beyond explanatory adequacy". For example, in vision it can be used to integrate evidence for depth perception from disparate sources. It can also be used to weld lyrics to music in building songs. Merge cannot perform either of these functions. If Unification is a general brain mechanism for achieving combinatoriality, it should be no surprise that language uses it too. (See Jackendoff in press for discussion of Merge vs. Unification.)

23.3 Conceptual Semantics

To work out any version of a parallel architecture, it is necessary to have theories of the individual components and the interfaces among them. Unlike other parallel architectures in the literature, and unlike mainstream linguistic theory, PA is grounded in a highly articulated theory of semantics, Conceptual Semantics, that answers to the concerns of the biolinguistic perspective and that also offers considerable (and continually increasing) empirical coverage. There is space here only to list some of the highlights of the theory.

First, Conceptual Semantics (like Cognitive Grammar) is thoroughly mentalistic: it is a theory of the information in a language user's mind/brain that is involved in understanding utterances, connecting them to perceptual evidence, and making inferences. It recasts the traditional philosophical concerns with reference and truth in mentalistic terms:

(13) a. Traditional formulation:
- i. A phrase P refers to an entity E in the world (or in a possible world).
- ii. A sentence S is true if it meets conditions $C_1, \ldots, C_n$ in the world.

b. Mentalistic formulation:
- i. A language user LU understands a phrase P to refer to an entity E in the world as LU conceptualizes it.
- ii. LU judges a sentence S true if S meets conditions $C_1, \ldots, C_n$ in the world as LU conceptualizes it.

The seeming objectivity of language, stressed by traditional philosophy of language, is a consequence of language users sharing a common (or near-common) conceptualization of the world, so that agreement can largely be taken for granted (Jackendoff 1983; 2002, chapters 9 and 10).

Second, Conceptual Semantics recognizes that many aspects of one's conceptualization of the world are independent of language. For instance, one can understand much of the behavior of physical objects ("naive physics") without any language at all. Decades of research on child development, linguistic and nonlinguistic, have shown that prelinguistic children bring a rich toolkit to the task of understanding the physical world, and that this understanding serves as a foundation for learning word meanings (e.g., solving Quine's *gavagai* problem). Thus the view of meaning espoused by Conceptual Semantics offers the potential of explanatory adequacy, i.e., helping to explain the innate basis from which children acquire lexicons (now including rules of grammar).

It also appears that other primates—especially apes—negotiate the physical world in much the same way we do; humans differ only in being able to talk about it. This provides an evolutionary underpinning for the semantic system of

978–0–19–954400–4 Heine-main-drv Heine-Narrog (Typeset by Spi, Chennai) 656 of 778 April 8, 2009 21:55

language: our ancestors had thoughts—as it were, things to talk about—before they could talk. This view of meaning, then, helps satisfy the goal of "beyond explanatory" adequacy: it helps explain why (some part of) the semantic system of language is the way it is, because it is built upon pre-existing primate cognition.

Within the MP, by contrast, the combinatorial properties of the "conceptual–intentional interface" arise through derivation from the syntactic component. On the face of it, this amounts to the claim that babies and apes cannot think combinatorially. It is possible to read certain passages of Chomsky as endorsing such a claim, but to my knowledge it has not been defended against the copious literature on primate intelligence. In a recent passage, Chomsky (2006) says "unbounded Merge provides only a language of thought, and the basis for ancillary processes of externalization". In a way this acknowledges the combinatorial character of thought, but it still does so in syntactocentric terms: the basic units of his "language of thought" are NPs and VPs; and Merge, the capacity for combinatoriality, is said to have arisen in the course of human evolutionary divergence from other primates.

In PA, by contrast, the "language of thought" is the combinatorial system in terms of which one understands the world. Its units are entities such as objects, events, properties, and trajectories. NPs and VPs are part of the combinatorial system of (narrow) syntax, which plays a role in the mediation between thought and sound, that is, as part of what Chomsky calls "processes of externalization". PA takes the combinatorial system of meanings to be universal (though use of the system can be biased by the means of expression if "Whorfian" effects prove to be genuine). It is just that meanings are not made of syntactic units. This approach is possible precisely because of the fundamental assumption of PA that language—and the mind in general—utilizes multiple sources of combinatoriality.

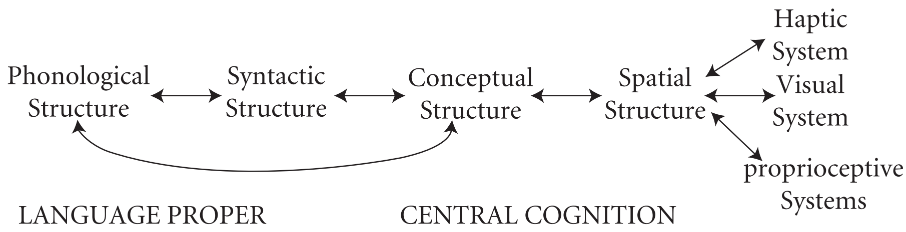
A third important aspect of Conceptual Semantics, again drawing on the Parallel Architecture, is that the system of meaning or "language of thought" is itself bifurcated into two linked combinatorial systems (at least). One of these, Spatial Structure, is quasi-geometric or topological in character. For a first approximation, it might be thought of as the highest level of the visual system. At this level, objects can be represented in terms of their detailed shape. However, shapes are encoded in a perspective-independent fashion, so that they can be recognized from any angle. Objects can also be represented schematically, so that, say, the action of sitting can be represented in terms of a generic or schematic human figure rather than a specific person.

In fact, though, Spatial Structure is not exclusively visual: it can also code shape and configuration that has been derived haptically (sense of touch) and proprioceptively (body sense), and both of these can be compared and integrated with visual input. Thus Spatial Structure is more abstract and general than a visual image—it is conceived of as a central level of cognition that codes the physical world in a relatively modality-independent fashion.

The second major division of meaning is Conceptual Structure, an algebraic structure built up in terms of discrete features and functions. It encodes distinctions that cannot be represented in the geometric/topological format of Spatial Structure, such as those in (14).

(14) a. the type-token distinction, distinguishing categories from individuals
b. taxonomic relations: 'X is an instance/subtype of Y'
c. temporal relations: 'X is past/future'
d. causal relations: 'X causes Y', 'X enables Y', 'X impedes Y', ...
e. modal notions: 'X is hypothetical/nonspecific/potential/fictional ... '
f. social notions: 'X is the name of Y', 'X is dominant to Y', 'X is kin to/friend of Y', 'X is member of group Z', 'X owns Y', 'X is obligated to perform act Y', 'action Y is of normative value Z', ...
g. theory of mind notions: 'X believes Y', 'X imagines Y', 'X intends Y', 'X is committed to norm Y', ...

The overall architecture looks like this:



Conceptual Semantics takes it that word meanings must be composite in order to encode relations among word meanings and in order to state properly general rules of inference. On the other hand, it differs from classical views of word meaning in admitting conditions other than necessary and sufficient. For instance, the conditions for color words must be encoded in terms of relative distance from central instances. In judging a hue between focal red and focal orange, two such conditions come into competition, and the judgment is therefore variable and to some degree context-dependent.

In addition, many word meanings contain multiple conditions interacting in "preference rule" fashion. For instance, sterotypical *climbing* involves moving (i) upward, (ii) in a clambering fashion. But one can *climb <u>down</u> a tree* (clambering but not moving upward), and an airplane can *climb into the clouds* (moving upward but not clambering). On the other hand, an airplane cannot *climb down out of the clouds*, because such motion is neither upward nor clambering. In other words, neither condition is necessary, either may be sufficient, and stereotypical cases satisfy both. This type of rule interaction produces so-called "cluster concepts", of which Wittgenstein's (1953) example of *game* is the most famous.

These characteristics of word meanings, even if strange according to standard philosophical preconceptions, are totally normal within the context of brain computation. As has been observed since the gestalt psychologists of the 1920s (Wertheimer 1923), conditions based on central instances and rule interactions with the characteristics of preference rules are standard in vision. They also appear in phonetic perception and in musical cognition, and essentially anyplace that multiple factors can either combine or conflict in producing a judgment.

Conceptual Semantics differs from most theories of semantics (but again, not from Cognitive Grammar) in that it denies a sharp division between linguistic meaning and encyclopedic meaning (or "knowledge of the world"). Every division that has been proposed turns out to eviscerate linguistic meaning to the point where it cannot serve as a basis for inference (see Jackendoff 2002, sections 9.6–9.7, as well as Bolinger 1965, Langacker 1987, and Levinson 2000).

A related point is that "semantics" and "pragmatics" do not involve distinct representations. Rather, there is a pair of mental representations, Conceptual Structure and Spatial Structure, that are the locus of sentence understanding. Some parts of these representations may come from the words in the sentence and their grammatical configuration; we may call these parts "semantic". Other parts come from nonlinguistic sources such as perception, inference, and "world knowledge"; we may call these parts "pragmatic." But these parts are often intricately interwoven in the representation in such a way that one cannot do the "semantics" first and paste in "pragmatics" afterward.

In Conceptual Semantics, the taxonomy of concepts ("a poodle is a kind of dog", "a dog is a kind of animal", etc.) grounds out in a fundamental ontology of concepts—the basic types of things that humans can conceptualize in the world. Traditional philosophy of language and formal semantics attempt to make do with an absolutely minimal ontology such as individuals and truth-values. Perhaps this makes sense if one thinks semantics is about the nature of reality and should ground out elegantly in fundamental physics. But if semantics is about the human conceptualization of the world, its fundamental units are the product of evolution building a brain equipped to guide an organism successfully through its life. Again "brain elegance" takes precedence over formal elegance.

One piece of evidence for the basic ontology comes from deictic expressions that pick out units in the visual field. Just as it is possible to point out objects for the hearer to identify, as in (15a), it is possible to pick out a wide range of other entities.

(15)	a. Please pick that [pointing] up.	[object]
	b. Please put your hat here [pointing].	[location]
	c. He went thataway [pointing].	[path/trajectory]
	d. Please don't do that [pointing] around here any more.	[action]

e.	Did you hear that?	[sound]
f.	I hope that [pointing] doesn't happen again.	[event]
g.	The fish I caught was this long [demonstrating].	[distance]
h.	There were about this many [gesturing] people here last night.	[amount/number]
i.	Can you walk like this [demonstrating]?	[manner]

Each of these ontological categories has its own conditions of individuation; many of them (but not all) allow a type-token distinction; many permit quantification. Adopting this relatively rich system from the start affords Conceptual Semantics a broad descriptive capacity and, to some extent, a better constrained relation between semantic and syntactic categories. Note also that (15) lists only ontological categories observable in the physical world; there are clearly others, such as information and value.

Once the ontological system is laid out, it becomes possible to recognize entities that subsist simultaneously in more than one ontological domain (the "dot-objects" of Pustejovsky 1995). For instance, a *book* is simultaneously a physical object and a body of information. These two characterizations, moreover, are in a preference rule relation, since there are blank (i.e., informationless) books and books stored on a computer (i.e., not laid out on paper pages). *Reading* is a "dot-action", in that it involves both the physical act of scanning the page with one's eyes and the informational act of receiving information off the page. Dot-objects are therefore multidimensional entities within Conceptual Structure.

Perhaps the most important case of a dot-object is a human being, who is conceptualized simultaneously as an animate physical object and as a person—an entity in the social domain. The two domains correspond to the (nearly universal) cultural conceptualization of people as composed of body and mind (or soul or spirit). The fact that people have faces and hands and livers falls into the physical domain; the social notions and theory-of-mind notions in (14f,g) above are predicated in the social domain. Again, in traditional beliefs at least, these two characterizations stand in a preference rule relation. For instance, a *zombie* is an animate physical object lacking conscious personhood; a *ghost* is a mind (or soul) lacking a physical body. Reincarnation and body-switching (both amply attested in human narratives) are one mind inhabiting different bodies in succession; multiple personality disorder is experienced as different personalities (i.e., different individuals) inhabiting the same body in succession (Jackendoff 2007*b*, chapter 5).

The combinatorial possibilities of Conceptual Structure arise from (at least) three principles of combination: argument satisfaction, modification, and binding. In the default case, argument satisfaction is expressed by syntactic complementation, and modification by syntactic adjuncts. For instance, in *John slept along the river*, *John* expresses an argument of *sleep*, and *beside the river* expresses a place modifier. However, there are exceptions to this typical configuration. For instance, in the sound+motion construction illustrated in (10b) above (e.g., *The trolley rumbled along the river*), the subject is a semantic argument not only of the verb but also of an unexpressed predicate of motion. The PP is also an argument of the predicate of motion, and the verb expresses a modifier of this predicate, i.e., "move *while rumbling*". A mismatch in the opposite direction is illustrated by *Bill buttered the bread with cheap margarine*. Here *cheap margarine* is syntactically an adjunct, but semantically it is an argument: it is what Bill put on the bread. Such mismatches are common.

Binding, a direct connection between one conceptual constituent and another, comes in two varieties: identity of reference and identity of sense. This is reflected in two kinds of anaphoric elements in language. Identity of reference binding is expressed by definite pronouns and also by anaphoric epithets, such as in *John wants to win, but the poor guy never will* (which does not display identity of sense). Identity of sense binding is expressed by *one*-anaphora and also by VP anaphora with expression like *do so*. These two types of binding must be distinguished in Conceptual Structure since they give rise to different inferences.

Using argument satisfaction to create semantic combinations requires functions whose arguments are to be satisfied. A number of broad families of functions have been investigated within Conceptual Semantics:

- Functions that encode spatial location, motion, and orientation. They all take two arguments: a Theme (the object being located or in motion) and a Location or Path: BE(Theme, Loc), GO(Theme, Path), STAY(Theme, Loc), ORIENT(Theme, Path), EXTEND(Theme, Path).
- Functions that encode Locations and Paths relative to a reference object: IN(X), ON(X), TO(X), FROM(X), TOWARD(X), NEAR(X), etc. Some of these involve imposing a reference frame on the reference object; e.g., BEHIND(X) must be specified as to whether one is speaking of the intrinsic back of X or its other side relative to the speaker. (This family has been heavily investigated within Cognitive Grammar as well.)
- Causative functions that encode a Causer (an Agent or Event) being causally connected to an Effect (another Event): CAUSE(Causer, Effect), LET(Causer, Effect), HELP(Causer, Effect), ENABLE (Causer, Effect), and others.
- Mereological functions that encode part–whole relations: PART-OF (legs, handles, noses), BOUNDARY-OF (edges, surfaces, ends, etc.), MEMBER-OF (members of aggregations), COMPOSED-OF (ingredients of mixtures).

A founding insight of Conceptual Semantics (due to Gruber 1965) is that all of these functions can be applied to semantic fields other than physical space. For instance, an object being owned by someone (a social relation) is often expressed crosslinguistically as the object "being at" the owner, and changes of possession are often expressed as the object "going" "from" the previous owner "to" the new owner. Similarly, just as we talk about the end of a rope, we can talk about the end of a speech, a relationship, or a genealogical line. This suggests that these Conceptual functions can be decoupled from their physical context (where they connect with Spatial Structure) so as to apply to more abstract domains as well. In addition to possession, they also extend to such fields as time, event structure (such as aspectuality and telicity), ascription of properties, and (in the case of causation) social coercion and logical entailment. (This insight is treated somewhat differently in Cognitive Grammar (Lakoff 1987), where it is taken to show that underlying linguistic expression is an extensive and powerful system of conceptual metaphor.)

Further functions that have been investigated (Jackendoff 2007*b*) involve the personal domain. They include:

- Theory-of-mind predicates, e.g., "X perceives Y (in various modalities)", "X is committed to proposition P" (belief), "X is committed to action A" (intention), "X is committed to norm N" (adherence to norms).
- Value predicates in various domains (affective, normative, quality, etc.): "X is of value V", "X is of value V to person Y".
- Predicates of exchange: "X does action A in exchange/return/retaliation for Y doing action B".
- Obligations, rights, and authority: "X is obligated to Y to perform action A", "X has a right to perform action A", "X has authority over Y's performing action A".

All of these functions are involved in constructing the *propositional tier* of Conceptual Structure. In addition, sentence meaning involves an *information structure tier*, which designates certain semantic constituents as topic, certain as focus, and the rest as common ground. Further differentiation of the propositional tier has also been proposed, for which there is no space here: a *referential tier* in Jackendoff 2002 (involved for instance in identity-of-reference anaphora, specificity, referential opacity, and quantification) and an *action tier* or *macrorole tier* in Jackendoff 1990, 2007*b*.

In short, Conceptual Semantics aspires to the formal richness necessary to encode the character of human concepts and their inferential affordances. It integrates comfortably with the Parallel Architecture, in that, although it is a combinatorial system, its units and principles of combination—as well as the resulting structures—are quite different from those of syntax. In particular, it is a multidimensional formal system, in that it involves both Spatial Structure and Conceptual Structure, the latter itself split into multiple tiers connected by interface components. Only through looking at semantics on its own terms, grounded in the character of nonlinguistic cognition, can the independence of these structures from language—and their psychological and biological grounding—be revealed. If meanings have this sort of structure, they certainly cannot be derived from the syntax of NPs and VPs.

978–0–19–954400–4 Heine-main-drv Heine-Narrog (Typeset by Spi, Chennai) 662 of 778 April 8, 2009 21:55

23.4 Simpler Syntax and the syntax–semantics interface

An advantage of a parallel architecture over a "single-engine" architecture is that no single level of structure has to carry the entire informational load. In a syntactocentric architecture, all semantic combinatoriality has to be derived from syntactic combinatoriality. Thus syntax is forced to be combinatorially at least as complex as semantics—if not more so, since it also has to answer to its own internal imperatives such as word order and agreement. And indeed this outcome has been achieved twice in the history of generative grammar: the first time, in the Generative Semantics movement of the late 1960s and early 1970s (Lakoff 1971), and the second time, in Government-Binding Theory of the 1990s and the Minimalist Program. In MP, the rules of grammar and the contents of UG have been reduced to a minimum (allegedly—though only through drastic cuts in empirical coverage), but the structures and derivations have increased steadily in size (see Culicover and Jackendoff 2005, chapters 2 and 3).

In PA, the combinatorial properties of meaning are a property of autonomous conceptual combinatoriality. From this perspective, syntax functions in the grammar not as the fundamental generative mechanism but rather as an intermediate stage in the mapping between meaning and sound (in either direction). Words are interface rules that provide small-scale mappings between meaning and sound. What remains to complete the mapping is the relationships among the words: the function–argument and function–modifier relations, as well as binding relations. Syntax can be thought of as a way of recoding the semantic relationships among the words in a phrase or sentence in terms that are visible to phonology, such as linear order, inflectional morphology, and anaphoric elements—as well as coding the overall semantic force of a clause, such as declarative vs. interrogative. However, there is no need for syntax to encode any more of semantic structure than is necessary in order to mediate the mapping between phonology and meaning.

In fact, many aspects of meaning are not supported by syntactic or lexical expression, for instance:

(16) a. Implicature:
Are you going to be going near a mailbox? (= "Will you mail some letters for me?")

b. Ellipsis:
It seems we stood and talked like this before. We looked at each other in the same way then. But I can't remember <u>where or when</u>. [Spoken to someone about to jump off a building] <u>Don't!!!</u>

978–0–19–954400–4 Heine-main-drv Heine-Narrog (Typeset by Spi, Chennai) 663 of 778 April 8, 2009 21:55

c. Constructional meaning:
The trolley rumbled around the corner. (="The trolley went around the corner rumbling") (cf. (10b))

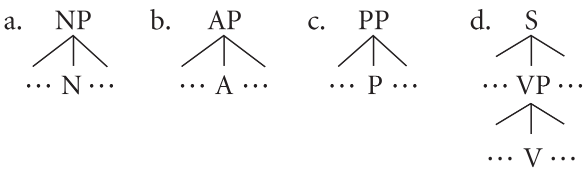
d. Coercion:
The ham sandwich over in the corner wants more coffee. (="guy with ham sandwich")
Plato is on the top shelf. (="book by/bust of Plato")
Joe jumped until the bell rang. (="jumped repeatedly")

Some of these are treated in mainstream theory in terms of syntactic (or PF) deletion of unexpressed elements; others are not treated in mainstream theory at all. Culicover and Jackendoff 2005 show that they are all best treated in terms of elements of semantics that have no syntactic realization.

Carrying this outlook consistently through the syntactic component leads to the approach of Simpler Syntax (Culicover and Jackendoff 2005): an attempt to cut syntactic structure down to the bare minimum necessary to accomplish the sound–meaning mapping. This is a "minimalist" approach to language, but with different premises about what is to be minimized than the Minimalist Program.

The basic stance of Simpler Syntax is that the complexity of semantics is independently necessary in order to explain inference and the relation to perception. Therefore semantics should play as large a role as much as possible in constraining grammaticality, and syntax as little as possible. On the other hand, the "generative engines" of syntax and morphosyntax are still necessary to account for differences among languages in word order, case marking, agreement, handling of long-distance dependencies, and the existence of special constructions. The resulting syntactic theory is by no means simple, but it is far simpl*er* than mainstream models.

The Simpler Syntax lexicon is as described in section 23.3: it contains words, regular affixes, idioms, constructions, and independent principles of phrase structure. Syntactic structures are as flat (i.e., as undifferentiated) as possible. Aside from linear order, there is no *syntactic* distinction between specifiers, arguments, and adjuncts, as this is already provided for in the semantics. The result is predominantly two-layer X-bar skeleta, as in (17a–c). The exception is S, which is a three-layer projection of V, as in (17d).

(17)


One price of this structural simplification is the need for trees with multiple branching nodes rather than strictly binary branching as in MP. Culicover and

Jackendoff 2005 give arguments why strictly binary branching is not an advantage, and in fact is often a disadvantage. Another price of this simplification is that some rules of grammar have to be sensitive to linear order as well as dominance. This is too often taken to be a disadvantage. But from a larger perspective it is actually an advantage. Linear order is given for free in the signal and hierarchical structure is not. So rules that depend in part on linear order ought actually to be easier for the child to learn.

Simpler Syntax makes use of almost no empty nodes in syntactic structure. This is desirable in principle, because empty nodes make heavier demands both on the learner and on processing. Most empty nodes in the classical theory are posited either for semantic reasons or to promote syntactic uniformity. For instance, the phonologically empty element PRO is posited to fill in a semantic subject of an infinitival VP where there is none at the surface, thereby giving all verbs a syntactic subject. Simpler Syntax instead allows infinitival VPs without syntactic subjects, and it uses the interface to identify their "understood" subjects in Conceptual Structure.

Similarly, ellipsis is not derived through empty nodes or deletion. Rather, elliptical configurations, especially when they are syntactically unusual (as in Gapping), are treated as meaningful constructions listed in the lexicon. The interpretation of an elliptical construction is derived from the Conceptual Structure of its antecedent—or from the Conceptual Structure of the context—not from a deleted syntactic structure. Culicover and Jackendoff show many cases of ellipsis for which there is no plausible syntactic antecedent, such as those in (16b).

A standard argument for syntactically derived ellipsis is that elliptical constructions often display syntactic properties that normally can arise only through syntactic licensing (so-called *connectivity*). For instance, in the dialogues in (18), the difference in the prepositions in the replies can be traced directly to the difference between the syntactic licensing of *proud* vs. *pride*.

(18) a. A: Bill is very proud.
B: Yes, especially <u>of</u> his stamp collection. [cf. *proud of/*in*]

b. A: Bill has a lot of pride.
B: Yes, especially <u>in</u> his stamp collection. [cf. *pride in/*of*]

However, similarly licensed syntactic properties appear even in sentences where there is no relevant linguistic context, such as *Do you like these?* [pointing at a pair of pants]. Simpler Syntax proposes a relation of *indirect licensing* that accounts for these effects.

Like other constraint-based theories, Simpler Syntax has no movement and no covert level of syntactic structure such as Logical Form. The effects ascribed to movement in mainstream theory are accounted for with a variety of mechanisms,

978–0–19–954400–4 Heine-main-drv Heine-Narrog (Typeset by Spi, Chennai) 665 of 778 April 8, 2009 21:55

most of them shared with other constraint-based theories, especially HPSG. These mechanisms include:

- Free phrase order (e.g., among adjuncts in VP, where the order is constrained only by prosody and focus).
- Alternative argument realizations (e.g., dative alternation).
- For long-distance dependencies, operator–trace relations along the lines of HPSG (trace is the only kind of empty node in Simpler Syntax). The constraints on long-distance dependencies arise from multiple sources, only some of which are syntactic. Others arise from processing complexity and from semantics, especially information structure and referential structure.
- Binding and control are relations over Conceptual Structure, not over syntactic structure, though they may involve syntactic conditions on the relation between anaphoric elements and antecedents.

In order to account for so-called A-movements, in particular passive and raising, it is unfortunately necessary to introduce extra machinery. Simpler Syntax proposes a *grammatical function tier* (*GF-tier*) that modulates the syntactic realization of semantic arguments expressed as NPs, that is, subjects, objects, and indirect objects. We are not too dismayed by this extra mechanism, as the principles behind it appear in every substantive syntactic theory: as f-structure in LFG, as essentially *all* of Relational Grammar, as the complement hierarchy in HPSG, and as abstract case in GB/MP.

The analysis is closest to that in LFG and HPSG. However, in these two theories, passive is a rule that converts active verbs into passive verbs in the lexicon, altering their argument structure. As mentioned earlier, this is not an option in PA, where the lexicon is where items are stored, and working memory is where structures are built online. Hence, in Simpler Syntax, passive is treated as a construction that alters argument realization online without altering the verb itself. The GF-tier is of course another piece of parallel architecture, this time a partial mediator of the syntax–semantics interface.

(19) illustrates the linking between the various structures in an example involving raising. The linking relations are notated as subscripts; for visual clarity, some of them are also notated redundantly by vertical association lines.

(19) *John seems to like scotch*:

$[\mathrm{SEEM}\ ([\mathrm{LIKE}\ (\mathrm{JOHN}_3,\ \mathrm{SCOTCH}_4)]_2)]_1$	Conceptual Structure
$[\mathrm{GF}_3]_1 \quad [\mathrm{GF}_3 > \mathrm{GF}_4]_2$	Grammatical Function Tier
$[_{\mathrm{S}}\ \mathrm{NP}_3\ [_{\mathrm{VP}}\mathrm{V}_1\ [_{\mathrm{VP}}\ \mathrm{to}_5\ \mathrm{V}_2\ \ \mathrm{NP}_4]_2]]_1$	Syntactic Structure
$\mathrm{John}_3\ \ \mathrm{seems}_1\ \mathrm{to}_5\ \mathrm{like}_2\ \mathrm{scotch}_4$	Phonological Structure

In Conceptual Structure, *JOHN* is an argument of *LIKE*. It links to the GF array associated with the subordinate clause (bracketed expression subscripted *2*). In turn, this GF is linked to a GF in the main clause array (subscript *1*), which is then linked to the subject of the main clause and its phonology. The linking through the GF-tier is the Simpler Syntax counterpart of an A-chain in classical syntax. But it is not in syntax proper, as there is no *syntactic* subject at all in the subordinate clause, only a GF-subject. (See Culicover and Jackendoff 2005 for more motivation and detail.)

23.5 Concluding remarks

An abiding issue between linguists and psycholinguists has been the competence–performance distinction. Mainstream linguistics tends to say that the grammar written by linguists is a description of competence, but it is somewhat obscure how it is utilized in performance. This has the effect of insulating linguistic theory from results in psycholinguistics. By contrast, in the Parallel Architecture, language processing consists of assembling pieces of structure stored in the lexicon to form a triple of phonological, syntactic, and semantic structures in working memory. As a result, there is no mystery to the competence–performance distinction. Competence theory describes the pieces of structure and their affordances for assembly, while performance theory describes how these very pieces are assembled in real time, starting from either phonetic input (perception) or conceptual input (production). Details of a performance model in such a vein appear in Jackendoff 2002, chapter 7 and Jackendoff 2007*a*.

The Parallel Architecture also offers an attractive vehicle for discussion of the evolution of the language capacity. It begins with the premise that some version of Conceptual Structure is present in apes, and therefore in our hominid ancestors. Bickerton 1990 and Givón 1979 have proposed that, prior to the development of modern language, there was a stage of "protolanguage", which persists in the human language capacity and emerges in situations such as pidgins and agrammatic aphasia. The defining characteristics of protolanguage are words concatenated into utterances, but lacking any syntactic organization beyond that afforded by linear order. A great deal of the informational load in such an utterance is carried by pragmatics. Within the Parallel Architecture, this form of language can be characterized in terms of a level of phonology linked to Conceptual Structure without the intervention of syntactic structure (Jackendoff 2002, chapter 8).

From this stage, the evolution of a syntactic capacity can be seen as adaptive: it is a canonical coding of semantic relationships among words for greater accuracy and efficiency. In any architecture, phonological and semantic structures have to be relatively rich, as they code the thousands of distinctions among words. In Simpler Syntax, syntactic structure is relatively lean: its elements comprise only a few parts of speech and phrasal categories, as might be expected of a relatively late evolutionary add-on. By contrast, in the mainstream architecture, an elaborate syntax would have had to evolve first before combinatorial phonology and semantics could be possible, a rather less enticing scenario.

To sum up, this chapter has shown many ways in which the Parallel Architecture, with its components Conceptual Semantics and Simpler Syntax, instantiates the biolinguistic outlook better than does the Minimalist Program. In particular, it offers the prospect of integrating linguistics fully with cognitive science. There still remain, of course, many challenges to the approach, among which perhaps the most important are integrating phonology, morphology, language variation, and language change into the model, so that it covers a broader range of linguistic phenomena. In addition, a theory of language acquisition has been sketched (Jackendoff 2002, chapter 6), but it remains a promissory note. It is dearly to be hoped that some of these challenges can be undertaken by practitioners of the relevant subdisciplines.

References

Bickerton, Derek (1990). *Language and Species*. Chicago: University of Chicago Press.

Bolinger, Dwight (1965). "The atomization of meaning". *Language* 41: 555–73.

Bresnan, Joan W. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.

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

—— (2001). *Beyond Explanatory Adequacy*. (MIT Occasional Papers in Linguistics, 20.). Cambridge, MA: MIT Dept. of Linguistics.

—— (2006). "General introductory remarks", in M. Piattelli-Palmarini, P. Salaburu, and J. Uriagereka (eds.). *Of Minds and Language: The San Sebastian Encounter with Noam Chomsky*. Ms., in preparation for Oxford University Press.

Culicover, Peter W. (1999). *Syntactic Nuts: Hard Cases in Syntax*. Oxford: Oxford University Press.

—— and Ray Jackendoff (2005). *Simpler Syntax*. Oxford: Oxford University Press.

Fillmore, Charles (1988). "The mechanisms of Construction Grammar". *Berkeley Linguistics Society* 14: 35–55.

Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Ghomeshi, Jila, Ray Jackendoff, Nicole Rosen, and Kevin Russell (2004). "Contrastive Focus Reduplication in English (The Salad-Salad Paper)". *Natural Language and Linguistic Theory* 22, 307–57.

Givón, Talmy (1979). *On Understanding Grammar*. New York: Academic Press.

Goldberg, Adele (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gruber, Jeffrey (1965). *Studies in Lexical Relations*. Ph.D. dissertation, MIT. Published in Gruber, *Lexical Structures in Syntax and Semantics*. Amsterdam: North-Holland, 1976.

Jackendoff, Ray (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.

—— (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

—— (1990). *Semantic Structures*. Cambridge, MA: MIT Press.

—— (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

—— (2002). *Foundations of Language*. Oxford: Oxford University Press.

—— (2007*a*). "A parallel architecture perspective on language processing". *Brain Research* 1,146: 2–22.

—— (2007*b*). *Language, Consciousness, Culture*. Cambridge, MA: MIT Press.

—— and Fred Lerdahl (2006). "The capacity for music: What's special about it?" *Cognition* 100, 33–72.

—— (in press). "Alternative minimalist visions of language". To appear in *Proceedings of Chicago Linguistics Society 41*. Also to appear in Robert D. Borsley and Kersti Börjars, *Nontransformational Syntax: A Guide to Current Models*. Oxford: Blackwell.

Joshi, Aravind (1987). "An introduction to tree-adjoining grammars", in A. Manaster-Ramer (ed.), *Mathematics of Language*, 87–114. Amsterdam: Benjamins.

Koch, Christof (2004). *The Quest for Consciousness*. Englewood, CO: Roberts.

Lakoff, George (1971). "On Generative Semantics", in D. Steinberg and L. Jakobovits (eds.), *Semantics: An Interdiscipliniary Reader in Philosophy, Linguistics, and Psychology*, 232–96. New York: Cambridge University Press.

—— (1987). *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.

Lamb, Sydney (1966). *Outline of Stratificational Grammar*. Washington: Georgetown University Press.

Langacker, Ronald (1987). *Foundations of Cognitive Grammar*, vol. I. Stanford, CA: Stanford University Press.

Lerdahl, Fred and Ray Jackendoff (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Levinson, Stephen (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Pollard, Carl and Ivan Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Postal, Paul (1972). "The Best Theory", in P. S. Peters (ed.), *Goals of Linguistic Theory*, 131–70. Englewood Cliffs, NJ: Prentice Hall.

Pustejovsky, James (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Sadock, Jerrold (1991). *Autolexical Syntax*. Chicago: University of Chicago Press.

Shieber, Stuart (1986). *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI Publications.

—— and Yves Schabes (1991). "Generation and synchronous tree adjoining grammars". *Journal of Computational Intelligence* 7: 200–28.

Van Valin, Robert, Jr. and Randy LaPolla (1997). *Syntax: Structure, Meanings, and Function*. Cambridge, Cambridge University Press.

Wertheimer, Max (1923). "Laws of organization in perceptual forms". Translated in W. D. Ellis (ed.), *A Source Book of Gestalt Psychology*, 71–88. London: Routledge & Kegan Paul.

Wittgenstein, Ludwig (1953). *Philosophical Investigations*. Oxford: Blackwell.