COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2015) 1–28

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12324

In Defense of Theory

Ray Jackendoff

Center for Cognitive Studies, Tufts University

Received 31 May 2015; accepted 29 July 2015

Abstract

Formal theories of mental representation have receded from the importance they had in the early days of cognitive science. I argue that such theories are crucial in any mental domain, not just for their own sake, but to guide experimental inquiry, as well as to integrate the domain into the mind as a whole. To illustrate the criteria of adequacy for theories of mental representation, I compare two theoretical approaches to language: classical generative grammar (Chomsky, 1965, 1981, 1995) and the parallel architecture (Jackendoff, 1997, 2002). The grounds for comparison include (a) the internal coherence of the theory across phonology, syntax, and semantics; (b) the relation of language to other mental faculties; (c) the relationship between grammar and lexicon; (d) relevance to theories of language processing; and (e) the possibility of languages with little or no syntax.

Keywords: Mental representation; Syntax; Lexicon; Rules of grammar; Language processing

1. Theories of mental representations

At the beginnings of cognitive science in the 1970s, when an interdisciplinary community of researchers began taking seriously the idea of the brain as a kind of information processing device, one of the important philosophical issues the field tried to confront was how to think about mental representations—the information the brain is encoding and computing (e.g., Dennett, 1978; Fodor, 1983; Haugeland, 1981; Margolis & Laurence, 1999). Over the years, my deepest concerns have remained with these basic questions.[†]

Correspondence should be sent to Ray Jackendoff, Center for Cognitive Studies, Tufts University, Medford, MA 02155. E-mail: ray.jackendoff@tufts.edu

[†]This article is largely based on my Rumelhart Prize lecture of the same title. I am deeply honored by this award, and I wish to express my profound thanks to the Cognitive Science Society and the selection committee, as well to Bob Glushko and the Glushko-Samuelson Foundation, who have made this prize possible. I am grateful to Peter Culicover, Eva Wittenberg, Steven Pinker, Jenny Audring, and an anonymous reviewer for many valuable comments on content and style.

Perhaps because of my background in linguistics, I tend to think of mental representations rather concretely:

- What is the repertoire of patterns that an organism can encode, and how does it acquire this repertoire?
- How do perceivers or learners internally represent what they see, hear, feel, smell, imagine, and so on?
- How do they represent their hypotheses about the world, such that they can compare them to the input they receive from the world?

Put in more customary terms of cognitive science:

- What are the different mental representations in terms of which an organism encodes its understanding of the world and formulates its actions?

I do wish to be careful, however, about this standard formulation. The proper way to conceive of mental representation is neither propositional nor intentional in the sense of Fodor (1981). Consider one especially well-understood variety of mental representation, phonological structure. The fact that English has a particular repertoire of speech sounds, a particular set of restrictions on combinations of speech sounds, and certain regularities of stress placement is surely not represented mentally in terms of propositions that can be true or false, defined in terms of truth conditions. Nor is it "about the world" in any straightforward sense: Phonological representations do not *represent* anything. Rather, they constitute part of how English speakers *construct* their understanding of the world and their actions in it. For instance, there are no "spr" clusters in the *physical* world—only in the world *as English speakers construe it*. So my preference is to think of mental representations more austerely, as data structures stored and manipulated in the brain, which in principle could be either digital or analog.

I fear that in recent years the term "mental representation" has often come to be used merely as a lazy abbreviation for something like "the information the brain uses, whatever that may be." Here, I would like to go back to the roots of cognitive science and ask for a thorough and explicit characterization of mental representations/data structures.

In the study of language, the characterization of mental structures is the domain of linguistic theory. Linguists in the mentalist tradition understand the "grammar of a language" as spelling out explicitly the repertoire of phonological, syntactic, and semantic structures in terms of which a speaker can construct utterances in comprehension and production. This tradition originates with generative grammar (Chomsky, 1965), but also includes numerous alternative approaches such as Optimality Theory (Prince & Smolensky, 2004), Cognitive Grammar (Lakoff, 1987; Langacker, 1994), and some strains of Construction Grammar (Goldberg, 1995, 2006).

In vision, David Marr (1982) was very insistent on the need for an explicit characterization of mental structures.[1] He himself formulated an elaborate theory of three kinds of mental representations involved in recognizing objects, and of the connections between these representations. Since Marr, other theories of vision have been proposed, but those that I am familiar with (e.g., Biederman, 1987; Edelman & Weinshall, 1991; Poggio &

Edelman, 1990; Riesenhuber & Poggio, 2000; Tarr, 1995; Ullman, 1998) fail to scale up to the full complexity of the problem of vision.

In another domain, Fred Lerdahl and I have developed a detailed theory of the representations involved in music—the structures that listeners construct that make music more than just a sequence of notes– and the principles listeners use to construct these structures (Jackendoff & Lerdahl, 2006; Lerdahl & Jackendoff, 1983). In yet another capacity, Neil Cohn proposes a theory of narrative structure that applies equally to linguistic narrative and narratives presented visually, for instance comic strips and Japanese manga (Cohn, 2013).

I think other domains of cognitive science call for the same approach. For instance, if one is studying episodic memory (e.g., Tulving, 1983, 2002), one ought to be asking what counts as an episode and how it is encoded. What are the units from which encoded episodes are constructed, and how are they composed? In what *form* do I code my memory of where I parked my car this morning or what transpired at the department meeting yesterday?

If one is studying embodied or grounded cognition (e.g., Barsalou, 2008; Clark, 2008; Varela, Thompson, & Rosch, 1991; Wilson & Foglia, 2011), one ought to be asking in what form our experience and especially our bodies are encoded by our brains—not to mention how the brain encodes all the other things that are claimed to be understood with reference to the body, such as logical reasoning and moral thought, and precisely how these encodings are related to the encoding of the body.

If one is studying item-based or exemplar-based learning—if one thinks that learning consists of encoding all the instances of whatever a person has experienced in some domain, and that understanding novel inputs is a matter of analogy or interpolation among instances (e.g., Bod, 2006; Bybee & McClelland, 2005; Pierrehumbert, 2002; Poggio & Edelman, 1990; Tomasello, 2003)—then one ought to be asking: Exactly how does the mind/brain encode instances it has encountered, and what are the dimensions available for encoding them? A crucial constraint on such a theory is that the dimensions in terms of which experience is encoded cannot themselves be learned. They form the *basis* for learning; they are what enable learning to take place *at all*. So they have to somehow be wired into the brain in advance of learning.

If one is studying statistical or probabilistic learning (e.g., Saffran, Aslin, & Newport, 1996; Solan, Horn, Ruppin, & Edelman, 2005), one ought to be asking: What is the repertoire of mental structures over which our brains run their statistics, and what is the form of the outputs of the statistics? If one is studying predictive or hypothesis-based perception (e.g., Hale, 2011; Kuperberg & Jaeger, 2015; Levy, 2008), one ought to be asking: In what terms are predictions and hypotheses encoded, and in what terms is the evidence encoded for hypotheses and for context?

I recognize that there is something right about all these approaches to the mind. We *do* code episodes, we do base abstract reasoning on principles of physical experience, we do learn on the basis of instances, and statistical and probabilistic factors are involved in our perception and behavior. But to my taste, something important is missing from all of these approaches, to the extent that I'm familiar with their literature: an explicit and sufficiently general theory of the mental structures being computed and learned.

I should stress as well that I don't want a theory of mental structures to live in a vacuum. Like Marr (1982), and unlike Chomsky (1965), who sets up a firewall between competence and performance, I want the formal theory of mental structures to embed gracefully in a theory of how moment-to-moment processing takes place and how learning takes place.

My working hypothesis is that memory, processing, attention, and learning are pretty much the same all over the brain. And I believe that structural features like grouping, sequencing—and even recursion—are common to many different kinds of mental representations (Jackendoff, 1987a, 2002, 2011; Lerdahl & Jackendoff, 1983; Pinker & Jackendoff, 2005). But that can't be the end of the story: One must also specify what makes vision different from language, and language different from music (Jackendoff, 2009; Patel, 2008), and music different from actions like washing dishes, and washing dishes different from morality (Jackendoff, 2007a). My working hypothesis here is that these differences are a consequence of the character of the mental structures appropriate to each; and that the similarities that cut across these domains are rather like the similarities we find across domains in the body: such matters as cell structure and metabolism, which are shared across muscle cells, stomach cells, and even brain cells. (See Gallistel & King, 2009, on the conservatism of neural mechanisms.) But this hypothesis cannot be tested seriously without a fair amount of detail about what these different mental structures are like.

It is also incumbent on a proper theory to say how the various forms of mental structure interact with each other to produce our capabilities for understanding and behavior. For instance, as pointed out by John Macnamara (1978), we obviously manage to talk about what we see. This means that the mental structures responsible for visual cognition have to interact with those responsible for language; these faculties cannot be informationally encapsulated from each other in the way that Fodor's (1983) putative language module is supposed to be. More generally, a theory of any faculty has to show how it makes contact with other faculties.

Beyond that, I would like the accounts of processing and learning eventually to be describable in terms of neural computation. As far as I know, at present we don't have the faintest idea of how neurons encode something as simple as a syllable, say /bæ/, much less the meaning of a sentence (Gallistel & King, 2009; Poeppel & Embick, 2005). We may have some idea *where* in the brain speech sounds are encoded, or where they are marshaled into words, but we have no idea *how*. Nevertheless, I think it is important to keep this desideratum in mind while developing theories of mental structure.

2. Considerations on the adequacy of theories

How are we to judge the adequacy of theories of mental structure? Consider mainstream generative linguistic theory (Chomsky, 1965, 1981, 1995), for instance. On the one hand, it has been extremely fruitful in leading to exploration of a huge number of linguistic phenomena in a huge number of languages. On the other hand, it has not fit very well with theories of language processing and language acquisition, and that has led many researchers (though not all by any means) to give up trying to use it to guide experimental work (Edelman & Christiansen, 2003; Ferreira, 2005; Poeppel & Embick, 2005).

This is not the only difficulty: Mainstream linguistic theory has not even been internally coherent. A typical graduate student in linguistics learns phonology in terms of Optimality Theory (Prince & Smolensky, 2004), syntax in terms of the Minimalist Program (Chomsky, 1995), and semantics in terms of truth conditions in possible worlds (Heim & Kratzer, 1998). None of these three theories has any substantive connection to the other two. And for the most part, this disconnect obtains in the theoretical literature as well.

Can this be the right overall theory of language? To me, it seems like a beast with the head of a cat, the body of a camel, and the legs of a turkey. The pieces don't fit together at all, much less relate to the mind in the way generative grammar originally aspired to.

For about 20 years I have been developing an alternative theory of language called the parallel architecture (PA: Jackendoff, 1997, 2002), which offers a more principled integration of the components of language with each other and with more general psychological concerns, including language processing and acquisition. The goal is to retain the insights of research on all the constituent parts, but to find a way to fit them together. Unlike mainstream generative grammar, it includes a substantive theory of meaning, Conceptual Semantics (Jackendoff, 1983, 1990), which is grounded in research on nonlinguistic cognition—both spatial cognition (Jackendoff, 1987b; Landau & Jackendoff, 1993) and social cognition, including things like joint intention, fairness, and morality (Jackendoff, 2007a). At the same time, PA addresses all the traditional concerns of generative syntax, through a theory of syntax and the syntax-semantics interface called Simpler Syntax (Culicover & Jackendoff, 2005). Like Head-Driven Phrase Structure Grammar (HPSG: Pollard & Sag, 1994) and Construction Grammar (Goldberg, 1995; Hoffmann & Trousdale, 2013), Simpler Syntax is much more "surfacey" than standard generative syntax, in that it has no movement, (almost) no empty nodes, and no covert levels of structure such as Logical Form. Jackendoff and Audring (unpublished data), building on work in Construction Morphology (Booij, 2010), extend PA to word structure and to how words, phrases, and the generalizations among them are stored, processed, and learned.

How ought one to adjudicate between these theories? And why does it matter? Each of them is rich and flexible, so finding a crucial experiment that can decide between them is hard, in fact probably impossible. What makes me think the PA's approach is on the right track, though, is that it opens up all sorts of interesting avenues for empirical research, for internal integration of linguistic theory, and for connections with the rest of cognitive science. Here, I wish to sketch four such directions:

- The architecture of the grammar
- The relation of the lexicon and the rules of grammar
- How the preceding two factors impact on a theory of language processing
- Languages that don't have much syntax

3. The architecture of the grammar

I begin by taking Chomsky's (1986) term "knowledge of language" actually more literally than he does—very psycholinguistically. "Knowledge" implies something stored in memory. So in these terms the major questions of generative grammar come out like this:

- What linguistic units and linguistic complexes (i.e., what mental representations) do speakers of a language store in long-term memory? The standard answer is *the lexicon*.
- What computational procedures do speakers use to construct novel composite structures—in particular structures that relate meaning to sound (or, in the case of sign languages, to gesture)? The standard answer is *the grammar*.
- What procedures do language learners use to construct the first two kinds of knowledge? The traditional answer within generative grammar is *Universal Grammar* or *the Language Acquisition Device*.[2]

Any mentalist theory of language has to address these questions. Here, the PA is totally in agreement with standard generative grammar. But it is far from agreement on the answers.

The first major difference is in the overall architecture. The received view is that the combinatorial structure of language arises in syntax, and that the combinatorial structures of phonology and of semantics are derived from syntax. Chomsky (1965, pp. 16, 17, 75, 198) explicitly states that this is an assumption, but he has never defended it against serious alternatives. Nevertheless, already in the 1970s, phonologists showed that phonology has its own independent combinatorial principles (e.g., Goldsmith, 1979; Liberman & Prince, 1977; Selkirk, 1984). And by the 1980s, practically every theory of semantics on the market (e.g., Heim & Kratzer, 1998; Lakoff, 1987; Langacker, 1987; Jackendoff, 1983; Schank, 1973) attributed to meaning its own primitives and principles of combination, independent of syntactic structure.[3]

Mainstream linguistics has for the most part ignored the combinatorial independence of phonology and semantics from syntax. But if we take it seriously, we arrive at a parallel architecture, in which phonology, syntax, and semantics are on an equal footing (Fig. 1). Each of the structures has its own combinatorial principles, but in addition they are linked by interface principles, which are what enable language to map between sound and meaning. A phrase or sentence is then a triple of well-formed phonological, syntactic, and conceptual structures, plus links among them established by the interface components.

One important reason to prefer this architecture to the standard one is that it allows us to situate the language faculty nicely within the rest of the mind. Consider again the question of how we talk about what we see. My hypothesis, inspired by Marr (1982), is that conceptual structure is connected by another interface to another form of mental representation that might be called *spatial structure*. This encodes one's understanding of the geometric configuration of the physical environment, and it is connected not just to

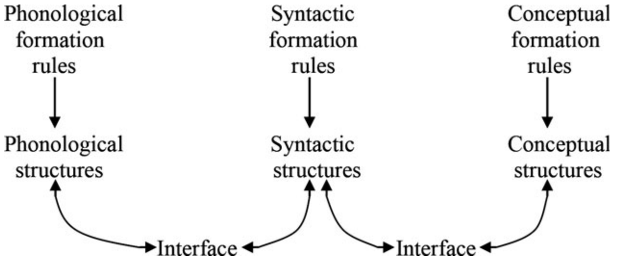


Fig. 1. The parallel architecture.

language, but also to multiple systems of perception—vision, haptic perception, and proprioception—as well as to the generation of action.

Fig. 2 shows how, on this conception, language fits into the overall ecology of the mind. The language faculty per se consists of phonology, syntax, and their interfaces to sensorimotor processes and to conceptual structure, which in turn encodes both linguistic meaning and nonlinguistic thought (Jackendoff, 1983, 2002, 2012).[4]

Now consider spatial structure. None of the representations that interface with it can be derived from it or from the others. Nor can any of these representations be derived from syntax. In particular, aside from phonology and syntax, all the components of Fig. 2 are evolutionarily older than language, so their structure cannot be dependent on that of language.[5] Within this overall layout, then, PA is perfectly natural: Its independent representations, coordinated by interfaces, are consonant with what has to be the case in the rest of the mind.

Notice also that the interface between conceptual structure and spatial structure is bidirectional. It is involved not only in talking about what we see, but in checking what we hear against visual input, and using what we see to disambiguate linguistic input, as in the visual world paradigm of Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995).

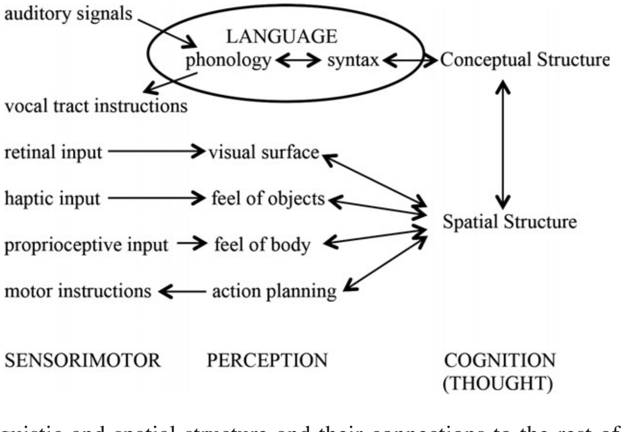


Fig. 2. Linguistic and spatial structure and their connections to the rest of the mind.

The layout in Fig. 2 makes it easy to plug in further domains of linguistic structure. For instance, orthography has its own internal principles such as the alphabet, plus an interface to the visual system, so one can see what's written, plus an interface to phonology or morphology that connects orthography to the linguistic system. Sign languages have, instead of vocal/auditory phonology, a phonology of structured gesture, with interfaces to vision and the motor system (Brentari, 1998; Perlmutter, 1992). And sociolinguistic variables such as formal versus casual speech register can be modeled as an interface between perceptions of social situation and the use of particular words or expressions (Jackendoff & Audring, in preparation). So, unlike standard generative grammar, the PA has the flexibility to link up new domains of cognition as the need arises. I take this to be an advantage of the theory.

4. Words and rules

A second major difference between the PA and standard generative grammar concerns the status of words and rules. As seen above, it is traditional to distinguish the lexicon, the list of words, from the grammar—the rules. This distinction is codified by Bloomfield (1933); Aronoff (1994) traces its roots back to the 17th century. From the earliest days of generative grammar, and especially from Chomsky (1965) onward, the lexicon-grammar distinction was taken for granted. The lexicon was take to encode static knowledge of idiosyncrasies, and the rules of grammar were viewed as general procedures or algorithms for constructing sentences step by step. Over the years, the main interest of generative linguistics has been in the step-by-step rules.

Of course, the rules of the grammar, like the words, do have to be encoded in memory in some form or another. To maintain the distinction between them, one might suppose that the lexicon and the grammar are stored differently in the brain, perhaps even in different places in the brain, as proposed by Ullman (2004).

I propose that there is a lot to be gained by abandoning the strict separation of lexicon and grammar and thinking of rules of grammar as *part* of the lexicon. The argument goes by a sort of slippery slope: We can find a succession of items stored in the lexicon that are progressively more and more rule-like. As we move through this progression, there is less and less reason to distinguish the items we find from things that everyone considers to be rules. Moreover, when the bottom of the slippery slope is reached, it turns out not to be as absurd as it might initially appear. (Versions of this argument have been made within HPSG [Pollard & Sag, 1994], Cognitive Grammar [Langacker, 1987], and Construction Grammar [Croft, 2001] as well as in the context of the PA [Jackendoff, 2002].)

Begin with words. In any mentalist theory, a word is an association of a piece of phonology, a piece of composite semantics, and some syntactic features, all stored in long-term memory. The word can be thought of as a little interface rule that links the three pieces of structure. Ex. (1) is the general form of the lexical entry for *cat*. (The notation *CAT* is simply a stand-in for the highly structured content of the concept "cat"; see Jackendoff [1983, 2002] for extensive discussion.)

(1) Phonology: /kæt/
Syntax: +N
Semantics: CAT

However, the lexicon also houses many deviations from this stereotypical case. For instance, words like those in (2) have phonology and semantics but no detectable syntax. They can occur alone as full utterances, and they occur in combination only in paratactic expressions like *Hello, Bill* and quotational contexts like "*Hello*," *she said*, where any expression at all, even non-English, can be inserted.

(2) *Phonology and semantics, no syntax:*
hello, ouch, yes, oops, dammit, upsey-daisy, allakazam, feh, uh-oh, shucks

English also has a few words such as the underlined words in (3) that have phonology and syntax but no semantics, and which function just as "grammatical glue."

(3) *Phonology and syntax, no semantics:*
a. <u>It</u>'s noisy in here.
b. I <u>did</u>n't see her.
c. a picture <u>of</u> Bill
d. I know <u>that</u> you came.

Nonsense words like those in (4) have legal phonology, but no syntax or semantics at all; their linguistic function is just to fill up metrical space in poetry or songs. If one happens to know the song, one recognizes these the same way one recognizes words.

(4) *Phonology, no syntax or semantics:*
hey diddle diddle, e-i-e-i-o, inka-dinka-doo, rickety-tickety-tin, bibbity-bobbity-boo

Hence, stored items can have different combinations of phonological, syntactic, and semantic features.

What else has to be stored in the lexicon? It is clearly necessary to store *idioms* and other fixed expressions in some form or another.[6] For instance, *kick the bucket* has phonological structure, plus a semantic structure approximately equivalent to *die*, plus the syntactic structure of a verb phrase. These have to be stored as a unit along the lines of (5), because the meaning is noncompositional.

(5) Phonology: /kɪk#ðə#bʌkət/
Syntax: $[_{VP}$ V $[_{NP}$ Det N$]]$
Semantics: DIE (X)

We know that (5) has the structure of a verb phrase, and is not simply an undigested string, because *kick* inflects just like an ordinary verb, for example, *he kick**ed** the bucket*, not **he kick-the-bucket**ed***. And we know it contains a noun phrase, because it admits at least limited adjectival modification: *kick the **old/proverbial** bucket.*[7] Moreover, idioms can have argument structure. For instance, those in (6) take a freely chosen direct object, just like ordinary transitive verbs, and this direct object goes exactly where a direct object *should* go—which happens to be in the middle of the idiom.

(6) take NP for granted
put NP on ice
give NP the once-over

A theory of knowledge of language cannot shrug idioms off as some kind of exceptional aberration. English has thousands of idioms, perhaps tens of thousands. Hence, we must give up the assumption that the lexicon is simply a list of words: It also contains items with internal syntactic structure.

Another class of idioms uses normal syntax, but to unusual semantic ends. Consider (7).

(7) Sam sneezed his way down the street.

Syntactically, (7) has a typical verb phrase. However:

- The verb *sneeze* doesn't usually take a direct object, but here we have this strange *his way* in the syntactic position normally occupied by the direct object.[8]
- *Sneeze* also doesn't normally occur with a path expression like *down the street* (**Sam sneezed down the street*).
- Even though there is no verb of motion, the meaning of the sentence in fact involves Sam *going* down the street.
- The verb *sneeze* describes the manner or concurrent action with which he goes down the street.In other words, (7) has a noncanonical mapping between syntax and semantics, marked by the phrase *his way* in object position. Any verb of the right semantic type is possible: You can eat your way across the country or knit your way through a conference. This sort of phenomenon might be called a "constructional idiom," and the case illustrated in (7) might be sketched informally as (8) (*pro* is a pronoun bound to the subject).

(8) *Way-construction*:
Syntax/Phonology: $[_{\mathrm{VP}}$ V pro's way PP$]$
Semantics: "NP goes PP while/by V-ing"

Two other mappings of the same sort, with different semantics, are found in *Fred drank the afternoon away* (= "Fred spent/wasted the afternoon drinking") and *Suzie sang her head off* (= "Suzie sang a lot/intensely"). Knowing these constructions is part of knowing English. For each one, a speaker has to learn and store its syntactic structure, how it links this syntax to meaning in other than the normal way, plus the phonology of the special words *way*, *away*, or *head off* that signal that something unusual is going on.

Other constructions, such as the three in (9), have no distinguishing phonological content, so they're not very wordlike at all.

(9)
a. *Sound+motion construction*
$[_{VP}$ V PP$]$
The bus rumbled around the corner. (= "The bus went around the corner, rumbling")
b. *Light verb construction*
$[_{VP}$ V NP NP$]$
Judy gave Henry a hug. (= "Judy hugged Henry")
c. *Casual paratactic conditional*:
[S, S]
You break my chair, I break your arm. (Edward Merrin, 12 June 2014)

The bus rumbled around the corner means that the bus *went* around the corner—even though there is no verb of motion—and that it was simultaneously making rumbling sounds. In the light verb construction, the direct object *a hug* provides the semantic content that is normally expressed by a verb, and the verb is more or less a dummy that provides argument positions and a host for tense. Finally, in the paratactic conditional, there is no *if-then*, but the conditional meaning is perfectly clear.

Again, a speaker has to learn and store these rule-like constructional idioms as associations between a syntactic complex and a semantic complex. But the basic formalism is the same as for words—except that the syntax is composite, not just a couple of features, and moreover idioms like (9) have no distinctive phonology. Constructions of this sort are a staple of Construction Grammar (Fillmore, Kay, & O'Connor, 1988; Goldberg, 1995; Hoffmann & Trousdale, 2013), and they also play an essential role in Simpler Syntax (Culicover & Jackendoff, 2005). The language turns out to be full of them, once one starts to look.

Mainstream generative grammar mostly ignores all these phenomena, and indeed they are not easy to account for in that framework. But within the PA and Construction Grammar, they emerge as variations on a theme: They are stored pieces of idiomatically linked syntactic and semantic structures, sometimes with accompanying phonology, sometimes not.

We now reach the crucial point in the argument. Using the very same formalism, a language's *phrase structure rules* can be stated as syntactic schemas, without any association to phonology or semantics. For example, (10) shows a partial rule for the transitive verb phrase in English. Tomasello (2003) calls this a "slot-and-frame schema," notating it as in (10a); Janet Fodor (1998) calls it a "treelet," in the alternative notation (10b). This is also the form of the basic combinatorial unit in Tree-Adjoining Grammar (Joshi & Schabes, 1997).

(10) a. 'slot-and-frame' notation: $[_{VP}$ V – (NP) …$]$

b. 'treelet' notation:
VP
V (NP)

In light of the phenomena in (2)–(9), (10) can be thought of as yet another type of lexical item. The lexicon contains items with just phonology like *fiddlededee*, items with just phonology and syntax like the *do* of do-support, items with just phonology and semantics like *hello*, and idioms and meaningful constructions with composite syntactic structure. (10) happens to be a lexical item with *only* syntactic structure—just one more logical possibility in the system. Its variables, like those in (6)–(9), license novel, freely combined phrases.

At this point we have slid to the bottom of the slippery slope from words to rules—and it actually makes some things work better, in particular the intermediate cases such as idioms and constructional idioms. The standard approach of generative grammar, in which words are static entities in the lexicon and rules are algorithmic processes, has to draw an unnatural line somewhere among these intermediate cases (and in practice, it has mostly ignored them). In contrast, this continuum is perfectly natural in PA's conception of rules. Words, rules, and everything in between are lexical items—pieces of stored structure. What makes an item word-like is that it is a grammatical word whose structure is fully specified. What makes an item rule-like is that some of its structure consists of variables, such as the V and NP in the VP schema.

However, (10) is not a rule of the standard sort—it is no longer one step in a sequence of procedures that create sentences out of smaller parts. Rather, it is a *schema*—a pattern that expresses a regularity in the language. This schema can be used in two ways. First, it can be used productively, as a piece of structure that can be "clipped" together with words, idioms, and other schemas to form novel utterances. Second, as we have seen, numerous idioms and constructions such as (5)–(9) have the form of transitive verb phrases. So another use of schema (10) is to motivate their structure—it says that the language includes a VP pattern to which lexical items can conform. This can be stated in terms of an inheritance hierarchy: the lexical item *kick the bucket* inherits its syntactic structure from the VP schema.

This second function, in which schemas are used to codify patterns, is especially important in morphology. Many morphologically complex words are much like idioms, in that they are semantically idiosyncratic but still conform to regular patterns. For instance, the meaning of the word *gorgeous* has nothing to do with the word *gorge*, and the word *impetuous* has no living relative along the lines of *impet*. Yet both of these conform to the common pattern of adjectives ending in *–ous*, like *joyous*, *glorious*, *incestuous*, and so on. The schema for adjectives with the *–ous* suffix expresses this pattern, and *gorgeous* and *impetuous*, along with *joyous* and so on, inherit structure from the schema.[9]

The treatment of phrase structure rules as schemas leads to the hypothesis that many—perhaps all—phenomena that have previously been called rules of grammar can now be reformulated as schemas stored in the lexicon. If so, the traditional issue of how complex expressions are constructed splits into two issues:

- How are words and schemas combined to create new utterances?
- How do schemas help organize or regulate the internal structure of words and phrases stored in the lexicon?

  Standard generative grammar tends to focus on Humboldt's "infinite use of finite means," and consequently attends primarily to the use of rules to construct novel utterances. But in a theory of mental representations like PA, in which lexical items have principled internal structure, and in which the rules are part of the lexicon, the latter issue assumes equal if not greater prominence.[10]

A nice consequence is that we can now think of learning a rule as adding a schema to the lexicon, based on generalizing over existing items. Such a schema will have constants where the related items are the same, and variables where they differ. The schema then forms an inheritance hierarchy with its instances. This is formally more natural than creating a rule in an entirely different format, as is necessary in a theory with the traditional declarative-procedural divide between words and rules.

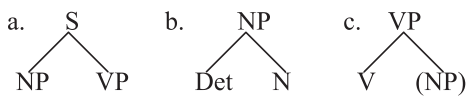
The hypothesis that rules are encoded as schemas invites a broader question. Recall the overarching hypothesis of Section 1: that general mental functions such as memory, attention, and learning are similar across many faculties. In light of this, we might be led to ask how closely long-term storage of language resembles long-term memory for other complex structures, such as memory for music and for structured actions. (Such an inquiry, of course, presumes that we have theories of these other faculties.) In this respect, the PA opens the prospect of a new sort of connection between linguistic theory and broader concerns of cognitive science. I take this to be an attractive feature of the approach, one that cannot be seriously addressed within the stance of standard linguistic theory, where the lexicon is an unstructured list.

5. The PA in language processing

Another payoff of treating words and rules as forming a continuum is in thinking about language processing. To begin with, remember what a word is: a long-term memory linking of pieces of phonological, syntactic, and conceptual structures, as illustrated in (1) above. It says that if one encounters *this* piece of phonology, say *cat*, it can be linked to *this* piece of syntax and *this* piece of conceptual structure.

This offers a simple way to think about lexical access (Jackendoff, 1987a, 2002, 2007b). In language comprehension, suppose phonetic perception creates some phonological string, say *cat*, in working memory. Then the processor sends a call to the lexicon: Does anyone sound like *this*? And various lexical entries are activated in parallel as competing candidates for what one is hearing. But once their phonology is activated, they also activate their syntax and semantics, which then come into play in attempting to construct an interpretation of the input. Similarly in production: one builds an intended message—a conceptual structure—in working memory, and the processor calls the lexicon: Does anyone there mean *this*?—and the conceptual structures of various entries are activated as candidates. Once these meanings are activated, they activate their syntax and phonology, and that gives the processor something to use in constructing the form of what one is going to say. This corresponds pretty well to the way many psycholinguists and computational linguists have thought about lexical access for a long time (e.g., Jurafsky, 1996; Levelt, 1989; Marslen-Wilson, 1987; Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979).[11]

This approach extends easily to syntactic parsing, along lines similar to Tree-Adjoining Grammar (Schabes, Abeillé, & Joshi, 1988) and Construction Grammar (Jurafsky, 1996).[12] As in these frameworks, PA describes tree structures by listing available treelets like these:

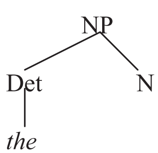
(11)


A tree is built by "clipping together" (more technically, unifying) these treelets at nodes they share, working from the bottom up, from the top down, or from anywhere in the middle, as long as the resulting tree ends up with S at the top and terminal symbols at the bottom. No order for building trees is logically prior to any other; the process of combining treelets is compatible with left-to-right, top-down, or bottom-up assembly, in particular with left-corner parsing (e.g., Hale, 2011).

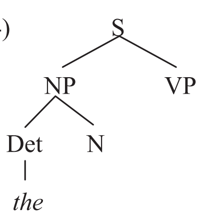
Suppose an utterance begins with the phonology *the*. This activates the lexical item *the*, which puts a Determiner node into play in syntax, yielding subtree (12).

(12)


Det calls to the lexicon, seeking structures it can fit into. Since NP contains a Det, the NP treelet (11b) activates and clips on to produce tree (13), predicting a noun to come.

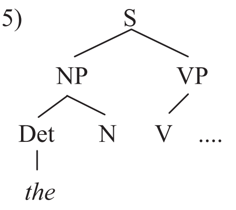
(13)


The NP then calls to the lexicon, seeking structures *it* can fit into. Since in this case the NP is sentence-initial, one of the treelets that activates is (11a), which gets clipped on, yielding (14), which predicts a VP to come.

(14)


Then the VP calls to the lexicon, and (11c) is activated and clipped on, predicting a verb to come, as in (15).

(15)



In turn, subsequent words and syntactic nodes can be clipped into the tree, fulfilling these predictions. In short, parsing with lexically stored treelets offers a simple account of anticipatory parsing along the lines of, for example, Levy (2008) and Hale (2011), setting up grammatical expectations or predictions on the basis of an initial word (a procedure that Phillips and Lau [2004] find "somewhat mysterious" in the context of Chomsky's (1995) Minimalist Program).

Bear in mind, however, that the predictions in (12)–(15) are not necessarily correct. The sentence could turn out to be something like (16), in which *the* is not a determiner at all.

(16) The more I read, the less I understand.

Here, *the* is a mark of a constructional idiom called the comparative correlative or CC (Culicover & Jackendoff, 2005; Fillmore et al., 1988). Like the constructional idioms mentioned earlier, the CC is stored as a schema in the lexicon. How does this schema come into play in parsing (16)? The simplest hypothesis is that *the* recruits the CC schema right along with the NP schema (11b), though at a much lower level of activation, thanks to its much lower frequency. But if the next word happens to be *more*, the CC is far better confirmed by the input and therefore increases in activation. On the other hand, the CC had better not entirely extinguish the NP schema, since the sentence *could* turn out to be either something like (17a,b), with an ordinary noun phrase, or (18a,b), instances of CC.

(17) a. The more attractive <u>theory</u> doesn't always win.
b. The more thorough but less immediately attractive <u>theory</u> often doesn't win.

(18) a. The more attractive <u>a</u> theory is, the better chance it has of catching on.
b. The more thorough but less immediately attractive <u>a</u> theory is, the worse its chances are.

Hence, the NP and the CC constructions have to be in competition for some distance into the utterance—as much as seven words in (17b) and (18b)—with no subjective sensation of strain or garden-pathing (although I know of no experimental work testing this). The uncertainty is eventually resolved by the appearance of a noun, underlined in (17), or a determiner, underlined in (18).

More generally, the PA predicts that everything that happens to words in processing also ought to happen to schemas such as syntactic treelets. For instance, in lexical access, multiple *words* are under consideration in parallel (Swinney, 1979; Tanenhaus et al., 1979). Similarly, it is generally taken to be established that alternative syntactic structures compete in working memory, trying to extinguish each other, so that in production one ends up *saying* one particular thing (barring speech errors), and in comprehension one (usually) ends up *hearing* one particular thing (Ferreira, 2003; Gibson & Pearlmutter, 2000; Hale, 2011; Jurafsky, 1996; Levy, 2008). Within PA, this is precisely what is to be expected, since words and treelets are both in the lexicon and play the same roles in processing.[13]

Similarly, in discussing the competition between the NP and CC schemas, I have assumed that the NP schema is initially more highly activated than the CC schema because of its higher frequency. Of course, it is a commonplace that more frequent *words* come to be more highly activated; but in fact this is true of syntactic structures as well (e.g., Bresnan & Ford, 2010). In the PA, where syntactic treelets and constructional idioms are part of the lexicon right along with words, this is a perfectly natural conclusion: Syntactic schemas should have activation levels just like words, and more frequent schemas should be easier to access than less frequent ones. But in a standard generative grammar, it would be difficult (though perhaps not impossible) to make this connection, since trees built by syntactic rules are entirely different from words.

Consider also syntactic priming, where, for instance, hearing a ditransitive sentence like (19a) increases the probability of producing another ditransitive like (19b) rather than a prepositional dative like (19c).

(19) a. He gave Bill the ball.
b. She sent Harry the package.
c. She sent the package to Harry.

Most psycholinguists now take syntactic priming for granted. But Bock's original work on syntactic priming (Bock, 1995; Bock & Loebell, 1990) observed that it was somewhat unexpected. Words prime, presumably from a transient increase in activation. But why, Bock asked, should there be priming on the application of phrase structure rules? In the prevailing framework, which she took for granted, words and rules were entirely different animals. However, from the perspective of PA, the ditransitive construction and the prepositional dative configuration are alternative schemas in the lexicon, and so they should be *expected* to prime, just like words.

More generally, the PA invites a direct connection between the structures posited by the theory and their behavior in performance. For instance, the theory identifies situations where syntax does not map to semantics in canonical fashion, such as the constructional idioms mentioned in section 4. Common sense might suggest that noncanonical mappings should lead to complexity in processing, and that we ought to be able to detect this complexity experimentally.

Here I will mention just one such situation, the light verb construction (9b). A light verb sentence like *Judy gave Henry a hug* is pretty much synonymous with *Judy hugged Henry*. The noun *hug* signifies an action, which would normally be denoted by a verb, whereas the actual verb functions as sort of a placeholder. Because of this mismatch between syntax and semantics, the PA predicts that processing ought to require some extra work, namely identifying *Judy* and *Henry* as Agent and Patient of the predicate expressed by the noun *hug*.

Such extra processing has been detected in a variety of experiments, summarized in Wittenberg et al. (2014).

- Using a cross-modal lexical decision task, Piñango, Mack, and Jackendoff (2006) find extra processing load 300 milliseconds after the presentation of the object of light verb constructions—in a time frame that is believed to be associated with semantic rather than syntactic processing.
- Wittenberg and Piñango (2011) replicate this finding in German, in a paradigm where the critical point is the clause-final verb. They find the same strain on processing, despite the fact that the verbs in question are more frequent as light verbs than in nonlight constructions.
- Wittenberg, Paczynski, Wiese, Jackendoff, and Kuperberg (2014), recording ERPs in German light verb constructions, find a widely distributed, frontally focused sustained negativity, 500–900 ms after presentation of the critical word. This response is quite different from the response to anomalous and ungrammatical sentences, and they attribute it to the extra processing required to align the thematic roles.
- Conventional generative grammar, if it addresses light verb constructions at all (e.g., Hale & Keyser, 2002), treats them as having a different syntax from an ordinary verb phrase like *give someone a book*. Wittenberg and Snedeker (2014) test this prediction by running a syntactic priming experiment. They show that double object nonlight verb constructions like *give someone a book* can be syntactically primed by double object light verb constructions like *give someone a hug* just as well as by another nonlight. This is evidence that the syntactic structure is the same in the two, as the PA claims.

  In short, then, the theoretical analysis of the PA has brought to attention a phenomenon that previously had received little experimental investigation, and the results bear out the PA's account over other theoretical alternatives. More generally, this section has shown that a theory of representation *can* interact meaningfully with issues in language processing, and that the road goes in both directions.

6. Languages without much syntax

An entirely different issue that can be approached within the PA's view of linguistic representation is languages that don't have much syntax.

To approach the problem, consider again the PA's interfaces in Fig. 1: an interface between phonology and syntax and one between syntax and semantics. This leaves open the possibility of another interface, a direct route between phonology and semantics, bypassing syntax, as in Fig. 3. Such an interface would be impossible in a traditional generative grammar, in which everything comes from syntax. But it is a natural possibility in the PA.

We have already seen this kind of linking: words like *hello* and *ouch* that have phonology and meaning, but no syntactic properties. This is also an attractive way of thinking about the effect of contrastive stress on topic and focus, as in (20), where only prosody signals the difference in meaning.

(20) a. Joe likes ICE cream.
b. Joe LIKES ice cream.
c. JOE likes ice cream.

But there is a deeper question: What would happen if the syntactic component were erased from the PA, so there would be *only* a connection between phonology and semantics, like Fig. 4? What would a language be like that was structured like this (Jackendoff & Wittenberg, 2014)?

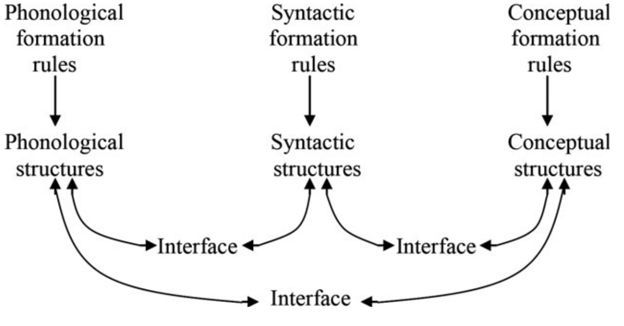


Fig. 3. The parallel architecture, with a phonology-semantics interface.

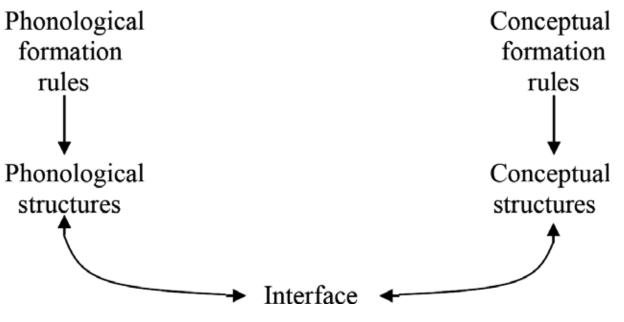


Fig. 4. The parallel architecture without syntax.

- A language like this could still have words that correlated pieces of sound with pieces of meaning. But the words couldn't have parts of speech: They could be distinguished only by whether they denote objects, actions, or properties, that is, by semantic features.
- A language like this could still have sequences of words in linear order, and it could have prosodic structure, that is, constituents delimited by intonation contours. But it couldn't have noun phrases and verb phrases, and subjects and objects, and clauses.
- This means it couldn't have *subordinate* clauses. So *syntactic recursion* would be absent (though there still might be *semantic* recursion, just not expressed that way syntactically).
- A language like this still might be able to assign semantic roles in terms of the linear order of words, through principles such as "Agent precedes Patient," "Focus Last," or "Modifier follows Modified."
- Words might have interior morphological structure, but this would have to be based on semantic rather than syntactic grounds.

  I'll call a grammar with these properties a *linear grammar*.

Linear grammars are not just a theoretical fantasy. There are actual languages that have these sorts of properties.

First, David Gil has written extensively about Riau Indonesian, a language spoken by several million people in an urban environment (e.g., Gil, 2005, 2009, 2014). He presents a persuasive case that this language has no parts of speech, almost no function words, virtually no morphology, and—on my analysis in terms of the PA–virtually no syntactic structure and definitely no recursion. The word order is determined directly by semantics, and interpretation typically requires a great deal of pragmatics and/or contextual support.[14]

Second, my reading of the literature on pidgins (e.g., Bickerton, 1981; Gil, 2014; Givón, 2009) suggests that they are structured this way as well, with if anything a greater degree of variability and necessity for pragmatic support. But they still have phonology and semantics.

Third, Klein and Perdue (1997) find that late second-language learners in naturalistic environments invariably pass through a stage they call the Basic Variety, which also corresponds rather closely with the characteristics of linear grammar.

More examples come from emerging sign languages such as Nicaraguan Sign Language (Kegl, Senghas, & Coppola, 1999), Abu-Sayyid Bedouin Sign Language (Aronoff, Meir, Padden, & Sandler, 2008; Sandler, Meir, Padden, & Aronoff, 2005), and a village sign language from Turkey, Central Taurus Sign Language, that came to light in 2012 (Caselli, Ergin, Jackendoff, & Cohen-Goldberg, 2014; Ergin et al., 2014). Especially at the earlier stages of these languages, there is no evidence for syntactic structure. Again there are semantically based principles of word order, no function words, little or no evidence of embedding or even constituent structure beyond that provided by prosody, and a great reliance on pragmatics (though semantically based morphology does develop relatively early).

In addition, looking from this perspective, one can find subsystems of English that have these sorts of properties. The clearest case is compounding (Jackendoff, 2010a), which simply jams two words together, and stipulates that normally the second one in linear order is the semantic head, for instance that *beef stew* is a kind of stew, but *stew beef* is a kind of beef. Beyond that, the proper interpretation is largely a matter of pragmatics, as can be seen from these well-known types of examples, where the extra semantic material is underlined.

(21) beef stew = "stew made of beef"
attack helicopter = "helicopter that attacks"
mosquito netting = "netting to protect from mosquitoes"
dog house = "house for a dog to live in"
bike helmet = "helmet to wear while riding a bike"
paper clip = "clip to hold papers together"
oil stove = "stove that runs on oil"

It is conceivable that principles of linear grammar are at work in normal sentence processing as well. They seem to correspond rather nicely to what people have called "perceptual strategies" (Townsend & Bever, 2001) or "good-enough parsing" (Ferreira, 2003) that show up in garden-pathing and other measures of parsing difficulty. These principles are not just general-purpose cognition, as they have often been characterized. They are still mappings of phonology to semantics—that is, language—just not via syntax. In the simplest cases, they reinforce syntactic analysis. But in such cases as reversible passives and object relatives, where the agent doesn't come first, they conflict with syntactic analysis, and this might help account for the difficulty.

These considerations suggest that there is no bright line that separates "true" languages from non-languages. It is just a matter of the extent to which a system goes beyond the resources offered by linear grammar. Early emerging sign languages use fewer resources than creoles and Riau Indonesian, which in turn use fewer resources than English, Russian, and Mohawk. And even in the latter languages, as we have seen, the resources of linear grammar do not go away.[15]

There is a lesson here. Some schools of thought in linguistics and cognitive science question the need for syntax in natural languages (e.g., Rumelhart & McClelland, 1986; Frank, Bod, & Christiansen, 2012; and many others). The investigation of linear grammar gives us a benchmark for what languages without syntax can be like: no function words, no inflectional morphology, no passive, no subordinate clauses, and a huge reliance on pragmatics. To me at least, it shows how essential syntax is to expressiveness in developed languages, while at the same time it acknowledges that there are circumstances (and even whole languages!) where speakers make do without it. This raises all sorts of interesting issues about language acquisition, innateness, poverty of the stimulus, critical periods, and the evolution of the language faculty that unfortunately I cannot even begin to discuss here.

7. Closing thoughts

To sum up, the PA grounds linguistic theory in the questions: What do speakers store in memory, and in what form? And how are these stored pieces assembled into utterances? The answer proves to be rather surprising from the point of view of mainstream linguistics: Words, schemas, and all sorts of intermediate phenomena are stored in a common format, namely as pieces of phonological, syntactic, and/or conceptual structure, linked by interface principles. The pieces are assembled not through a step-by-step bottom-up or top-down derivation, but by an operation of opportunistically "clipping" pieces together, which can be directly instantiated in a rather natural model of processing.

Many of the ideas brought up here are of course found in other approaches to linguistic theory and language processing. I have tried to find the insights in each approach and put them to work. What is new here, I think, is the synthesis of these insights into a coherent story—a story that offers a chance for better integration, both within linguistic theory itself and between linguistic theory, psycholinguistics, and cognitive science at large.

At more of a meta-level, I have tried to demonstrate what it can be like to engage detailed theoretical issues about mental representations, and especially what it can be like to uncover assumptions in the field that are so deeply buried that they are not even noticed. My measure of success is that, when you get it right, all kinds of interesting prospects for research open up, both theoretical and experimental. In this sense, I hope to have shown that the PA in fact does get it right.

So I want to conclude by going back to my challenge to the field of cognitive science as a whole: to develop comparable theories of mental representations in other domains of human cognition. I hope the discussion here might encourage some readers to take questions of representation more seriously, in their own domains of expertise, and to recognize how theoretical work of this sort can stimulate more comprehensive and ecologically plausible experimental and modeling research.

Notes

1. There may be some disagreement as to whether Marr intended this characterization to belong to the computational or the algorithmic level of description. But there is no question that he saw it as of the highest importance.
2. The biggest dispute here over the years has been what aspects of these procedures are domain-general and what aspects, if any, are domain-specific. Unfortunately, there has been a lot of polarization on this issue (e.g., Elman et al., 1996; Hauser, Chomsky, & Fitch, 2002; Lakoff, 1987; Tomasello, 1995). I think it's an empirical question, not a matter of ideology (Jackendoff, 2002, 2011; Pinker & Jackendoff, 2005).

3. Exceptions are Latent Semantic Analysis (Landauer & Dumais, 1997), FrameNet (https://framenet.icsi.berkeley.edu/fndrupal/), and Natural Semantic Metalanguage (Wierzbicka, 1996), which all propose to explicate meaning purely in terms of natural language words.
4. I consider conceptual structure to be outside the language faculty per se because, like spatial structure, it is demonstrably present in nonlinguistic organisms such as babies and apes. Nevertheless, it is doubtless hugely enriched in adult humans through its connection to language. In any event, nothing much hangs on where one draws the boundaries of the ellipse in Fig. 2.
5. It also bears mention that the representations to the left of spatial structure in Fig. 2 are potential vehicles for various modalities of grounded cognition (Barsalou, 2008), whereas spatial structure and conceptual structure are amodal, in the sense of not being tied to a particular modality. Jackendoff (1987b, 1996) discusses how the work of amodal cognition is divided up between the quasi-geometric format of spatial structure and the quasi-algebraic format of conceptual structure. These two are perhaps the present framework's counterpart of Paivio's (1971) Dual Routes, but more abstract in conception.
6. Di Sciullo and Williams (1987) use the term *listeme* for all items listed in the lexicon, including idioms and other multiword expressions; Pinker (1999: 24) proposes to use the term *word* in a similar sense. I prefer reserving *word* for single words, and I will use *lexical item* or *lexical entry* for Di Sciullo and Williams's listemes.
7. Does the entry for *kick the bucket* have to specify its phonology? At the very least, it has to include pointers to the phonology of the individual items *kick*, *the*, and *bucket*. However, Jackendoff (1975) and Jackendoff and Audring (unpublished data) show that in general such maximally impoverished entries are insufficient, and that lexical entries, even when they are built from other items, have to include something close to a full specification of phonology and meaning.
8. Even though *his* way does not have a thematic role assigned by *sneeze*, we know it is in direct object position, because a transitive verb cannot appear in the construction with an object of its own.

   (i) *Bill drank bourbon his way into a stupor. *but* Bill drank his way into a stupor.

   It is as if *his way* has "hijacked" syntactic object position for its own purposes.
9. Pinker (1999) equivocates between treating rules as combinatorial procedures or as treelets (though my impression is that the former is predominant). Jackendoff (2013; Jackendoff & Audring, unpublished data) reanalyzes Pinker's distinction between productive and unproductive patterns in terms of a distinction between schemas (i.e., treelets) with productive versus nonproductive variables.
10. A potential obstacle to treating rules as schemas in the lexicon is that this requires all of syntax to be reformulated in terms of schemas rather than in terms of the step-by-step derivations of the standard generative approach. In particular, it is necessary to eliminate the very seductive procedural notion of movement, such as the idea of moving the auxiliary in front of the subject in questions, moving wh-words to the front in questions and relative clauses, and moving the object to sub-

ject position in passives. Fortunately, HPSG, LFG (Bresnan, 1982, 2001), Construction Grammar, Simpler Syntax, and other frameworks all have been working this out for some decades, with a good deal of success; I wouldn't dream of trying to summarize the arguments here.

11. Jackendoff (2002, sections 3.5 and 7.3; 2007b) observes that activation cannot simply amount to activation of nodes in a semantic or connectionist-style network. In particular, structures can be built that contain two or more instances of the same lexical item. For instance, *the black cat hissed at the white cat* contains two occurrences of *the*, two occurrences of *cat*, two nouns, and two NPs. These occurrences cannot be kept distinct if each consists simply of an activation of the lexical entry in long-term memory. Rather, there has to be a separate facility in the brain corresponding to a more traditional working memory, where pieces of stored material can be combined into larger structures. Long-term memory in this sense might correspond to Hagoort's (2013) notion of memory; working memory might correspond to his unification component. I want to emphasize that this need for a separate working memory is not an artifact of PA; it is a constraint on any theory of language processing—and on theories of other faculties as well. Consider, for instance, what it takes for a scene analysis to identify two identical forks in a visual array and encode them *as forks*.
12. Hale's (2011) "Rational Parser" also assembles treelets, although he considers them to have been built in the course of parsing by procedural phrase structure rules. The account could easily be reframed in the present terms.
13. MacDonald, Pearlmutter, and Seidenberg (1994) motivate this parallelism by treating all syntactic structure as projected from words, so that *cat*, for instance, contains within its lexical entry all the syntactic possibilities for noun phrases such as determiners and prenominal adjectives. I find this hypothesis unlikely. It leads to the claim that the lexical entry for every noun in the language encodes the possibility of determiners, adjectives, postnominal prepositional phrases, and relative clauses; and the lexical entry for every verb encodes the possibility of auxiliaries, direct objects, adverbs in many positions, plus time, place, instrument, manner, and means expressions. In other words, MacDonald et al. have offloaded the generalizations captured by phrase structure rules onto thousands of individual words. The PA account preserves MacDonald et al.'s lexicalism but by another route: downloading phrase structure rules *as such* into the lexicon.
14. I would be remiss not also to mention the notorious case of Pirahã, which, according to Everett (2005, 2009), lacks syntactic recursion. Nevins, Pesetsky, and Rodrigues (2009) show that under their interpretation of the data, this language *might* have subordinate clauses; but they do not show that a clause may be embedded within a clause that is itself embedded—the true proof of recursion. In fact, R. Futrell, L. Stearns, D. Everett, S. Piantadosi, and E. Gibson "A corpus investigation of syntactic embedding in Pirahã", unpublished) find no evidence of embedding in a sizable corpus of transcriptions of Pirahã speech. The language

does, however, have a great deal of morphology and perhaps some limited phrase structure, so it is not quite a pure linear grammar.

15. At one presentation of this material, a prominent creolist was moved to remark that "pidgins aren't natural languages." But if they're not natural languages, what are they? They're still mappings between sound and meaning, just less structured.

References

Aronoff, M. (1994). *Morphology by itself*. Cambridge MA: MIT Press.

Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2008). The roots of linguistic organization in a new language. In D. Bickerton & M. Arbib (Eds.), *The roots of linguistic organization in a new language*, Special Issue of *Interaction Studies* (pp. 133–149).

Barsalou, L. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–147.

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Bock, K. (1995). Sentence production: From mind to mouth. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition: Vol. XI. Speech, language, and communication* (pp. 181–216). Orlando, FL: Academic Press.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *The Linguistic Review*, *23*, (Special Issue on Exemplar-Based Models of Language), 291–320.

Booij, G. (2010). *Construction morphology*. Oxford, UK: Oxford University Press.

Brentari, D. (1998). *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Bresnan, J. (Ed.) (1982). *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.

Bresnan, J. (2001). *Lexical-functional syntax*. Malden, MA: Blackwell.

Bresnan, J., & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language*, *86*, 168–213.

Bybee, J., & McClelland, J. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, *22*, 381–410.

Caselli, N., Ergin, R., Jackendoff, R., & Cohen-Goldberg, A. (2014). The emergence of phonological structure in Central Taurus Sign Language. Conference talk: From Sound to Gesture, Padua, Italy.

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

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford, UK: Oxford University Press.

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.

Croft, W. (2001). *Radical construction grammar*. Oxford, UK: Oxford University Press.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. Oxford, UK: Oxford University Press.

Dennett, D. C. (1978). *Brainstorms*. Cambridge, MA: Bradford Books/MIT Press.

Di Sciullo, A., & Williams, E. (1987). *On the definition of word*. Cambridge, MA: MIT Press.

Edelman, S., & Christiansen, M. H. (2003). How seriously should we take Minimalist syntax? *Trends in Cognitive Science*, *7*, 60–61.

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple-view representation of 3D objects. *Biological Cybernetics*, *64*, 209–219.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Ergin, R., Caselli, N., Padden, C., Ilkbasaran, D., Meir, I., Cohen-Goldberg, A., & Jackendoff, R. (2014). Argument structure in a newly documented village sign language. Sixth Conference of the International Society for Gestural Studies, San Diego.

Everett, D. (2005). Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology*, *46*, 621–646.

Everett, D. (2009). Pirahã culture and grammar: A response to some criticisms. *Language*, *85*, 405–442.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F. (2005). Psycholinguistics, formal grammar, and cognitive science. *The Linguistic Review*, *22*, 365–380.

Fillmore, C., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, *64*, 501–538.

Fodor, J. A. (1981). *Representations*. Cambridge, MA: Bradford Books/MIT Press.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research*, *27*, 339–374.

Frank, S. L., Bod, R., & Christiansen, M. (2012). How hierarchical is language use? *Proceedings of the Royal Society B: Biological Sciences*, *279*, 4522–4531.

Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain*. Malden, MA: Wiley-Blackwell.

Gibson, E., & Pearlmutter, N. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research*, *29*, 231–240.

Gil, D. (2005). Word order without syntactic categories: How Riau Indonesian does it. In A. Carnie, H. Harley, & S. A. Dooley (Eds.), *Verb first: On the syntax of verb-initial languages* (pp. 243–263). Amsterdam: John Benjamins.

Gil, D. (2009). How much grammar does it take to sail a boat? In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable*. Oxford, UK: Oxford University Press.

Gil, D. (2014). Sign languages, creoles, and the development of predication. In F. Newmeyer & L. Preston (Eds.), *Measuring grammatical complexity* (pp. 37–64). Oxford, UK: Oxford University Press.

Givón, T. (2009). *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution*. Amsterdam: John Benjamins.

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

Goldberg, A. (2006). *Constructions at work*. Oxford, UK: Oxford University Press.

Goldsmith, J. (1979). *Autosegmental phonology*. New York: Garland Press.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*, 416. doi:10.3389/fpsyg.2013.00416

Hale, J. (2011). What a rational parser would do. *Cognitive Science*, *35*, 609–642.

Hale, K., & Keyser, S. J. (2002). *Prolegomena to a theory of argument structure*. Cambridge, MA: MIT Press.

Haugeland, J. (Ed.) (1981). *Mind design*. Cambridge MA: Bradford Books/MIT Press.

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

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford, UK: Blackwell.

Hoffmann, T., & Trousdale, G. (2013). *The Oxford handbook of construction grammar*. Oxford, UK: Oxford University Press.

Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, *51*, 639–671. Reprinted with commentary in Jackendoff (2010b), 35–84.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, R. (1987a). *Consciousness and the computational mind*. Cambridge, MA: Bradford Books/MIT Press.

Jackendoff, R. (1987b). On beyond zebra: The relation of linguistic and visual information. *Cognition*, *26*, 89–114. Reprinted with commentary in Jackendoff (2010b), 88–111.

Jackendoff, R. (1990). Semantic Structures. Cambridge, MA: MIT Press.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 1–30). Cambridge, MA: MIT Press. Reprinted in Jackendoff (2010b), 112–137.

Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.

Jackendoff, R. (2002). *Foundations of language*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2007a). *Language, consciousness, culture*. Cambridge MA: MIT Press.

Jackendoff, R. (2007b). A Parallel Architecture perspective on language processing. *Brain Research*, *1146*, 2–22.

Jackendoff, R. (2009). Parallels and non-parallels between language and music. *Music Perception*, *26*, 195–204.

Jackendoff, R. (2010a). The ecology of English noun-noun compounds. In Jackendoff (2010b), 413–451.

Jackendoff, R. (2010b). Meaning and the Lexicon. Oxford: Oxford University Press.

Jackendoff, R. (2011). What is the human language capacity? Two views. *Language*, *87*, 586–624.

Jackendoff, R. (2012). *A user's guide to thought and meaning*. Oxford, UK: Oxford University Press.

Jackendoff, R. (2013). Constructions in the parallel architecture. In Hoffman and Trousdale, 2013, 70–92.

Jackendoff, R., & Audring, J. (In preparation). *The texture of the mental lexicon: Morphology in the parallel architecture.*

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*, 33–72.

Jackendoff, R., & Wittenberg, E. (2014). What you can say without syntax: A hierarchy of grammatical complexity. In F. Newmeyer & L. Preston (Eds.), *Measuring linguistic complexity* (pp. 65–82). Oxford, UK: Oxford University Press.

Joshi, A., & Schabes, Y. (1997). Tree-adjoining grammars. In G. Rozenberg and A. Salomaa (Eds.), *Handbook of formal languages* (pp. 69–123). Berlin: Springer-Verlag.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194.

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language creation and language change* (pp. 179–238). Cambridge, MA: MIT Press.

Klein, W., & Perdue, C. (1997). The basic variety (Or: Couldn't natural languages be much simpler?). *Second Language Research*, *13*, 301–347.

Kuperberg, G., & Jaeger, F. (in press). What do we mean by prediction in language comprehension?. Cognition and Neuroscience.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*, 217–238.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Langacker, R. (1987). *Foundations of cognitive grammar* (vol. 1). Stanford, CA: Stanford University Press.

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

Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, *8*, 249–336.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Macnamara, J. (1978). How do we talk about what we see? Manuscript, McGill University.

Margolis, E., & Laurence, S. (Eds.) (1999). *Concepts: Core readings*. Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102.

Nevins, A., Pesetsky, D., & Rodrigues, C. (2009). Pirahã exceptionality: A reassessment. *Language*, *85*, 355–404.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Patel, A. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Perlmutter, D. (1992). Sonority and syllable structure in American Sign Language. *Linguistic Inquiry*, *23*, 407–442.

Phillips, C., & Lau, E. (2004). Foundational issues (review article on Jackendoff 2002). *Journal of Linguistics*, *40*, 1–21.

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory phonology VII*, 101–139.

Piñango, M., Mack, J., & Jackendoff, R. (2006). Semantic combinatorial processes in argument structure: Evidence from light verbs. *Proceedings of Berkeley Linguistics Society*, *32*. Available at http://linguistics.berkeley.edu/bls/past_meetings/bls32/abstracts/bls32_pinangoetalLA.pdf. Accessed May 31, 2015.

Pinker, S. (1999). *Words and rules*. New York: Basic Books.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*, 201–236.

Poeppel, D., & Embick, D. (2005). Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 103–120). Hillsdale, NJ: Lawrence Erlbaum.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*, 263–266.

Pollard, C., & Sag, I. (1994). *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishers.

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience supplement*, *3*, 1199–1204.

Rumelhart, D., & McClelland, J. (1986). On learning the past tense of English verbs. In J. McClelland, D. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing* (vol. ii, pp. 216–271). Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar in a new sign language. *Proceedings of the National Academy of Sciences*, *102*(7), 2661–2665.

Schabes, Y., Abeillé, A., & Joshi, A. K. (1988). Parsing strategies with "lexicalized" grammars: Application to tree adjoining grammars. Technical Reports (CIS), Paper 691. Available at http://repository.upenn.edu/cis_reports/691. Accessed May 31, 2015.

Schank, R. (1973). Identification of conceptualizations underlying natural language. In R. Schank & K. Colby (Eds.), *Computer models of thought and language* (pp. 187–248). San Francisco: Freeman.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Solan, Z., Horn, D., Ruppin, E., & Edelman, S. (2005). Unsupervised learning of natural languages. *Proceedings of the National Academy of Sciences*, *102*(33), 11629–11634.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.

Tanenhaus, M., Leiman, J. M., & Seidenberg, M. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*, 427–440.

Tanenhaus, M., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects." *Psychonomic Bulletin & Review*, *2*, 55–82.

Tomasello, M. (1995). Language is not an instinct (Review of Steven Pinker, *The Language Instinct*). *Cognitive Development*, *10*, 131–156.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Townsend, D., & Bever, T. (2001). *Sentence comprehension*. Cambridge MA: MIT Press.

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

Ullman, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, *67*, 21–44.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford, UK: Oxford University Press.

Wilson, R. A., & Foglia, L. (2011). Embodied cognition. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition). Available at http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/. Accessed May 31, 2015.

Wittenberg, E., Jackendoff, R., Kuperberg, G., Paczynski, M., Snedeker, J., & Wiese, H. (2014). The processing and representation of light verb constructions. In A. Bachrach, I. Roy, & L. Stockall (Eds.), *Structuring the argument. Multidisciplinary research on verb argument structure*. Philadelphia: John Benjamins (pp. 61–80).

Wittenberg, E., Paczynski, M., Wiese, H., Jackendoff, R., & Kuperberg, G. (2014). The difference between "giving a rose" and "giving a kiss": Sustained neural activity to the light verb construction. *Journal of Memory and Language*, *73C*, 31–42.

Wittenberg, E., & Piñango, M. (2011). Processing light verb constructions. *The Mental Lexicon*, *6*, 393–413.

Wittenberg, E., & Snedeker, J. (2014). It takes two to kiss, but does it take three to give a kiss? Categorization based on thematic roles. *Language, Cognition and Neuroscience*, *29*(5), 635–641.