View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by PhilPapers

Croatian Journal of Philosophy
Vol. XI, No. 31, 2011

Psychological and Computational Models of Language Comprehension: In Defense of the Psychological Reality of Syntax

DAVID PEREPLYOTCHIK
Department of Philosophy,
Baruch College, CUNY

In this paper, I argue for a modified version of what Devitt (2006) calls the Representational Thesis (RT). According to RT, syntactic rules or principles are psychologically real, in the sense that they are represented in the mind/brain of every linguistically competent speaker/hearer. I present a range of behavioral and neurophysiological evidence for the claim that the human sentence processing mechanism constructs mental representations of the syntactic properties of linguistic stimuli. I then survey a range of psychologically plausible computational models of comprehension and show that they are all committed to RT. I go on to sketch a framework for thinking about the nature of the representations involved in sentence processing. My claim is that these are best characterized not as propositional attitudes but, rather, as subpersonal states. Moreover, the representational properties of these states are determined by their functional role, not solely by their causal or nomological relations to mind-independent objects and properties. Finally, I distinguish between explicit and implicit representations and argue, contra Devitt (2006), that the latter can be drawn on "as data" by the algorithms that constitute our sentence processing routines. I conclude that Devitt's skepticism concerning the psychological reality of grammars cannot be sustained.

Key words: psychological reality of language, mental representation, sentence processing, computational parsing models, personal and subpersonal states

§1. *Introduction*

Michael Devitt's book, *Ignorance of Language*, marked a new chapter in the debate concerning the psychological reality of language. Like

language itself, the debate is multi-faceted, bringing into the fold core issues in epistemology, metaphysics, and psychology. Devitt's discussion is comprehensive; he covers much ground and makes bold claims about the epistemic status of linguistic intuitions, the ontology of linguistic theory, and the proper characterization of innateness, *inter alia*. It's hardly surprising, then, that Devitt's critics have challenged his views from a wide variety of angles. For instance, Culbertson and Gross (2009) mount a critique of Devitt's views on intuitions, and Rey (2006, 2008) challenges Devitt's quasi-nominalism regarding linguistic entities like words and phrases. Similarly, Pietroski (2008) argues that Devitt's position is at odds with well-founded claims about language acquisition and Cain (2010) sheds doubt on Devitt's characterization of public-language conventions.

Joining the fray, I tackle what I take to be the main thesis of *Ignorance of Language*, viz., that the rules or principles of grammar are not mentally represented. In contrast to the commentators mentioned above, my approach will center primarily on the results emerging from psycholinguistic research into the character of human parsing and comprehension. I therefore leave to one side debates about intuitions, ontology, and innateness. My claim is that, however those debates turn out, Devitt's main thesis cannot be sustained. I'll argue that our best psycholinguistic theories of comprehension are committed to the claim that the rules or principles of *some* grammar are psychologically real, in the sense of being *represented* in the minds/brains of linguistically competent individuals.

Devitt (2006) formulates what he calls the *Representational Thesis* (RT) as follows:

(RT) A speaker of a language stands in an unconscious or tacit propositional attitude to the rules or principles of the language, which are represented in her language faculty (p. 4).

It's difficult to say whether any psycholinguistic theory is committed to RT. As stated, the thesis contains several technical terms, the import of which is presently in dispute. In particular, it is a matter of live debate what exactly a propositional attitude is,[1] what it takes for a propositional attitude to be held tacitly,[2] and what it is for a propositional attitude to be nonconscious.[3] Further, there is widespread disagreement regarding the boundaries of the "language faculty," and, indeed, regarding the very existence of such a faculty, and of mental "faculties"

[1] For an account of the structure of propositional attitudes that departs substantially from the influential view developed by J. A. Fodor (1987), see Cummins (1996).

[2] An early discussion of tacit knowledge appears in J. A. Fodor (1968). A much more helpful approach to issues concerning this difficult notion can be found in Davies (1987, 1989, 1995). I draw heavily on this work in §4.

[3] Philosophers working in the field of consciousness studies tend to focus exclusively on qualitative states. For a lucid discussion of consciousness as it pertains to propositional attitudes, see Rosenthal (2005).

more generally.[4] Finally, while a number of theories of representation, intentional content, reference, aboutness, and meaning are currently on offer,[5] it's not at all clear which, if any, is applicable to the case here in question—i.e., the relation between a speaker/hearer and the grammar of his or her language. I will address many of these issues in §4, where I conclude that a suitably modified version of RT enjoys an abundance of empirical support and is likely to remain a core commitment of any viable research program in psycholinguistics well into the future.

I develop my case for this conclusion in two steps. The first step is to show that, in the course of comprehension, the human sentence processing mechanism (henceforth, the **HSPM**) constructs what I will call *mental phrase markers*—i.e., causally active mental representations of the syntactic structures of linguistic stimuli. I secure this premise in §2 by appeal to evidence from priming studies, EEG experiments, and various types of garden-path effect. The second step is to show that the parsing routines that deliver these representations draw on mental representations of grammatical rules or principles. To support this premise, in §3 I survey a range of psychologically plausible computational parsing models, and examine their commitments regarding what data structures are necessary for an effective parsing procedure.

§2. *The Psychological Reality of Mental Phrase Markers*

The claim that the HSPM constructs mental phrase markers during comprehension underlies virtually all of the work in contemporary psycholinguistics. Here are two typical statements of it by leading researchers in the field:

> [L]et us suppose (as we surely should, until or unless the facts dictate against it) that the human sentence processing routines compute for a sentence the very structure that is assigned to it by the mental "competence" grammar. (Fodor, J. D., 1989: p. 157).

> Most models of human language comprehension assume that the processor incorporates words into a grammatical analysis as soon as they are encountered. … We assume that sentence processing involves the computation of

[4] J. A. Fodor (1983) is a classic discussion of faculty psychology and its discontents. More recent work on the topic can be found in Pinker (1999), Coltheart (1999), J. A. Fodor (2002), Carruthers (2006), and Prinz (2006).

[5] Nearly every publication on the topic of intentional content begins with a series of objections to each of the other available theories. Cummins (1991) provides a useful, if somewhat outdated, summary; a more up-to-date catalogue of objections can be found in Neander (2006). For a version of the popular "use theory of meaning," see Horwich (1998). Inferential-role semantics has been forcefully defended by Sellars (1963a,b) and extended by Brandom (1998, 2008). The "asymmetric dependence" approach is an alternative developed by J. A. Fodor (1987, 1990). Fodor's theory builds on the "informational semantics" introduced by Dretske (1981). Teleosemantic theories owe much of their popularity to Millikan (1984). An interestingly different version of teleosemantics is advanced by Cummins (1996). Finally, there are "two-factor" views, e.g., Block (1986), which seek to incorporate the virtues of both causal/informational theories and use theories of meaning.

> dependencies between the words and phrases that are encountered. For example, in the sentence *The troops found the enemy spy*, the relations include information that *the troops* is the subject of *found*. Often, words are incorporated directly into the representation without breaking existing dependencies. For example, when the main verb *found* is encountered, the processor forms the dependency between *the troops* and *found*, and does not need to break any other dependencies (e.g., that between *the* and *troops*). (Sturt, Pickering, and Crocker, 2001: p. 283)

In this section, I present several arguments for the psychological reality of mental phrase markers. I begin by taking a brief look at what is currently known about the neural processes underlying language comprehension, with a particular focus on EEG studies that employ the so-called "violation paradigm." Turning to behavioral studies, I discuss the results of a number of experiments designed to examine a phenomenon known as structural priming. The data from both the EEG work and the structural priming studies provide strong support for the claim that the HSPM constructs representations of distinctly *syntactic* properties of the input. I then turn to the psycholinguistic experiments that reveal "garden-path" phenomena in sentence processing. These are cases in which the HSPM encounters a locally ambiguous input and resolves the ambiguity in a way that turns out to be incorrect relative to the completion of the sentence. I discuss three principles of ambiguity resolution, which form the foundation of most psychologically plausible parsing models. Finally, I discuss evidence for the presence of empty categories—specifically, *wh*-traces—in the mental phrase markers that the HSPM constructs. The data currently available suggest that *wh*-traces are psychologically real and that the HSPM employs rather sophisticated strategies in searching for them, making use of the cues provided by their antecedents as well as considerable knowledge of grammatical constraints.

2.1 *The Argument from Neurophysiological Data*

Let us begin by surveying some of the results emerging from the field of neurolinguistics that bear on the psychological reality of mental phrase markers. In arguing for the claim that "real-time processes assemble syntactic representations that are *the same as those motivated by grammatical analysis*" (emphasis mine), Phillips and Lewis (forthcoming) say the following:

> [S]tudies that use highly time-sensitive measures such as event-related brain potentials (ERPs) have made it possible to track how quickly comprehenders are able to detect different types of anomaly in the linguistic input. This work has shown that speakers detect just about any linguistic anomaly within a few hundred milliseconds of the anomaly appearing in the input. Different types of grammatical anomalies elicit one or more from among a family of different ERP components, including an (early) left anterior negativity ('(e)LAN'; Neville et al., 1991; Friederici, Pfeifer, & Hahne, 1993) … [F]or current purposes the most relevant outcome from this research is that

> more or less any grammatical anomaly elicits an ERP response within a few hundred milliseconds. If the on-line analyzer is able to immediately detect any grammatical anomaly that it encounters, then it is reasonable to assume that it is constructing representations that include sufficient grammatical detail to detect those anomalies.

Below, I describe two of the experiments mentioned in this passage.

EEG devices measure what are known as *event-related potentials* (ERPs). These are small voltage differences between electrical activities in the brain, recorded by electrodes placed on the scalp. In studies of linguistic comprehension, a typical piece of EEG data will have the format exemplified in Figure 1. There, we see a comparison between the ERPs evoked by two distinct stimuli—the critical regions of two German sentences, one grammatical, the other not.

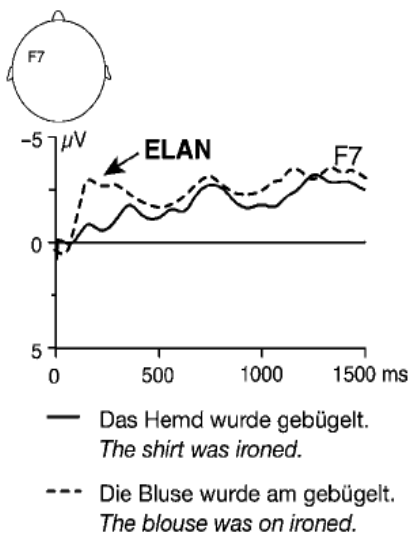

Figure 1: A typical display of EEG data, showing the latency, degree, polarity, location, and distal cause of the neuronal signal. At the top left, the gross location of the activity is specified (in this case, by the symbol 'F7'). The values on the y-axis represent the degree of the signal and its polarity (positive or negative). The values on the x-axis represent the signal's latency—i.e., time at which it occurs, relative to the onset of the stimulus. In this case, the signal is an ELAN—an early negativity in the left anterior region of the brain. As the graph shows, the ELAN occurs roughly 125 milliseconds after the onset of the critical stimulus. Like the studies discussed in the main text, the experiment from which this data was derived used two German sentences as stimuli. The first, indicated by the unbroken line, is a grammatical sentence. The second, indicated by the broken line is *un*grammatical—i.e., it exhibits a basic phrase structure violation. The critical stimulus is the word 'ironed'. The ERP associated with the grammatical sentence are significantly different from the one associated with the ungrammatical sentence. The graph shows a distinct negativity (conventionally plotted upward on y-axis). *Source:* Bornkessel-Schlesewsky and Schlesewsky (2009: p. 110)

The most widely used experimental paradigm in neurolinguistic ERP studies is known as *the violation paradigm*. In this paradigm, participants are shown a variety of sentences, some of which contain violations with respect to one or another linguistic property. For instance, in an early ERP study, Neville *et al.* (1991) used the following materials:

(1) *The man admired Don's <u>of</u> sketch the landscape. syntactic violation
(2) *The man admired Don's <u>headache</u> of the landscape. semantic/pragmatic violation[6]
(3) The man admired Don's sketch of the landscape. control sentence (no violation)

As subjects read such sentences, an EEG device monitors their brain activity, particularly at the crucial regions, underlined in (1) and (2). The ERPs evoked by the anomalous stimuli differ from those evoked by the well-formed stimuli. This yields information about where and when in the brain specific kinds of violation are detected. For instance, Neville *et al.* found that sentence (1) evoked a negative-polarity response in the left anterior region of the brain approximately 125 milliseconds after the onset of the word 'of'. This fast negative-polarity response has come to be known as the ELAN—early left anterior negativity. (See Fig. 1 above.) By contrast, the non-syntactic violation in sentence (2) evoked a negative-polarity response approximately 400 milliseconds after the onset of the word 'headache'. This has been dubbed the N400. Using the materials in (4)-(6), Friederici, Pfeifer, and Hahne (1993) found the same pattern—a replication that is especially striking given the fact that, unlike Neville *et al.*, Friederici *et al.* used German rather than English sentences and presented them auditorily rather than visually.

(4) *Der Freund wurde im besucht. syntactic violation
the friend was in-the visited
(5) *Die Wolke wurde begraben. semantic/pragmatic violation
the cloud was buried
(6) Der Finder wurde belohnt. control sentence (no violation)
the finder was rewarded

Indeed, virtually the same pattern has been observed in dozens of subsequent studies. The natural interpretation is that incoming words are incrementally incorporated into a mental phrase marker, with syntactic information being accessed quite early—125 milliseconds after stimulus onset—while other properties of the stimulus are recovered significantly later.

Consider now sentences like (7), in which a semantic violation is *combined* with a syntactic violation.

(7) *Das Gewitter wurde im gebugelt. combined syntactic and semantic violation
the thunderstorm was in-the <u>ironed</u>

Hahne and Friederici (2002) found that sentence (7) evokes an ELAN, which is characteristic of syntactic violations, but not an N400, which

[6] In view of the notoriously shaky status of the semantics/pragmatics distinction, I simply slur over it in what follows, labeling various properties of sentences 'semantic' regardless of whether they would be classified as semantic or pragmatic by a theorist who insists on drawing the distinction.

seems to be correlated with semantic violations. It appears, then, that the (presumably syntactic) ELAN is capable of *blocking* the (presumably semantic) N400. Crucially, it has also been discovered that the reverse is *not* true; a semantic N400 evoked *prior* to a syntactic ELAN cannot "block" the ELAN. Researchers have thus concluded that "existing ERP findings provide strong converging support for the assumption that constituent structure information *hierarchically dominates* other information types such as semantics/plausibility" (Bornkessel-Schlesewsky and Schlesewsky, 2009: p. 113).

2.2 *The Argument from Structural Priming*

Fodor, Bever, and Garrett (1974) reported a range of studies that they interpreted as demonstrating the psychological reality of mental phrase markers. Among these were the famous "click" experiments, in which subjects monitoring a speech stream misheard short clicks as if they occurred at constituent boundaries. However, as Jurafsky and Martin (2008: pp. 424–5) point out, many of these studies failed to control for semantic biases that correlate with syntactic structure. After all, effects that can be explained by the hypothesis that the HSPM groups words into a *syntactic* perceptual unit can oftentimes be equally well explained by the hypothesis that the grouping is a *semantic* one. Convincing arguments for the psychological reality of syntactic constituency must, therefore, be based on data that can be shown to be independent of semantic effects. Recent evidence from priming studies fits the bill.

The logic of priming phenomena is this: A mental representation activated at time *t*—say, in response to a stimulus—continues to be active for some time after *t*, influencing cognitive processing as long as it persists. In the following passage from Pickering and Ferreira (2008), the authors discuss the importance of priming to recent studies of sentence processing.

> In the past couple of decades, research in the language sciences has revealed a new and striking form of repetition that we here call *structural priming*. When people talk or write, they tend to repeat the underlying basic structures that they recently produced or experienced others produce. This phenomenon has been the subject of heavy empirical scrutiny. Some of this scrutiny has been because, as in other domains in cognitive psychology (e.g., priming in the word-recognition literature; e.g., McNamara, 2005), the tendency to be affected by the repetition of aspects of knowledge can be used to diagnose the nature of that knowledge. [T]he tendency to repeat aspects of sentence structure helps researchers identify some of the representations that people construct when producing or comprehending language. As we shall see, much structural priming is unusually abstract, evidently reflecting the repetition of representations that are independent of meaning and sound. This is therefore informative about how people represent and use abstract structure that is not directly grounded in perceptual or conceptual knowledge. One possibility is that the representations that it identifies can be equated with the representations assumed in formal linguistics. (p. 427)

An early and influential study in this vein is reported in Bock and Loebell (1990). The researchers were careful in eliminating the semantic and lexical confounds mentioned above, by constructing their stimulus materials in such a way as to vary syntactic structure independently of lexical and semantic structure, and vice versa. This was made possible by the fact that some verbs in English are ditransitive—capable of being used in semantically identical but syntactically distinct expressions. Examples of ditransitive verbs include 'give', 'sell', and 'send'. Sentence (8) illustrates a double-object dative construction, i.e., a V–NP–NP structure, while (9) is an example of a prepositional dative construction, i.e., a V–NP–PP structure.

(8) Quentin [$_{V}$ gave/sold/sent [$_{NP}$ Oliver [$_{NP}$ a toy]]].
(9) Quentin [$_{V}$ gave/sold/sent [$_{NP}$ a toy [$_{PP}$ to Oliver]]].

Bock and Loebell's experiment made use of the picture-description paradigm. Participants were first asked to read some sentences out loud. Unbeknownst to them, these served as primes and were selected by the experimenters for having a preposition after the verb (i.e., a V–NP–PP structure), but differing from (9) in their semantics and lexical constituency. For instance, although a sentence like (10) has the same *syntactic* structure as (9), it has almost none of the same words as (9) and, crucially, has a different semantic interpretation—e.g., the preposition carries a locative meaning, as against the dative meaning of the preposition in (9).

(10)IBM [$_{V}$ moved [$_{NP}$ a bigger computer] [$_{PP}$ to the Sears store].

Having read out loud sentences like (10), participants were shown pictures and asked to describe them. The pictures depicted events that involve an agent giving something to someone. Such events can be described equally well by sentences that employ the double-object dative construction (V–NP–NP) and ones that employ the prepositional dative construction (V–NP–PP).

Bock and Loebell found a strong priming effect. Participants who initially read out loud sentences that employ the double-object dative construction were more likely to employ that same construction in describing the events depicted in the pictures. Similarly, those who had recently read out loud sentences that employ the prepositional dative construction were more likely to employ *that* construction in describing the events depicted in the very same pictures. This strongly suggests that the participants in the experiment mentally represented the syntactic properties of the sentences that they were initially asked to read, and then used those representations in repeating those sentences out loud. Having been activated twice—once in the comprehension of the written sentences and once in their spoken production—the representation then remained active in their sentence processing mechanisms, hence more likely to be reused later in production. This priming effect accounts for the participants' choice of construction.

This experiment was one of the first in what has become a long line. Indeed, the area of structural priming research is currently thriving, largely because the results are so robust and the data so telling.[7] Bock and Loebell's initial conclusions have been refined and extended in a number of ways. Pickering and Ferreira (2008) discuss studies which show that structural priming is not restricted to the constructions mentioned above—e.g., it occurs also with active-passive pairs. Moreover it is not due to the presence of common closed-class words in the stimulus materials—e.g., the preposition 'to' in sentences (9) and (10). For instance, Bock (1989) finds priming across sentences with *different* prepositions—e.g., 'to' and 'for'. Nor is structural priming restricted to a single language; the phenomenon has been observed in German, and bilingual English-German speakers even exhibit *cross-linguistic* priming effects—i.e., primed production across their two languages, with respect to suitable constructions. Children and aphasics also exhibit structural priming effects, ruling out the possibility that the phenomenon is restricted to some special set of language users. Further studies rule out the possibility that subjects produce forms similar to the prime because they want to stay in the same rhetorical register, e.g., formal speech. Similarly, structural priming is independent of both prosody and argument structure (i.e., θ-assignment), and can be elicited cross-modally from spoken to written language and vice versa.[8] Finally, the same findings have been replicated using experimental paradigms other than the picture-description paradigm. These include sentence recall, written sentence completion, and spoken sentence completion. Having ruled out a broad range of possible confounds, Pickering and Ferreira write:

> In conclusion, taken together, these results provide compelling evidence for autonomous syntax: The production of a sentence critically depends upon an abstract syntactic form that is defined in terms of part of speech forms (e.g., nouns, verbs, prepositions) and phrasal constituents organized from those (noun phrases, verb phrases, prepositional phrases), and this abstract syntactic form has a large influence upon structural priming. (p. 431)

We may conclude, in a similar vein, that studies of structural priming effects shore up decisive evidence in favor of the psychological reality of mental phrase markers and provide a glimpse into their role in language comprehension and use.

[7] Pickering and Ferreira (2008) even entertain the "intriguing possibility that *all* levels of processing that occur during production show priming and therefore that the absence of priming suggests the absence of a corresponding level of representation" (p. 429, emphasis added).

[8] This contradicts a contention of Devitt (2006) to the effect that there may well be no modality-neutral language faculty. Pickering and Ferreira discuss what they take to be "strong evidence that at least those aspects of structural knowledge that underlie structural priming are modality independent—they are used in the same way both when speaking and when writing" (p. 439).

2.3 *The argument from garden-path effects*

A classic form of argument for the psychological reality of mental phrase markers begins with the observation that competent language users have problems reading and understanding sentences like the following.

(11) Daniel tells students he intrigues to stay.
(12) Jake knows the boy hurried out the door slipped.
(13) The soldier persuaded the radical student that he was fighting in the war for to enlist.
(14) Aron gave the man who was eating the fudge.

From the point of view of formal syntax, all of these sentences contain a local ambiguity that is resolved at or before the end of the sentence. What could explain the fact that even proficient readers encounter measurable processing difficulties with regard to such sentences? The standard explanation appeals to the on-line construction of mental phrase markers.

A parsing routine that computes phrase markers incrementally will update its representation of a sentence in accordance with the words or phrases that it encounters, up to the point at which the sentence becomes syntactically ambiguous. At that point, the parser has to make a choice among the possible ways of continuing the phrase marker that it has thus far constructed.[9] Any ambiguity resolution strategy will sometimes lead a parser to make *incorrect* choices—i.e., choices that give rise to expectations that the remainder of the sentence will serve to disconfirm. Herein lies the explanation of the aforementioned processing difficulties. In sentences like (11)–(14), the human sentence processor's preferred structural assignment turns out to be incorrect by the time the sentence is complete. Additional computational load is then incurred in revising the mental phrase marker—a process known as *reanalysis*.[10] This additional processing burden shows up in behavioral and neurophysiological indicators of processing difficulty, such as error rates and delayed reaction times. The success of this explanation of the observed processing difficulties constitutes evidence in favor of the claim that human sentence processing routines construct mental phrase markers.

Countless instances of this explanatory strategy can be found in the sentence processing literature. Consider for instance, the following sentences.

[9] Here, I assume a serial architecture. Parallel models will bifurcate their processing at this point, building multiple grammatically licensed structures and ranking them.

[10] When a serial model has made a mistake, it will incur extra computational load by being forced to backtrack. Parallel models will likewise incur extra computational load, on account of their having to re-rank the parses that they have constructed and stored in working memory, if only by deleting the disconfirmed parse. For further discussion of this issue, see Crocker, Pickering and Clifton (2000), ch. 1.

(16)The spy saw the cop with a revolver, but the cop didn't see him.
(17)The spy saw the cop with the binoculars, but the cop didn't see him.

Rayner, Carlson, and Frazier (1983) examined the eye movements involved in reading these sentences and others like them.[11] The data show that readers fixate immediately after the word 'revolver' in sentence (16) for a significantly longer time than they do after the word 'binoculars' in sentence (17). (See Fig. 2, below.) This is indicative of a mild hiccup in processing. Rayner *et al.* argue that the increased fixation time is a result of the fact that the HSPM prefers to construct a representation of sentence (16) in which the prepositional phrase 'with a revolver' attaches to the verb 'saw', not the noun 'the cop'. (See Fig. 3.) Fractions of a second later, this initial attachment preference comes into conflict with the reader's encyclopedic information—specifically, a belief to the effect that one is much less likely to use a revolver to see something than to see a person who is in possession of a revolver. This gives rise to a time- and resource-consuming reanalysis of the sentence, in the course of which the prepositional phrase is attached to the noun 'the cop'. By contrast, in the case of sentence (17), the initial attachment preference is consistent with the semantic interpretation of the sentence, so processing is not delayed.

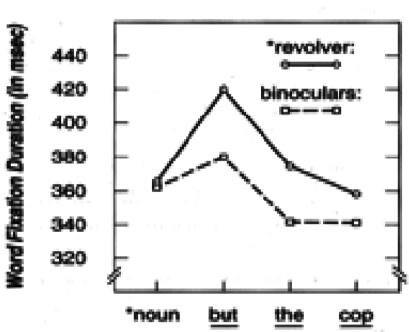

Figure 2: Eye-tracking data from Rayner, et al. (1983). The graph depicts a significant difference between the fixation times at the critical regions of sentences (16) and (17).

The Rayner *et al.* study and numerous others like it serve to illustrate two points. First, there is the now-familiar point that phrase structure is recovered from the linguistic input in the course of processing. That is, a mental phrase marker is constructed, in a manner that is sensitive to the syntactic properties of the stimulus. The second point is that the HSPM seems to construct mental phrase markers in accordance with a quite general ambiguity resolution strategy, known in the literature as *Minimal Attachment*. Minimal Attachment is a least-effort principle, according to which the parser will attach incoming material

[11] For an extensive discussion of the eye-tracking paradigm, see Rayner (1998).

into the existing mental phrase marker in such a way as to minimize the number of nonterminal nodes in the resulting structure. (See Figure 3, below.) The parser's adherence to this strategy also accounts for the garden-path effect associated with the famous sentence (18) and its cohorts, (19)–(21).

(18)The horse raced past the barn fell.
(19)The ship floated down the river sank.
(20)The dealer sold the forgeries complained.
(21)The man sent the letter cried.

The verb-forms 'floated', 'raced', 'sold', and 'sent' are ambiguous. They can serve as either past-tense verbs that are part of a main clause or as past participles that serve to introduce a relative clause. (In these cases, the optional complementizer 'that' has been omitted.) The locally ambiguous structures associated with these sentences are illustrated in Figure 4.

Minimal Attachment is one of three principles that, taken together, explain a wide range of the HSPM's ambiguity resolution preferences. The second such principle is Late Closure, which dictates that the parser will incorporate newly encountered material into the most recent phrase or clause of the mental phrase marker that it has already constructed. Late Closure is invoked to explain the HSPM's preference in cases like the one illustrated in Figure 5.[12]

[12] The tree structures in the diagrams presented below are far simpler than would be posited in contemporary syntactic theories, but they are sufficient to illustrate the structural distinctions relevant to the present discussion.

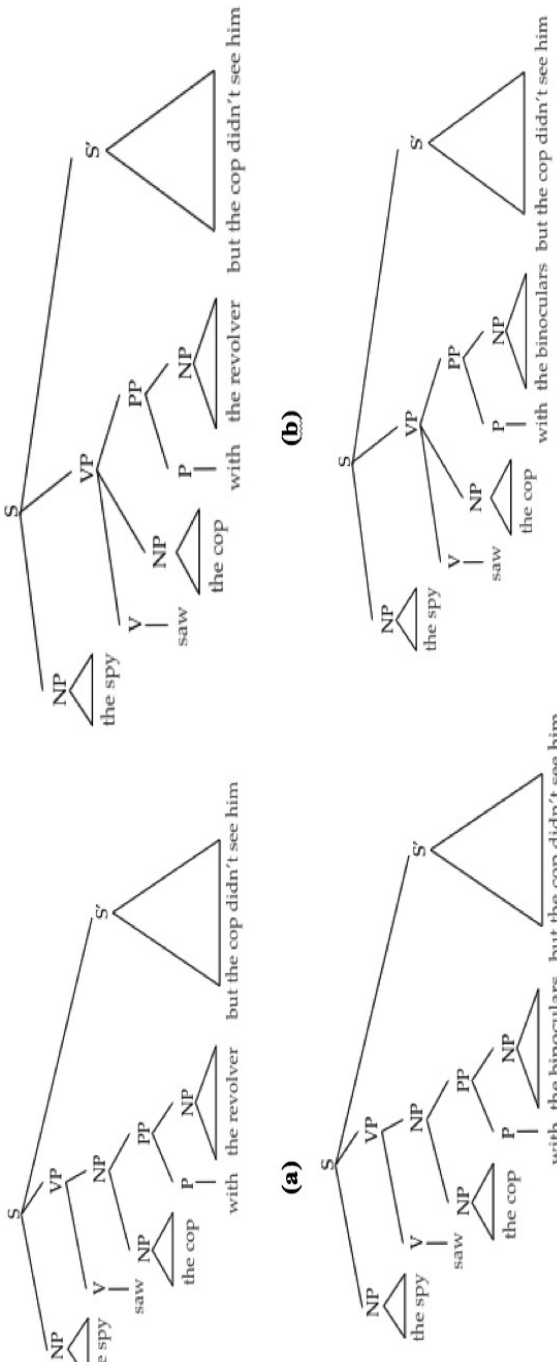

Figure 3: In accordance with the principle of Minimal Attachment, when sentence (16) is the input, the HSPM will initially fail to build the plausible and grammatical structure in **(a)**, preferring instead the semantically implausible structure in **(b)**. The reason is that, in **(a)**, the verb phrase (VP) dominates more nodes than it does in the structure depicted in **(b)**. The additional nonterminal node in **(a)** is an NP, which is needed to accommodate the attachment of the prepositional phrase 'with the binoculars' to the NP 'the cop'. Minimal Attachment rules against constructing a mental phrase marker with this additional internal structure. When sentence (17) is the input, Minimal Attachment likewise dictates avoiding the structure depicted in **(c)** and building instead the structure shown in **(d)**. In this case, the attachment preference turns out to be semantically plausible.

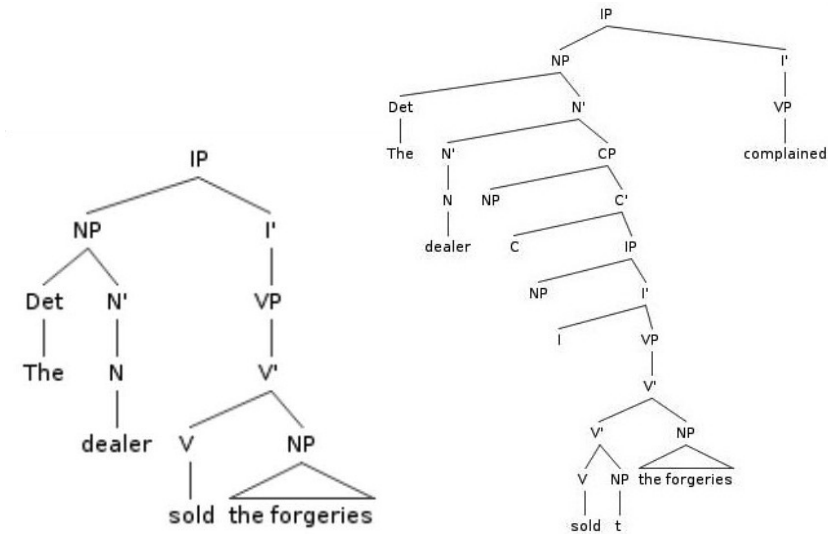

Figure 4: In accordance with the principle of Minimal Attachment, the HSPM assumes that the noun 'the dealer' is the subject of that verb 'sold'. It thus attaches the verb to the existing structure in the manner depicted in the left panel. Subsequent input reveals that the HSPM's assumption was incorrect, thus necessitating reanalysis. The difficulty of reanalysis in this case is a function of the sheer amount of additional structure needed to accommodate a passive participle reading.

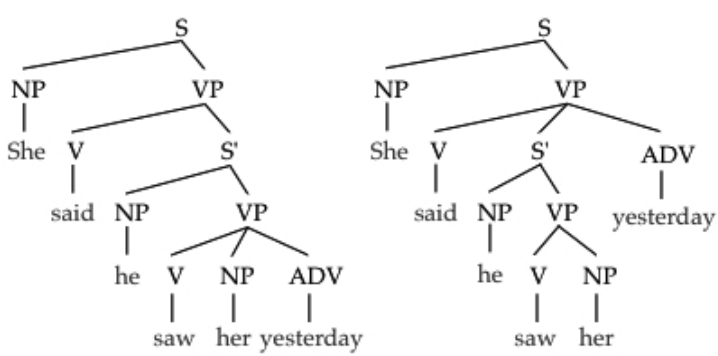

Figure 5: Having built a structure for the input 'She said he saw her', the HSPM receives the adverb 'yesterday'. The grammar licenses two possible attachments, represented by the left and right panels. In accordance with the principle of Late Closure, the HSPM resolves this ambiguity in favor of the structure depicted in the left panel. That is, it attaches the adverb to the most recent phrase of the structure it had already built—in this case, the verb phrase 'saw her'.

The third ambiguity resolution principle, known as the Minimal Chain Principle, pertains to the processing of so-called "filler-gap" constructions, to which we now turn.

2.4 *The Argument from Filler-Gap Processing*

An important feature of the dominant framework in generative linguistics—often referred to as "Principles and Parameters" theory (P&P, henceforth)—is that it posits so-called "empty categories." These are linguistic entities that are not overtly present in writing and speech, but are assumed to occupy a position in the underlying syntactic structures of many types of sentence. Empty categories play a role in explaining why passive sentences, certain kinds of ellipsis, and *wh*-questions (to name just a few constructions) have the semantic interpretations that they do, despite the fact that the relations between the nouns and the verbs in these constructions are not the canonical ones that syntacticians presume to be encoded in the lexicon. For example, the lexical entry for the verb 'consult' specifies that this verb requires a subject to its left and an object to its right. In sentence (22), however, the object of the verb does not occupy its canonical position. Nevertheless, the word 'whom' bears a structural relation to the object position—a relation that is crucial for correctly interpreting the question. To capture this relation, P&P grammars posit an empty category known as "*wh*-trace" in the object position, and *co-index* this entity with 'whom', formally representing this relation with the subscript 'i'.[13]

(22)[Whom]$_i$ did your parents consult ***wh***-trace$_i$ before buying the guitar?

The word 'whom' serves as the *antecedent* of this *wh*-trace. In the jargon of psycholinguistics, 'whom' is said to be the "filler" and the *wh*-trace is said to be the "gap."

While the notion of an empty category helps the syntactician capture significant structural relationships, it poses special challenges for the psycholinguist's theory of comprehension. A model of the HSPM must accommodate the processing of input that is not overtly present in the sound stream or on the printed page. Hence, one might wonder whether empty categories are *psychologically real*—i.e., whether the HSPM bothers to include empty categories in the mental phrase markers that it builds. This question gains urgency in light of the fact that alternative descriptive grammars—rivals of the P&P approach, such as Lexical Functional Grammar and Generalized Phrase-Structure Grammar—encode the relevant structural relations *without* positing empty categories.[14] Thus, a demonstration of the psychological reality

[13] See Haegeman (1994) for detailed coverage of the rich study of empty categories.

[14] This is an oversimplification. For details regarding the way in which LFG and HPSG treat of wh-constructions and relative clauses see Bresnan (2001) and Pollard and Sag (1994) respectively. For treatments of this issue that highlight its relevance to psycholinguistics, see Fodor, J. D. (1989: 177–186) and Featherston (2001).

of empty categories can be used to argue that grammars emerging from the P&P framework more closely resemble the grammar employed by the HSPM than do rival grammatical formalisms.

Fodor (1989, 1995) summarizes a number of experiments designed to address this issue. To illustrate, consider what psycholinguists call "the filled gap effect." Sentences (23) and (24) demonstrate that a *wh*-trace can appear in a variety of locations in the input.

(23) Who$_i$ could the little child have forced ***wh***-trace$_i$ to sing those stupid songs for Jennifer last year?

(24) Who$_i$ could the little child have forced us to sing those stupid French songs for ***wh***-trace$_i$ last year?

Eye-tracking studies reveal a minor hitch in processing at the word 'us' in (24), as compared with the analogous position in (23). The explanation for this, which again appeals to the construction of mental phrase markers, runs as follows: The HSPM is sensitive to the fact that the word 'Who' is an antecedent, to which a *wh*-trace will have to be bound—i.e., a filler awaiting a gap. Thus, it actively searches for legitimate positions at which to posit the *wh*-trace. The HSPM predicts that the gap will occur after 'forced', in both (23) and (24). In the case of (23), the prediction is correct, so comprehension proceeds smoothly. But in the case of (24), the prediction leads the parser astray, giving rise to measurable processing difficulties at just the point in the sentence where the word 'us' occupies the predicted position of the *wh*-trace.

Fodor (1989) notes that the success of this explanation rests on our having independent reason to believe that the HSPM actively hunts for positions at which to posit a gap. There is, after all, no *obvious* reason why the HSPM should predict a gap where there is none, instead of simply waiting to see whether some overt material occupies that position. Nevertheless, as she goes on to argue, the HSPM *does* make active predictions, and these are by no means random or blind. On the contrary, the HSPM appears to be well informed about where a gap would be licensed, which in turn strongly suggests that its parsing routines draw on a grammar to guide its predictions. Phillips and Lewis (forthcoming) echo this conclusion in the following passage:

> [One] body of on-line studies has examined whether on-line structure building respects various grammatical constraints, i.e., whether the parser ever creates grammatically illicit structures or interpretations. Many studies have found evidence of immediate on-line effects of grammatical constraints, such as locality constraints on wh-movement (Stowe, 1986; Traxler & Pickering, 1996; Wagers & Phillips, 2009), and structural constraints on forwards and backwards anaphora (Kazanina et al., 2007; Nicol & Swinney, 1989; Sturt, 2003; Xiang, Dillon, & Phillips, 2009). Findings such as these imply that the structures created on-line include sufficient structural detail to allow the constraints to be applied during parsing.

Let us briefly examine one of the studies that Phillips and Lewis mention. Nicol and Swinney (1989) made use of an experimental paradigm known as *cross-modal priming*. To see how this works, consider sen-

tence (25), which contains a relative clause with a *wh*-trace in the object position of the verb 'accused'.

(25) The policeman saw the boy$_i$ [that$_i$ the crowd at the party accused ***wh***-trace$_i$ of the crime].

In (25), there is only one position at which the *wh*-trace is grammatically licensed and only one noun phrase that can legitimately serve as the antecedent of that *wh*-trace—viz., 'the boy'. However, the sentence contains a number of other noun phrases—'the policeman', 'the crowd', and 'the party'—any of which a linguistically "ignorant" processor might take to be the antecedent. Let's call these *distracters*. To test whether the HSPM is temporarily fooled into taking any of the distracters as the antecedent of the *wh*-trace, Nicol and Swinney had participants listen to sentences like (25) while looking at a computer screen. When the word 'accused' was spoken, participants saw a word appear on the screen. In some trials the word was semantically related to 'boy', which is the antecedent of the *wh*-trace that appears after 'accused'. For example, some participants saw the word 'girl'. In other trials, participants saw words that were comparable in length to 'girl', but semantically related to one of the distracters. Participants were asked to read this word out loud and their reaction times were measured.

Nicol and Swinney made the following assumption: If the HSPM posits a *wh*-trace after 'accused', then it will activate the meaning of the antecedent 'boy' at that point, which would in turn prime the recognition of semantically related words, like 'girl', thus speeding up the participants' reaction times in reading those words.[15] By contrast, the recognition of words that are semantically related to one of the distracters would *not* be primed. And this is precisely what they found. Participants were significantly faster at reading the words that bear a strong semantic relation to 'boy' than they were at reading words that have closer semantic affinities with the distracters. The results were quite robust and have been replicated a number of times.[16] Nicol and Swinney concluded that *wh*-traces are psychologically real and that the HSPM uses sophisticated, *grammatically informed* strategies in actively predicting their occurrence and determining their relation to other items within the syntactic structure of the incoming stimuli.[17]

[15] Priming studies had, by this time, demonstrated quite clearly that the recognition of a word primes the recognition of semantically related words.

[16] It is worth pointing out that other aspects of the data from Nicol and Swinney's experiment provides evidence for a "modular" or "syntax-first" processing architecture. In reporting their findings, they write: "When structural information cannot serve to constrain antecedent selection, then pragmatic information may play a role, *but only at a later point in processing*" (p. 5, emphasis added).

[17] Subsequent research raised an important question about whether these results, and others like them, should be seen instead as semantic rather than syntactic effects. Although a decisive resolution of the ensuing debate is currently out of reach, it is worth noting that recent experiments reported in Featherston (2001) provide strong grounds in favor a view that locates filler-gap effects at a *syntactic* level of representation. These experiments also provide grounds for attributing psychological

On the basis of the findings described above, we can add a processing principle to the list that already contains Minimal Attachment and Late Closure. Frazier and Clifton (1989) referred to the HSPM's strategy for dealing with filler-gap constructions as the "Active Filler Strategy"—a label that reflects the fundamental point that, upon discovering a filler, the HSPM makes active and informed predictions about the occurrence of the corresponding gap in the linguistic input. Soon after, deVincenzi (1989) proposed a more general principle that subsumes the Active Filler Strategy by making use of the notion of a syntactic chain from Government and Binding theory. According to what she called the "Minimal Chain Principle," the HSPM will "avoid postulating unnecessary chain members at S-structure, but [will] not delay required chain members" (199). As deVincenzi pointed out, the second conjunct of this principle is equivalent to the Active Filler Strategy. It states, in essence, that when the HSPM recognizes some aspect of the input as an antecedent, it will posit a gap in the very first position at which a gap is licensed by the grammar.

For our purposes, it is noteworthy that the Minimal Chain Principle makes ineliminable reference to abstract grammatical notions—e.g., 'position at which a gap is licensed'.[18] Note, moreover, that the principle entails that the HSPM will predict a gap in positions where the information provided by the antecedent, *combined with an internal representation of a grammar*, provides sufficient grounds for doing so. Without an internal representation of a grammar, the mere presence of an antecedent would not be sufficient for the HSPM to venture any guesses about where in the input a gap might be found, nor what relations that gap bears to other items in the syntactic structure that the HSPM has already constructed.

2.5 *Summary and further reflections*

The findings reviewed above, garnered from ERP studies, structural priming experiments, and research concerning garden-path and filler-gap processing, all point to the same conclusion: In the course of comprehension, the HSPM constructs explicit representations of the syntactic structure of linguistic input. It goes without saying that these studies are all subject to further scrutiny. Nevertheless, at present, they underwrite a number of powerful arguments for the psychological reality of mental phrase markers.

This conclusion is further supported by morals drawn from the AI literature, specifically the failure of the so-called "mostly-semantics" models developed by Roger Schank and his colleagues in the 1980s.[19] These models eschewed the computation of syntactic dependencies and

reality to other empty categories, in particular *NP*-trace and *PRO*.

[18] If recent versions of Minimalist syntax are correct, then the relevant licensing condition is the Empty Category Principle. Haegeman (1994) and Chomsky (1995) present several formulations of this principle.

[19] For an overview, see Schank and Birnbaum (1984).

attempted to analyze linguistic input by first identifying the thematic structure of predicates in the input and then using the *linear order* of the words in the input string to determine which noun phrases play which of the required thematic roles. In his critique of such models, Marcus (1984) demonstrated that this strategy fails to handle a wide variety of multi-clause sentences, complex passives, and sentences that contain what we've been referring to as "gaps" (which includes most *wh*-questions). Plainly, the inability of mostly-semantics models to cope with such ubiquitous phenomena disqualifies them as plausible candidates for a model of human sentence processing. Moreover, it is arguable that the "brute-causal" model tentatively advanced by Devitt (2006), according to which phonetic representations are mapped *directly* into thoughts, faces the same insuperable difficulties.[20] On the whole, we can be reasonably sure that there is no hope for models that fail to compute mental phrase markers in the course of comprehension.

Our discussion also points to a second and more profound conclusion: The HSPM is not a naïve mechanism. The explanatory success of principles like Minimal Attachment, Late Closure, and the Minimal Chain Principle suggests that the HSPM builds mental phrase markers in a way that is *linguistically informed*. The aforementioned principles all make ineliminable reference to the proprietary notions of an independently motivated syntactic theory—e.g., the notions *number of nonterminal nodes* and *position at which an empty category is licensed*. Thus, in addition to explaining a broad range of experimental data, these principles make it possible to see, if only in dim outline, how one might incorporate a formal syntactic theory into a model of sentence processing—an idea on which Fodor, Bever, and Garrett (1974) cast doubt, for reasons that have come to be recognized as spurious.[21] This has important implications for our understanding of the relation between the HSPM and grammar. As Fodor (1989) observes,

> long-distance binding of traces should provide many pitfalls for a rough-and-ready processor which relies on informal strategies rather than consulting the information provided by the grammar (except perhaps as an emergency back-up). The hypothesis that the [HSPM] is such a device (Fodor, Bever, and Garrett, 1974; Bever et al., in press) becomes quite implausible in the face of the speed and accuracy with which the [HSPM] interprets traces. We can conclude, instead, that the [HSPM] is very closely attuned to the grammar of the language. If that is so, then differences in how the processor responds to different (putative) empty categories can be taken seriously as evidence of how they are treated by the grammar. (Fodor, 1989: p. 205)

The final remark in this passage is particularly significant for the linguist who seeks to formulate a grammar that is not only descriptively adequate but also *psychologically real*.[22]

[20] See Pereplyotchik (forthcoming) for further discussion.

[21] See Berwick and Weinberg (1984: ch. 2), Phillips (1994), and Phillips and Lewis (forthcoming).

[22] Needless to say, descriptive adequacy is a worthy goal, in and of itself, to pursue in constructing a grammar. And, as Devitt (2006) argues at length, it seems

It is important to be clear about the status of the processing principles discussed above. To my knowledge, it has never been suggested that Minimal Attachment, Late Closure, or the Minimal Chain Principle are *represented* in the HSPM, or that the HSPM has *knowledge* of them, however tacit. Rather, these are intended to be *descriptive* principles—they are *true of* the HSPM in much the same way that the principles of celestial mechanics are *true of* our solar system. Still, given that the principles make ineliminable use of the proprietary notions of syntactic theory, it follows that the HSPM works *in accordance* with a particular theory of syntax. There is, of course, a difficult question regarding how to properly cash out this notion of "working in accordance with." An opponent of the Representational Thesis might, at this point, insist that the HSPM doesn't need to *represent* a grammar—implicitly *or* explicitly—in order to act "in accordance" with it. If asked how the HSPM manages to build *just the right structures* and make *just the right predictions* in the course of comprehension, the opponent might reply: "It just *does*!" To fully appreciate the inadequacy of this reply, one must delve into the details of the existing computational models of parsing and comprehension, both classical and connectionist, and to understand *why* all psychologically plausible models require the grammar to be represented—again, either implicitly *or* explicitly. As noted at the outset, the notion of representation is, itself, "up for grabs," so to speak. Accordingly, I will conclude in §4 by clarifying the notions of implicit and explicit representation.

§3. *A Survey of Parsing Algorithms*

In this section, we examine what Devitt (2006) calls the "processing rules" of a comprehension system by surveying computational parsing models that have some claim to psychological plausibility. In these models, internal representations of the rules or principles of a grammar are consulted from the very outset of the parsing process. I will argue that there is simply no way of building mental phrase markers without consulting an internal representation of a grammar. *There is no such thing as a parser without an internally represented grammar.*

3.1 *Context-free grammars, the Earley Algorithm, and the CKY algorithm*

The starting point of many contemporary syntactic theories consists of a list of *context-free phrase structure rules*, examples of which appear below.[23]

to be the *only* goal that many syntacticians pursue. Nevertheless, it's plain that finding a descriptively adequate *and* psychologically real grammar is a much more exciting prospect. Note also that I am passing over Chomsky's important notion of *explanatory* adequacy. The distinction between descriptive and explanatory adequacy raises issues that are well beyond the scope of the present discussion.

[23] The context-free grammar displayed here is adapted from Jurafsky and Martin (2008: p. 394). The symbol '|' expresses exclusive disjunction. Context-free rules

Lexical rules	**Grammatical rules**
Noun → flight | trip | morning	*S → NP VP*
Verb → is | prefer | like | need	*S → VP*
Adjective → cheapest | non-stop	*NP → Pronoun*
Pronoun → me | I | you | it	*NP → Proper-Noun*
Proper-Noun → Chicago | United	*NP → Det Nominal*
Determiner → the | a | an | this	*Nominal → Nominal Noun*
Preposition → from | to | on | near	*Nominal → Noun*
Conjunction → and | or | but	*VP → Verb NP*

A context-free grammar (henceforth, **CFG**) can be used to demarcate a class of well-formed sentences in some language and to describe their hierarchical structure. For instance, the grammar presented above describes sentence (26) as having the structure shown in (27).[24]

(26) I prefer a morning flight.
(27) $[_{S}\ [_{NP}\ [_{Pronoun}$ I]] $[_{VP}\ [_{V}$ prefer] $[_{NP}\ [_{Det}$ a] $[_{Nom}\ [_{N}$ morning] $[_{Nominal}\ [_{N}$ flight]]]]]]

CFGs form the foundation of a wide range parsing models. These can be divided into two broad classes: *top-down* and *bottom-up* parsing.[25] The distinction between these two approaches reflects a more basic distinction, with which philosophers are well acquainted.

> Regardless of the search algorithm we choose, there are two kinds of constraints that should help guide the search. One set of constraints comes from the data, that is, the input sentence itself. … The second kind of constraint comes from the grammar. … These two constraints … give rise to the two search strategies underlying most parsers: top-down or goal-directed search, and bottom-up or data-directed search. These constraints are more than just search strategies. They reflect two important insights in the western philosophical tradition: the rationalist tradition, which emphasizes the use of prior knowledge, and the empiricist tradition, which emphasizes the data in front of us. (Jurafsky and Martin, 2008: p. 433)

comprised the "base" of nearly every transformational grammar prior to the Minimalist program. Even X-bar theory is trivially expressible in a context-free format.

[24] This structure can be represented in bracket notation (as in (27)), tree notation, or in ordinary English. While the brackets render any such claim shorter, trees render it easier to read. Whatever the notation, such representations serve only to make claims about the structure of a sentence. Without further argument, nothing at all follows about the psychology of a language user, nor about the internal operations of a computational system designed to parse sentences of a language. This observation leads Devitt (2006) to draws an important distinction between structure rules and processing rules. Jurafsky and Martin (2008) make plain their recognition of this distinction when they write, "Syntactic parsing … is the task of recognizing a sentence and assigning a syntactic structure to it. This chapter focuses on the kind of structures assigned by context-free grammars ... [S]ince they are based on a purely declarative formalism, context-free grammars don't specify *how* the parse tree for a given sentence should be computed. We'll therefore need to specify algorithms that employ these grammars to produce trees" (431).

[25] Fodor, Bever, and Garrett (1974) refer to top-down and bottom-up techniques as *analysis-by-analysis* and *analysis-by-synthesis*, respectively. Computer scientists sometimes use the terms *recursive-descent* and *shift-reduce*. Mixed strategies, such as left-corner parsing, have also been explored and found to be psychologically plausible in a number of important respects. See, e.g., Abney and Johnson (1991).

A bottom-up parser begins by immediately looking at the input and searching the lexicon in order to determine all of the possible grammatical categories to which each word in the input can belong. Once these are ascertained, the parser begins to build up all of the syntactic structures compatible with those categories and the grammar of the language. The parsing process is completed when the partial structures that were built out of the items in the input are integrated into a full sentence—a structure with an S node at its root. At each step of the process, the parser "looks for places in the parse in progress where the right-hand side of some rule might fit" (Jurafsky and Martin, 2008: p. 434). This, of course, entails that the parser has an explicit representation of the rules and uses them as a template to determine which parses are grammatically licensed. Indeed, the pseudocode of the bottom-up "CKY" algorithm reveals explicit reference to a data-structure in which the rules of a grammar are explicitly represented.[26]

```
function CKY-PARSE(words, grammar) returns table                    ←
        for j←from 1 to LENGTH(words) do
                table[ j−1, j]←{A | A → words[ j] ∈ grammar }       ←
                for i←from j−2 downto 0 do
                        for k←i+1 to j−1 do
                                table[i,j]←table[i,j] ∪
                                        {A | A → BC ∈ grammar,      ←
                                                B ∈ table[i,k],
                                                C ∈ table[k, j] }
```

In contrast to a bottom-up parser, a top-down parser uses its internally represented grammar to issue predictions about the input, *prior* to examining it. For instance, if the internally represented CFG is the one displayed above, then the parser will predict that the sentence consists either of an NP and a VP, or solely of a VP. (These are the only two expansions of the S node that the grammar allows.) In the next phase, the parser "unpacks" these predictions further by predicting, e.g., that the NP consists of either a pronoun, a proper noun, or a determiner and a nominal. Eventually, the parser looks at the input and weeds out all of the predictions that are incompatible with what it finds. The process is complete when all of the input has been accounted for and at least one of the analyses has not been falsified.

In practice, many parsing algorithms adopt a mixed strategy, issuing top-down predictions and then using a bottom-up, data-driven

[26] This elegant algorithm was discovered in the late 1960s by three separate researchers: John Cocke, Tadao Kasami, and Daniel Younger. The algorithm is frequently labeled 'CKY', in honor of its discoverers. The pseudocode displayed above is taken from Jurafsky and Martin (2008), ch. 13. I have used bold arrows to indicate the explicit reference to the rules of the grammar. For instance, the string '{*A* | *A* → *BC* ∈ *grammar*}' can be read as "all nonterminal nodes A, such the grammar contains a rule that expands A into B and C." It is notable that the CKY algorithm has been implemented in a connectionist architecture; see Hale (1999) for details.

approach in mid-sentence to weed out false predictions on the fly. Moreover, to avoid generating the same partially successful predictions again and again, or backtracking through prior choice points and reduplicating work already done, many parsers employ the "dynamic programming" technique of storing successful partial parse trees in a data-structure known as a "chart" or a "well-formed substring table." The successful partial parses are then called on when needed, instead of having to be reconstructed anew. A well-known algorithm developed by Earley (1970) proceeds along these lines. The details are fascinating, but for our purposes the important point is that each of the three main functions of the Earley algorithm—*Predict, Scan,* and *Complete*—draw on an internally represented grammar. This is made clear in the pseudocode, which explicitly refers to a data structure that the authors have aptly labeled GRAMMAR-RULES.[27]

```
procedure PREDICTOR((A → a • B b , [i, j]))
        for each (B → g ) in GRAMMAR-RULES-FOR(B, grammar) do      ←
            ENQUEUE((B → • g , [ j, j]), chart[j])
procedure SCANNER((A → a • B b , [i, j]))                          ←
        if B ⊂ PARTS-OF-SPEECH(word[j]) then
            ENQUEUE((B → word[ j], [ j, j+1]), chart[j+1])
procedure COMPLETER((B → g •, [ j,k]))
        for each (A → a • B b , [i, j]) in chart[j] do             ←
            ENQUEUE((A → a B • b , [i,k]), chart[k])
```

The Earley parser constructs all of the possible parses *in parallel*, which typically requires a great deal of memory capacity when a broad-coverage grammar is applied. Despite this problem, the Earley algorithm is still widely used and has been applied to probabilistic extensions of context-free grammars (Hale, 2001; 2003), as well as to Minimalist and other mildly context-sensitive grammars (Harkema, 2001). Indeed, Hale (2001, 2003) presents evidence for the psychological plausibility of an Earley parser that draws on an internally represented probabilistic CFG.[28] Hale shows that such a model predicts two kinds of processing difficulties that the HSPM is known to exhibit, having to do with the main-clause/relative-clause ambiguity (cf. sentences (18)–(21) above),

[27] The pseudocode displayed here is taken from Jurafsky and Martin (2008: ch. 13). See also Pereira and Warren (1983), Shieber, Schabes, and Pereira (1995), Hale (2001), Harkema (2001).

[28] Hale (2001) makes clear his commitment to what he calls the *strong competence hypothesis*: "What is the relation between a person's knowledge of grammar and that same person's application of that knowledge in perceiving syntactic structure? … *The relation between the parser and grammar is one of strong competence.* Strong competence holds that the human sentence processing mechanism directly uses rules of grammar in its operation, and that a bare minimum of extragrammatical machinery is necessary. This hypothesis, originally proposed by Chomsky (Chomsky, 1965, p. 9) has been pursued by many researchers (Bresnan, 1982) (Stabler, 1991) (Steedman, 1992) (Shieber and Johnson, 1993), and stands in contrast with an approach directed towards the discovery of autonomous principles unique to the processing mechanism" (p. 1, emphasis in the original)

and the asymmetry between subjects and objects in unreduced relative clauses. Examples of the latter, taken from Gibson (1998), appear below.[29]

(28)[S The reporter [S' who [S the senator attacked]] admitted the error].
(29)[S The reporter [S' who [S attacked the senator]] admitted the error].

3.2 *Parsing as Deduction*

An exciting development in computational linguistics, known as the *Parsing as Deduction* approach (**PD**, henceforth), construes the parsing process as a species of natural deduction in either first-order logic or related formalisms. On this approach, parsing routines run through an explicit proof procedure that takes the rules of a grammar as axioms and derives theorems concerning the syntactic structure of input strings. PD constitutes the most concrete implementation of the Representational Thesis, treating rules as *truth-evaluable claims*, which the parser then uses as premises in the course of its inferential procedures.

Pereira and Warren (1983) show how the Earley algorithm discussed above can be reinterpreted in the Parsing as Deduction framework. The basic functions of the parser—*Scan*, *Predict*, and *Complete*—are interpreted as inference rules, on a par with *modus ponens* and *existential instantiation*. Shieber, Schabes, and Pereira (1995) extend this treatment to the CKY algorithm mentioned above and a variety of other algorithms.

> [D]eduction can provide a metaphor for parsing that encompasses a wide range of parsing algorithms for an assortment of grammatical formalisms. We flesh out this metaphor by presenting a series of parsing algorithms literally as inference rules, and by providing a uniform deduction engine, parameterized by such rules, that can be used to parse according to any of the associated algorithms. … As we will show, this method directly yields dynamic-programming versions of standard top-down, bottom-up, and mixed-direction (Earley) parsing procedures. (p. 4)

They also apply the PD approach to syntactic formalisms other than the context-free grammars that we've been considering thus far. Likewise, Harkema (2001) and Hale (2003) apply the approach to Minimalist grammars. These formalisms are widely taken to be capable of providing a more descriptively adequate treatment of natural language than context-free grammars. We shall see shortly that principle-based

[29] Gibson cites the results of a number of psycholinguistic experiments that establish the reality of this processing difficulty: "The object extraction is more complex by a number of measures including phoneme monitoring, on-line lexical decision, reading times, and response-accuracy to probe questions ... In addition, the volume of blood flow in the brain is greater in language areas for object-extractions than for subject-extractions … , and aphasic stroke patients cannot reliably answer comprehension questions about object-extracted [relative clauses], although they perform well on subject-extracted [relative clauses]…" (p. 2).

formalisms, such as the grammar of Government and Binding theory, likewise receive a natural interpretation in the PD approach.[30]

From the point of view of a computational linguist, the PD approach has an immediate payoff in that it makes possible the application of well-known programming techniques in *LISP* and *Prolog* to the task of parsing natural language. For philosophical purposes, the payoff is quite different, though no less intriguing. It is difficult to find up-and-running examples of psychological processes *convincingly* modeled as deductive operations defined over truth-evaluable statements. The PD approach shows that such an example is available in the case of at least one aspect of natural language comprehension. If the CKY or Earley algorithms play a role in a plausible model of the HSPM, then there is a perfectly workable account according which the neural mechanisms that underpin comprehension can be said to be carrying out deductive procedures. The availability of such an account can be brought to bear on the recent debate in the philosophy of linguistics, concerning whether knowledge of language—i.e., grammatical competence—is strictly-speaking *propositional*.[31] The fact that a rich array of successful parsing models can be represented in a familiar propositional format shows, at the very least, that it is not incoherent to suppose that knowledge of language should be propositional. Nevertheless, I will argue in §4 that it is better to regard the HSPM as a set of *subpersonal* processes, hence not consisting of propositional attitudes, in the full-blooded sense of the term.

3.3 *Principles and Parameters in Syntax and Parsing*

The Standard Theory of transformational grammar (Chomsky 1957, 1965) posited transformational rules over and above context-free phrase structure rules. Each linguistic construction—passive, question, etc.—was associated with a distinct transformational rule. In addition, many transformations were language-specific; their inputs and outputs would differ from one language to the next. This proliferation of transformational rules came to be seen as a serious problem. Subsequent work, eventuating in the Principles and Parameters theory (P&P, discussed above), heralded a striking revision of transformational grammar in the 1980s. The new theory did away with most transformational rules. Only a single rule remained: *Move*-α. This rule says that any constituent appearing at D(eep)-structure can be moved *anywhere* in the phrase-structure tree on the way to S(urface)-structure. Left unconstrained, *Move*-α would generate a great many S-structure phrase markers that fail to correspond to anything one finds in natural language. To eliminate these unwanted results, P&P grammars invoke a set of syntactic principles, e.g., the Projection Principle, the Theta Criterion, the Case Filter, the three Binding Principles, and the Empty Category Principle, which act as a sequence of filters.

[30] See Berwick (1991a,b) and Johnson (1989).
[31] See Knowles (2000) and Rattan (2002).

With the rise of the P&P framework in formal syntax, it was only natural that computational linguists would embark on the project of implementing these novel ideas in parsing models. From the start, principle-based parsers made heavy use of the PD approach.

> deductive inference is still perhaps the clearest way to think about how to 'use' knowledge of language. In a certain sense, it even seems straightforward. The terms in the definitions like [that of the Case filter] have a suggestive logical ring to them, and even include informal quantifiers like *every*; terms like *lexical NP* can be predicates, and so forth. In this way, one is led to first-order logic or Horn clause logic implementation (Prolog) as a natural first choice or implementation, and there have been several such principle-based parsers written… Parsing amounts to using a theorem prover to search through the space of possible satisfying representations to find a parse… (Berwick, 1991a)

One important feature of principle-based parsers is their flexibility with regard to different languages. Yang and Berwick (1996) point out that "[t]raditional parsing technologies utilize language-particular, rule-based formalisms, which usually result in large and inflexible systems." By contrast, a principle-based parser can be used to assign structure to sentences from a variety of typologically distinct languages, simply by swapping out one lexicon for another and setting the parameters to the values characteristic of the target language.

> It is believed that languages are constrained by a small number of universal principles, with linguistic variations largely specified by parametric settings. The merit of principle-based parsing is two-fold. As a tool for linguists, it is directly rooted in grammatical theories. Therefore, linguistic problems, particularly those that involve complex interactions among linguistic principles, can be cast in a computational framework and extensively studied by drawing directly on an already-substantiated linguistic platform. It is designed from the start to accommodate a wide range of languages — not just 'Eurocentric' Romance or Germanic languages. Japanese, Korean, Hindi and Bangla have all been relatively easily modeled in PAPPI (Berwick and Fong 1991, …). Differences among languages reduce to distinct dictionaries, required in any case, plus parametric variation in the principles. … Because the PAPPI system implements its model linguistic theory faithfully, adapting new languages is expected to be quite minimal, as our implementation shows. (Yang and Berwick, 1996)

Besides their flexibility with respect to distinct languages, principle-based parsers exhibit tolerance with respect to ungrammatical input. On its way from X-bar analysis to the ultimate assignment of structure, linguistic input passes through a series of filters. Rather than grinding to a halt when presented with ungrammatical input, a principle-based parser simply takes note of which principles of the grammar the input fails to satisfy. This gives rise to variable-strength judgments of ungrammaticality; the more principles are violated, the more ungrammatical the input is judged to be. Nevertheless, as it runs this gauntlet, almost all linguistic input is assigned *some* interpretation, in a way that mirrors what we observe in human performance.

In recent years, the P&P framework has moved away from the formalism of Government and Binding theory. Seeking a more elegant and compact account of syntactic structure, Chomsky (1995) and others have developed a variety of Minimalist grammars, in which the basic operations of *Merge* and *Move* forge hierarchical relations between lexical items, conceived as sets of features. Unsurprisingly, computational linguists have built parsing models that incorporate these grammars as well. For example, the algorithms developed by Harkema (2001) and Hale (2003) employ the PD approach in constructing Earley and CKY algorithms for parsing with Minimalist grammars. In addition, Berwick (1997) and Weinberg (1999), both attempt to account for a number of psycholinguistic results by seeing parsing processes as "the incremental satisfaction of grammatical constraints" imposed by Minimalist grammars.

3.4 *Efficiency*

It is well known that, *in their purest and simplest forms*, both top-down and bottom-up parsing algorithms are so inefficient as to be practically unusable for realistic natural language processing. Devitt (2006) puts the point vividly in the following passage:

> How can the represented rules be used as data in language use? Consider language comprehension. Suppose that, somehow or other, the processing rules come up with a preliminary hypothesis about the structure of the input string. In principle, the represented rules might then play a role by determining whether this hypothesis could be correct (assuming that the input is indeed a sentence of the language). The problem in practice is that to play this role the input would have to be tested against the structural descriptions generated by the rules and *there are just too many descriptions*. The "search space" is just too vast for it to be plausible that this testing is really going on in language use. This led Fodor, Bever, and Garrett to explore the idea that heuristic rules not representations of linguistic rules govern language use. (p. 209)

Unlike the authors of standard texts in computational linguistics, Devitt takes this to be a powerful argument against models that make use of internally represented grammars. But this is a mistake. The stark inefficiency of *simplistic* models does not constitute grounds for rejecting *every* model that is committed to internal representations of a grammar.[32]

Computational linguists have devised an array of strategies for minimizing the inefficiency that Devitt points to. Some of these are

[32] Incidentally, the approach that Devitt mentions—i.e., the heuristic strategies proposed by Fodor, Bever, and Garrett (1974)—is known to be untenable. As Pritchett (1992: pp. 22–26) points out, their "canonical sentoid strategy" and associated heuristics are simply not adequate for explaining a wide range of processing data. Psycholinguists have abandoned Fodor, Bever, and Garrett's approach, pursuing more promising avenues of research, from which a variety of processing principles have emerged, including those discussed in §2. These cut down the parser's search space, but in a way that *presupposes* the parser's internal representation of a grammar.

best characterized as implementations of Minimal Attachment, Late Closure, and the Minimal Chain Principle (§§2.3, 2.4). Others, like the Earley and CKY algorithms, employ clever computational tricks, such as the storage and reuse of partial solutions (§3.1). Still others have a different flavor, making heavy use of statistical methods to reduce error and avoid ambiguity. For instance, a common approach involves enriching the context-free phrase structure rules discussed above with information about the probability of their application in a given context. Probabilistic grammars encode the frequency with which lexical items, syntactic categories, and even phrase structures appear in a corpus. It has been hypothesized that such frequencies have an explanatory relation (as yet not fully understood) to the ambiguity resolution principles employed by HSPM.[33] For present purposes, the important point is that in order to implement such frequency-based principles, a model must still draw on internal representations of a grammar.[34] The probabilities have to be attached to *something*, and that "something" must be represented, either explicitly or implicitly. Thus, statistical methods are not *alternatives* to the parsing strategies that make use of an internally represented grammar. Rather, they are *extensions* of those strategies.

Similarly, in the early stages of their development, principle-based parsers faced a number of challenges, centering mostly on issues to do with computational inefficiency. But here, too, impressive gains in efficiency were made possible by the judicious application of clever programming techniques, such as the "co-routining," "interleaving," and formal compilation of Government and Binding principles.[35]

> We are able to parse sentences with the range of structures including Wh-movement, the Binding Theory, Quantifier Scoping, the BA-construction to complex NP (clausal, possessive, and numeral/ classifier). All testing sentences are correctly analyzed: LF logical form representations are computed for the grammatical sentences and the ungrammatical ones are ruled out ones with linguistic principle violation(s) shown. Each parse takes no more than 2 seconds on a Sparc10 workstation. (Yang and Berwick, 1996: p. 370. Note that today's processors would require only a fraction of the time. –D.P.)

In short, the models currently being explored in both psycholinguistics and computational linguistics are premised on the idea that structure rules are represented and used "as data" for the purpose of parsing and comprehension. Contrary to some of the remarks in Devitt (2006), this idea has *not* been abandoned. Rather, it has been taken as a starting point, to which various modifications are made, in an effort to increase both efficiency and psychological plausibility.

[33] See Jurafsky (2003) for discussion.

[34] Jurafsky and Martin (2008: ch. 14) and Hale (1999, 2001) provide working examples of such models.

[35] See Berwick (1991a,b) for in-depth discussion of these techniques.

3.5 *Connectionist Models of Sentence Processing*

In recent years, a number of connectionist models of sentence processing have been developed.[36] The pioneering work of Elman (1992) led to a number of follow-up studies, which yielded interesting results concerning the ability of simple recurrent networks to perform as if they were explicitly representing grammatical dependencies, without actually doing so. There are, of course, familiar worries about whether such models can be scaled up to achieve broad coverage, or to perform tasks that are more realistic (from the point of view of psychology) than predicting the lexical category of a word on the basis of prior input. In light of our conclusions in §2, it is reasonable to suppose that no connectionist model can attain psychological plausibility without computing phrase markers for a wide range of natural language inputs.

Fortunately, such models are, in fact, available. The PRISM model presented in Hale (1999) is a connectionist implementation of the CKY algorithm discussed earlier (§3.1). Hale demonstrates how a context-free grammar, suitably represented, can be used by a connectionist network to compute syntactic structures. Adopting a hybrid classical-connectionist approach, Stevenson (1994) presents a model in which lexical items compete for attachment in a syntactic structure—a competition governed by connectionist principles, but in accordance with an internally represented phrase-structure grammar. Stevenson's model provided the basis for the more recent work reported in Stevenson and Smolensky (2006), where the authors argue that "an [Optimality Theoretic] grammar that is well-motivated from the perspective of theoretical syntax can explain on-line parsing preferences of comprehenders, as evidenced by empirical data on the processing of sentences which, at intermediate positions, have various structural ambiguities" (p. 829). Finally, Gerth and beim Graben (2009) mobilize a representation of a Minimalist grammar in a connectionist parsing model that achieves some measure of psychological plausibility, again by replicating the processing difficulties that we find in human comprehension.

Connectionist research has, moreover, enriched our stock of tools for assigning interpretations to the internal states of neural networks. Statistical techniques such as *Principal Components Analysis* reveal how well-defined regions of a network's vector space can track abstract linguistic properties, e.g., information about a verb's subcategorization frame. Moreover, in a landmark study, Tabor and Tanenhaus (1999) demonstrated how concepts from dynamical systems theory can increase our understanding of what interpretations can be reasonably assigned to the activation vectors of the hidden nodes in a connectionist network.[37] All in all, the interpretive limitations that we currently

[36] See Rohde (2002), ch. 2, for a historical overview.

[37] To be fair, it should be noted that Tabor and Tanenhaus would strongly resist the claim that their models contain internal representations of grammatical principles; they believe that such models are tracking only complex statistical

face seem unlikely to constitute a *principled* problem; such limitations will almost certainly be overcome as research progresses. And the more we learn about how to interpret the inner workings of successful connectionist models, the more opportunity we will have to see that their success is due to the fact that they represent grammatical principles, either implicitly or explicitly.[38] Progress toward this goal is aided by productive efforts on the part of philosophers to spell out the conditions on implicit representation. Drawing on the particularly helpful taxonomy developed by Davies (1995), I will argue in the next section that it is incorrect to say that connectionist networks must, *by their very nature*, fail to represent a grammar. Hence, even if connectionist models ultimately capture the empirical findings about human parsing performance better than competing classical models, it would still not be correct to say, as Devitt (2006) does, that there is "no sign of the structure rules of the language governing the process of comprehension" (240).

§4. *Representation, Functionalism, and Subpersonal Psychology*

Recall that Devitt (2006) argues against the representational thesis (RT), repeated here:

(RT) A speaker of a language stands in an unconscious or tacit propositional attitude to the rules or principles of the language, which are represented in her language faculty (p. 4).

The thesis I wish to defend is a modified version of RT that eschews Devitt's relational conception of representation, as well as his appeal to the notion of a propositional attitude. The relational conception emerges most clearly in the following passage, which is the only characterization of representation that one finds in Devitt (2006).

> [T]alk of representing rules raises a question: What sense of 'represent' do I have in mind in RT? The sense is a very familiar one illustrated in the following claims: a portrait of Winston Churchill represents him; a sound / the President of the United States / represents George W. Bush; an inscription, 'rabbit', represents rabbits; a certain road sign represents that the speed limit is 30 mph; the map on my desk represents the New York subway system; the number 11 is represented by '11' in the Arabic system, by '1011' in the binary system, and by 'XI' in the Roman system; and,

properties of the items in the corpus of training data. I am not persuaded that their experiments reveal this. Putting aside serious worries about their methodology—particularly the construction of their training corpus—it remains the case that statistical frequencies in any natural language corpus are confounded with a wide range of syntactic, semantic, and pragmatic effects. Further experimental work will be needed to decide this issue.

[38] See, e.g., Lawrence, Giles, and Fong (2000) for an attempt to extract rules, in the form of deterministic finite-state automata, from a variety of recurrent neural networks that have been trained to distinguish grammatical from ungrammatical sentences.

> most aptly, a (general-purpose) computer that has been loaded up with a program represents the rules of that program. Something that represents in this sense has a *semantic content*, a *meaning*. When all goes well, there will exist something that a representation refers to. But a representation can fail to refer; thus, nothing exists that 'James Bond' or 'phlogiston' refer to. Finally, representation in this sense is what various theories of reference—description, historical-causal, indicator, and teleological—are attempting to partly explain. (p. 5)

In this section, I explain why propositional attitudes are *not* the kind of state involved in the early stages of language comprehension and I sketch a framework for thinking about the kinds of representations that do play a role in sentence processing. Furthermore, I argue that what Devitt (2006) calls "merely embodied" rules should be seen as *represented* rules.

4.1 *Representation and the Personal/Subpersonal Distinction*

I take as my starting point a functional-role theory of representation, according to which the representational properties of an event, state, or process are exhaustively determined by three factors: i) the environmental conditions under which it is typically elicited, ii) the causal relations it typically bears to other events, states, or processes, and iii) its typical behavioral consequences. Contrary to the standard versions of behaviorist, covariationist, and information-theoretic accounts of representation, I believe that none of the three factors is individually sufficient; only a combination of all three can be both necessary and sufficient for something to have the representational properties that it does.[39]

Following Sellars (1963a), I hold that the functional-role theory of *intentional content* applies, in the first instance, to speech acts. But the need to account for speech and rational action gives rise to the theoretical posit of internal states—the propositional attitudes. These are personal-level states, whose role in a creature's cognitive economy is captured by a suitable formulation of folk psychology.[40] The folk-psychological posit of propositional attitudes suffices only for a relatively coarse-grained way of describing a creature. Digging deeper, one wants to know how a creature can so much as *have* such states. Here, the strategy of attributing *sub*personal mechanisms becomes useful. Focusing specifically on the case of language comprehension, we can say the following: Whereas folk psychology lets us talk about the sensation of a sound giving rise to an act of linguistic comprehension—judging that a speaker said that *p*—cognitive psychology tells us what happens *between* the sensation and the judgment.

Dennett (1987) draws a useful distinction between taking "the intentional stance"—i.e., using folk psychology to predict, explain, and

[39] For compelling arguments to this effect, see Brandom (1998).
[40] Lewis (1972)

describe a creature's behavior—and adopting "the design stance," which involves thinking of a creature as an aggregate of purposeful mechanisms, each of which has the function of performing a specialized task. It's from the design stance that we attribute subpersonal states to the HSPM. These states have *some* of the features that we take to be characteristic of personal-level states. In particular, they bear systematic relations to the environment, to behavior, and to one another. This is what makes it both reasonable and useful to think of them as representations—i.e., to interpret them in something like the way that we interpret the speech acts we find in a syntax textbook. Doing so allows us to abstract away from the largely unknown neural mechanisms that underpin subpersonal states and to see the causal relations between these states as resembling the *inferential* relations that hold between propositional attitudes. This, in turn, allows us to *rationalize* subpersonal mechanisms—i.e., to understand them as being engaged in purposeful activities and as taking reasonable steps toward accomplishing their goals. These similarities between personal and subpersonal states make it tempting to think of the latter as a kind of propositional attitude. But that would be a mistake, for there are salient and well-known differences between the two kinds of state.

First, subpersonal states are not "inferentially integrated" with personal-level states. We cannot draw personal-level inferences whose premises are the rules or principles of the grammar that we subpersonally represent (or, to use Chomsky's neologism, "cognize").[41] Second, subpersonal states like those involved in the early stages of language comprehension are not expressible in speech. Third, subpersonal states are *always* nonconscious, whereas personal-level states are sometimes conscious and sometimes not. (For this reason, the conscious/nonconscious distinction should *not* be confused with the personal/subpersonal distinction.) Fourth, subpersonal states are susceptible to a computational description, whereas there are well-known and potentially insurmountable problems—loosely captured under the label "the frame problem"—for the enterprise of giving a computational description of personal-level states.[42] Fifth, there is no reason to believe that subpersonal states are composed of *concepts*, in any sense of that vexed term, whereas personal-level states are paradigmatically conceptual.

Finally, subpersonal states are best characterized by their natural functions—the purpose for which they were selected—whereas the possession of a great many personal-level states (e.g., thoughts about quarks) cannot be explained by a straightforward appeal to natural selection. As Dennett makes clear, adopting the design stance involves assuming only that a system has a purpose, and that it is not currently malfunctioning, whereas taking the intentional stance involves making more weighty assumptions—e.g., that the system is a rational

[41] See Stich (1978) for further elucidation of this point.

[42] See Fodor (1983: p. 107). See also Putnam (1988) and Brandom (2008), ch. 3.

agent whose terms mostly refer, whose judgments are mostly true, and whose inferences are mostly good.

Functionalism in the philosophy mind is arguably the reigning orthodoxy. But there are, broadly speaking, two kinds of functionalism, and it matters for present purposes which of the two we adopt. One kind stems from the work of Sellars and Lewis, who emphasize personal-level states and folk-psychological descriptions, pitched from the intentional stance. The other, which stems from the early work of Hilary Putnam and Jerry Fodor, deals with computational states, of the sort the cognitive psychologist posits. The latter brand of functionalism has led philosophers astray, by blurring the distinction between personal and subpersonal states and making it seem as though the states involved in language processing are propositional attitudes.[43] Chomsky (2000), Matthews (2007) and Collins (2008: ch. 5) expose the errors of this view.

One such error is embodied in a widespread commitment to a relational view of representation, according to which representation is a special kind of causal or nomological link between a symbol and something in the world. The project of naturalizing this "relation" has given rise to no shortage of theories, all of which are known to face grave difficulties. (See §1, fn. 5.) I believe that the classic work of Sellars (1974), as well as the more recent views of Brandom (1998), Chomsky (2000), and Matthews (2007), provides compelling grounds for a *non-relational* view of representation. On this view, to say of some state S that it represents X is *not* to claim that the S bears some special relation to X. Rather, to say that S represents X is to mark that state as *belonging to a particular type*—i.e., as playing a distinctive role in the creature's cognitive and behavioral economy. There are, of course, many interesting relations between a creature's psychological states and its environment. And such relations may even turn out to be key ingredients in our account of those states' representational properties. (Indeed, on the functional-role view that I favor, these relations may well constitute two thirds of such an account.) Nevertheless, there is no single, unified "representation relation." No mind-world relation deserves that title, regardless of its causal, nomological, or naturalistic credentials.

I extend this view to the notion of reference. Referring is best seen as something that people (and perhaps other creatures) *do*—a kind of communicative act, not a relation between linguistic expressions and objects, properties, or events.[44] I urge that we see representation as a

[43] A particularly striking instance of this conflation can be found in Fodor, Bever, and Garrett (1974), ch. 1.

[44] To be clear, no "idealism" or "anti-naturalism" is in play here. Naturalism should not be held hostage to the view that any of the numerous relations that we bear to the extra-mental world in perception, thought, and behavior is usefully labeled "the reference relation." Nor need there be such a relation for us to sustain the eminently reasonable doctrine of scientific realism. See Devitt (1997) and Chomsky (2000) for two defenses of these claims.

more inclusive notion than reference. Whereas talk of reference has its home in personal-level descriptions of speech acts and propositional attitudes, extending the notion to other cases—e.g., tree rings, computer programs, maps, pictures, and subpersonal psychological states—yields awkward consequences. ("The tree ring refers to the age of the tree." "The portrait refers to Plato.") By contrast, the notion of representation extends comfortably to all such cases.

4.2 *Computational Psychology and Embodied Representation*

Taking on board the main claims of the foregoing discussion, let us restrict our focus solely to the subpersonal level of description, and employ a view of representation as "functional classification" (to use Sellars' handy term). We are now in a position to be more precise about the kinds of commitments that a psychologically plausible computational model might make about the representations involved in language processing.

Stabler (1983) defines a variety of claims that can be made within the framework of computational psychology. On his view, a computational system is one that goes into a physical state that represents the output of some function whenever the system is in a state that represents the corresponding input to that function. We say that a system computes a function when this causal pattern is "regular and predictable" enough for it to be "convenient" for us to so describe it; the description is used because it is "clear and useful." (The quoted phrases are from Stabler, fn. 1.) In Stabler's terms, saying that a system computes a function is giving a *first-level theory*. This doesn't tell us *how* the system performs the computation. If we want a more substantial description of the system, we might specify the *program* that the system uses in computing the function. The program consists of the instructions that the machine carries out at each step between the input and the output. Each instruction is associated with a more basic function, and the sequential computation of basic functions produces the output of the target function. Specifying such a program is giving a *second-level theory*. Some systems—e.g., the modern-day personal computer—compute many functions by representing, in their memory banks, the very programs that they are computing. Such systems have "control states" that encode the instructions of a program and are causally involved in the inner workings of the machine. Saying which programs are *encoded* in the system (rather than merely computed by it) is giving a *third-level theory*.[45] Importantly, not all computers are "stored-program" systems of this kind. Some are hardwired circuits, for which we can give a second-level but not a third-level theory. Such circuits can be said to "embody," rather than to encode, a program.

[45] This is what Davies (1995) means by "explicit representation." Davies also draws a distinction between two notions of implicit representation, both of which have their home in what Stabler would call a second-level theory.

If we say that a grammar is psychologically real in the sense that *(i)* a speaker computes a function from linguistic stimuli to judgments about what was said, and *(ii)* the grammar is *true of* those stimuli, then we are offering a first-level theory—or what Devitt (2006) calls "position (M)" ('M' for 'minimal'). This theory would be true, as far is goes, but it wouldn't go very far; first-level theories are almost totally uninformative, telling us only what rules a device *conforms* to, extensionally. Following Chomsky, we should aim for a more interesting thesis to the effect that grammatical rules and principles are *mentally represented*, that language processing is *governed* by the rules of the grammar (rather than merely conforming to them), and that the mentally represented grammar is used to *generate* mental phrase markers, in the causal sense of the term. Stabler's three-level framework allows us to distinguish two forms that such a thesis might take. According to what I will call the "strong thesis," we can give a third-level theory of the HSPM, on which it contains an explicit representation of a grammar and draws on this representation "as data" in the course of comprehension. By contrast, the "weak thesis" has it that the HSPM is susceptible only to a second-level theory, on which the grammar is embodied but not encoded in the hardwired circuitry of the brain.[46]

Now, a number of leading figures in psycholinguistics have expressed what appears to be a commitment to the strong thesis. Consider, for instance, the following passage from Frazier and Fodor (1978), in which the authors draw out the consequences of their parsing model.

> when making its subsequent decisions, the executive unit of the parser refers to the geometric arrangement of nodes in the partial phrase marker that it has already constructed. It then seems unavoidable that *the well-formedness conditions on phrase markers are stored independently of the executive unit, and are accessed by it as needed*. That is, *the range of syntactically legitimate attachments at each point in a sentence must be determined by a survey of the syntactic rules for the language*, rather than being incorporated into a fixed ranking of the moves the parser should make at that particular point . . . (322n, emphases added).

But, as Stabler (1983) points out, it's provable that any computation performed by a stored-program computer can also be performed by an assembly of hardwired circuits that don't *encode* any instructions—a system for which can give only a second-level theory. Hence, it may well be that the HSPM is such a device, in which case only the weak thesis is warranted. Moreover, Stabler goes on to argue that we do not, at present, have any behavioral or neurophysiological data to support

[46] The strong and the weak theses bear a close resemblance to what Stabler (1983) calls "(Hd)" and "(H2)" respectively and also to what Devitt (2006) calls "position (ii)" and "position (iii)" respectively. There are, however, subtle differences between these claims. First, it's possible that Stabler would not take (Hd) to carry any commitment to a third-level theory, though his reasons for this are obscure, as many authors in the BBS peer commentary point out. Second, Devitt sees position (iii) as entailing the claim that a grammar is not mentally *represented*, whereas it's not clear that Stabler would agree. Indeed, I argue below that he should *not* agree.

the stronger of the two theses. On the basis of Stabler's conclusions, Devitt (2006) concludes that we have no grounds for supposing that a grammar is mentally *represented*. It is this last inference that I shall challenge in the remainder of the present discussion.

Suppose that Stabler were correct in saying that we have no persuasive evidence for the strong thesis that a grammar is explicitly encoded in the brain. We would then retreat to the weaker thesis that a grammar is merely embodied. But would it follow from this weaker thesis that the rules of the grammar are not mentally *represented*, as Devitt concludes? I think it would not. The inference from 'embodied' or 'not encoded' to 'not represented' is, in general, unwarranted. There are important distinctions to be made between kinds of representation, even at the subpersonal level of description. Encoding is, of course, a *species* of representation—what we might call "explicit representation"—but it is not the only one. There is, in addition, a notion of *implicit* representation of rules, which Davies (1995) elaborates as follows: A device that effects transitions between a set of inputs and their respective outputs can be said to have an implicit representation of a rule just in case the device contains a state or a mechanism that serves as a *common causal factor* in all of those transitions. In a series of clever examples, Davies demonstrates that there are ways of satisfying this definition which are logically weaker than explicit encoding but, at the same time, *stronger* than mere "conformity" with a rule.[47]

Armed with Davies' account of implicit representation, we may conclude that the rules of a grammar may be represented in a system, even if that system is susceptible only to what Stabler calls a second-level theory. As long as it has the right kind of structure—i.e., one that admits of explanations that appeal to common causal factors—even a hardwired circuit that computes a function from linguistic stimuli to judgments about what speakers express can be both usefully and reasonably interpreted as *representing* a grammar, albeit implicitly. This conclusion, significant in its own right, has a bearing on how we regard the connectionist models of sentence processing discussed above (§3.5). For, in many cases, it is useful to think of the trained-and-frozen connectionist network as an instance of a hardwired circuit, susceptible only to a second-level theory.

I end by addressing one final point in Devitt's argument against RT. Devitt (2006) claims that the structure rules of a language—i.e., grammatical rules and principles—may be the "wrong sort of rules to govern [the process of comprehension]" (53). Plainly, quite a bit hangs on how we construe his use of the term 'govern'. On one reading, Devitt's claim is certainly plausible. For, as we have seen (§3.1, fn. 24), the structure rules of a language are not to be confused with the processing

[47] As noted above (fn. 45), Davies makes further distinctions *within* the category of implicit representation. Moreover, his account has close affinities with the one developed in Peacocke (1989). I will pass over these subtleties here. Note also that Devitt (2006) cites Davies' work with approval (p. 52, fn. 10).

rules that constitute a parsing algorithm. Rather, the structure-rules are treated *as data* by such algorithms. This, in turn, yields a different reading of 'govern'—which Devitt also recognizes—according to which grammatical principles govern sentence processing by being used as premises in the "deductions" that constitute a parsing routine.[48]

Devitt seems to assume that grammatical principles can only be drawn on as data if they are *explicitly* represented. But this assumption is groundless. To see this, consider a device, D, that computes mental phrase markers by implementing the Earley or the CKY algorithm (§3.1), which draws on explicitly represented grammatical rules as data. Suppose that D is susceptible to a third-level theory, such that its parsing algorithm is likewise explicitly represented. As mentioned above, it is provable that one can construct a hardwired circuit, D*, that performs the same computations as D, without explicitly representing the algorithm. For present purposes, the important point is that constructing D*—i.e., "hardwiring" the original algorithm—involves *also* "hardwiring" *all of the data structures on which that algorithm drew*. Hence, the grammatical rules that were used as data in D would still be used as data in D*, though they would be *implicitly* represented—i.e., embodied, rather than encoded.

Conclusion

We saw in §2 that a range of behavioral and neurophysiological data supports the claim that the HSPM constructs mental phrase markers in the course of comprehension. In §3, we surveyed a number of psychologically plausible computational models of sentence processing. I argued that all such models—both classical and connectionist—draw on representations of the rules or principles of a grammar. Finally, in §4 I sketched a framework for thinking about these representations, arguing that they are best seen as subpersonal states, whose representational properties are determined by their functional role in a computational system. I distinguished between explicit and implicit representation and argued that Devitt (2006) is wrong to think that implicit representations cannot serve as data for an algorithm. I conclude that Devitt's skepticism concerning the psychological reality of grammars cannot be sustained.[49*]

[48] Devitt calls this "position (ii)" on the psychological reality issue.

[49*] I am indebted to a great many people for very helpful conversations on the topics addressed here, as well as for comments on previous drafts of this paper. Many thanks to Jake Berger, Jennifer Corns, Michael Devitt, Janet Dean Fodor, Jon Golin, Bob Matthews, Mike Nair-Collins, Georges Rey, David Rosenthal, Ed Stabler and all of the participants in the 2010 "Mental Phenomena" conference at the IUC in Dubrovnik, Croatia. Thanks also to Dunja Jutronić, both for organizing the conference and for her help and encouragement throughout the publication process.

References

Abney, S. P. and Johnson, M. (1991). "Memory Requirements and Local Ambiguities of Parsing Strategies," *Journal of Psycholinguistic Research*, 20, (3), 233–250.

Berwick, R. C. and Weinberg, A. S. (1984). *The Grammatical Basis of Linguistic Performance,* Cambridge: MIT Press.

Berwick, R. C. (1991a). "Principles of Principle-based Parsing," in *Principle-Based Parsing: Computation and Psycholinguistics*, Berwick, R. C., Abney, S. P., and Tenny, C. eds., Dordrecht: Kluwer Academic Publishers, 1–37.

Berwick, R. C. (1991b). "Principle-Based Parsing," in *Foundational Issues in Natural Language Processing,* P. Sells, S.M. Shieber, T. Wasow, eds., Dordrecht: Kluwer Academic Publishers, 115–226.

Berwick, R. C. (1997). "*Syntax Facit Saltum*: Computation and the Genotype and Phenotype of Language," *Journal of Neurolinguistics,* Vol. 10, No. 213, 231–249.

Berwick, R. C., Abney, S. P., and Tenny, C., eds. (1991). *Principle-Based Parsing: Computation and Psycholinguistics*. Dordrecht: Kluwer Academic Publishers.

Bever, T. G., Carroll, J. M., and Miller, L.A., eds. (1984). *Talking Minds: The Study of Language in Cognitive Science*. Cambridge: MIT Press.

Block, N. (1986). "Advertisement for a Semantics for Psychology," in French, P. A., *et. al.*, eds., *Midwest Studies in Philosophy, Vol. X*. Minneapolis: Univ. of Minnesota Press, 615–78.

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

Bornkessel-Schlesewsky, I. and Schlesewsky, M. (2009). *Processing Syntax and Morphology, A Neurocognitive Perspective*, Oxford University Press.

Brandom, R. (1998). *Making It Explicit*. Harvard University Press.

Brandom, R. (2008). *Between Saying and Doing*. Oxford University Press.

Bresnan, J. (2001). *Lexical-functional Syntax*, Oxford: Blackwell.

Carruthers, P. (2006). "The Case for Massively Modular Models of the Mind," in *Contemporary Debates in Cognitive Science*, R. Stainton (ed.), Malden: Blackwell, 3–21.

Cain, M.J. (2010). "Linguistics, Psychology and the Scientific Study of Language," *dialectica* Vol. 64, N° 3, 385–404.

Chomsky, N. (1957/2002). *Syntactic Structures*, 2nd edition. De Gruyter Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*, Cambridge: MIT Press.

Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.

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

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*, Cambridge: Cambridge University Press.

Collins, J. (2008). *Chomsky: A Guide for the Perplexed*. London: Continuum Press.

Coltheart, M. (1999). "Modularity and cognition." *Trends in Cognitive Sciences*, 3, 115–120.

Crocker, M. W., Pickering, M., and Clifton C., eds. (2000). *Architectures and Mechanisms for Language Processing*, Cambridge: Cambridge University Press.

Culbertson J. and Gross, S. (2009). "Are Linguists Better Subjects?" *British Journal for the Philosophy of Science*, 60, 721–736.

Cummins, R. (1991). *Meaning and Mental Representation*. Cambridge: MIT Press.

Cummins, R. (1996). *Representations, Targets, and Attitudes*. Cambridge: MIT Press.

Davies, M. (1987). "Tacit Knowledge and Semantic Theory: Can a Five per cent Difference Matter?" *Mind*, 96, 441–62.

Davies, M. (1989). "Tacit Knowledge and Subdoxastic States," in *Epistemology of Language*, George, A. (ed.), Oxford University Press, 131–52.

Davies, M. (1995). "Two Notions of Implicit Rules," *Philosophical Perspectives,* 9, *AI, Connectionism, and Philosophical Psychology,* Tomberlin, J. E. (ed.). Cambridge: Blackwell.

Dennett, D. C. (1987). *The Intentional Stance*, Cambridge: MIT Press.

DeVincenzi, M. (1991). "Filler-gap dependencies in a Null Subject Language: Referential and Non-Referential WHs," *Journal of Psycholinguistic Research*, vol. 20, no. 3, 197–213.

Devitt, M. (1997). *Realism and Truth* (2nd edition), Princeton: Princeton University Press.

Devitt, M. (2006). *Ignorance of Language*, Oxford: Clarendon Press.

Dretske, F. (1981). *Knowledge and the Flow of Information*, MIT Press: Cambridge, MA

Elman, J. L. (1992). "Grammatical Structures and Distributed Representations," in *Connectionism: Theory and Practice*, *Vol. 3,* Davis, S. (ed.), Oxford University Press.

Featherston, S. (2001). *Empty Categories in Sentence Processing*. Amsterdam: John Benjamins Publishing Company.

Fodor, J. A. (1968). "The Appeal to Tacit Knowledge in Psychological Explanation," *The Journal of Philosophy*, 65, (20), 627–640.

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

Fodor, J. A. (1987). *Psychosemantics*. Cambridge: MIT Press.

Fodor, J. A. (1990). *A Theory of Content, and Other Essays*. Cambridge: MIT Press.

Fodor, J. A. (2002). *The Mind Doesn't Work That Way*. Cambridge: MIT Press.

Fodor, J. A., Bever, T., and Garrett, M. (1974). *The Psychology of Language*, New York: McGraw Hill.

Fodor, J. D. (1989). "Empty Categories in Sentence Processing," in *Language and Cognitive Processes*, 4, 155–209.

Fodor, J. D. (1995). "Comprehending Sentence Structure," in *An Invitation to Cognitive Science: Volume 1, Language*, Gleitman and Liberman (eds.), Cambridge: MIT Press, 209–246.

Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*, PhD Dissertation, available at: http://digitalcommons.uconn.edu/dissertations/AAI7914150/

Frazier, L. and Clifton C. (1989). "Successive Cyclicity in the Grammar and the Parser," *Language and Cognitive Processes*, 4, (2), 93–126.

Frazier, L. and Fodor, J.D. (1978). "The Sausage Machine: A New Two-Stage Parsing Model," *Cognition,* Vol. 6, 291–325.

Friederici, A. D., Pfeifer, E., and Hahne, A. (1993). "Event-related brain potentials during natural speech processing: Effects of semantic, morphological, and syntactic violations," *Cognitive Brain Research*, 1, 183–92.

Friederici, A. D., Hahne, A., and Saddy, D. (2002). "Distinct Neurophysiological Patterns Reflecting Aspects of Syntactic Complexity and Syntactic Repair," *Journal of Psycholinguistic Research*, Vol. 31, No. 1.

Gerth, S. and beim Graben, P. (2009). "Unifying syntactic theory and sentence processing difficulty through a connectionist minimalist parser," *Cognitive Neurodynamics*, 3, 297–316.

Gibson, E. (1998). "Linguistic Complexity: Locality of Syntactic Dependencies," *Cognition*, 68, 1–76.

Haegeman, L. (1994). *Introduction to Government and Binding Theory*, 2nd ed., Oxford: Blackwell.

Hale, J. (1999). "Dynamical Parsing and Harmonic Grammar," unpublished report, available at the following URL: courses.cit.cornell.edu/jth99/prism.ps

Hale, J. (2001). "A Probabilistic Parser as a Psycholinguistic Model," in Proceedings of the NAACL. Available at: http://www.aclweb.org/anthology/N/N01/N01-1021.pdf

Hale, J. (2003). *Grammar, Uncertainty and Sentence Processing*, Ph.D. dissertation, Johns Hopkins University. Available at: http://web.jhu.edu/bin/s/q/Hale_dissertation.pdf

Harkema, H. (2001). *Parsing Minimalist Languages*, PhD Dissertation, UCLA. Available at: http://www.linguistics.ucla.edu/people/stabler/paris08/Harkema01.pdf

Horwich, P. (1998). *Meaning*. Oxford: Clarendon Press.

Johnson, M. (1989). "Parsing as Deduction: The Use of Knowledge of Language," *Journal of Psycholinguistic Research*, Vol. 18, 1, 105–128.

Jurafsky, D. (2003). "Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production," in *Probabilistic linguistics* Bod, Hay, & Jannedy, eds., MIT Press: Cambridge.

Jurafsky, D. and Martin, J. H. (2008). *Speech and Language Processing: An Introduction to Speech Recognition, Computational Linguistics, and Natural Language Processing (2nd edition)*, Prentice Hall.

Knowles, J. (2000). "Knowledge of Grammar As a Propositional Attitude." *Philosophical Psychology* 13(3), 325–53.

Lawrence, S., Giles, L., Fong, S. (2000). "Natural Language Grammatical Inference with Recurrent Neural Networks," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 12, No. 1, 126–140.

Lewis, D. K. (1972). "Psychophysical and Theoretical Identifications," *Australasian Journal of Philosophy*, 50, 249–258.

Marcus, M. P. (1984). "Some Inadequate Theories of Human Language Processing," in *Talking Minds: The Study of Language in Cognitive Science* Bever, T. G., Carroll, J. M., and Miller, L. A., eds., Cambridge: MIT Press, 253–278.

Matthews, R. J. (2007). *The Measure of Mind*. Oxford University Press.

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories: New Foundations for Realism*. Cambridge: MIT Press

Neander, K. (2006). "Naturalistic Theories of Reference," in *The Blackwell Guide to the Philosophy of Language*, Devitt, M. and Hanley, R. (eds.), Blackwell Publishing, 374–391

Neville, H.J. Nicol, J., Barss, A., Forster, K., and Garrett, M.F. (1991). "Syntactically based sentence processing classes: Evidence from event-related potentials," *Journal of Cognitive Neuroscience,* 6, 233–55.

Nicol, J. & Swinney, D. (1989). "The role of structure in coreference assignment during sentence comprehension," *Journal of Psycholinguistic Research,* 18, 5–19.

Peacocke, C. (1989) "When is Grammar Psychologically Real?" in *Reflections on Chomsky*, George, A. (ed.), Oxford: Basil Blackwell, 111– 130.

Pereira, F.C.N. and Warren, D.H.D. (1983). "Parsing as Deduction," *Proceedings of the 22nd Annual Meeting of the Association for Computational Linguistics*, 137–144.

Pereplyotchik, D. (forthcoming). *The Psychological Import of Syntactic Theory*, PhD Dissertation, City of New York (CUNY) Graduate Center.

Phillips, C. (1994). *Order and Structure*, unpublished PhD thesis, MIT.

Phillips, C. and Lewis, S. (forthcoming). "Derivational Order in Syntax: Evidence and Architectural Consequences," in *Directions in Derivations*, C. Chesi (ed.), Elsevier.

Pickering, M.J. and Ferreira, V.S. (2008). "Structural Priming: A Critical Review," *Psychological Bulletin*, Vol. 134, No. 3, 427–459.

Pietroski, P. (2008). "Think of the Children: Comments on Ignorance of Language, by Michael Devitt," *Australasian Journal of Philosophy*, 86, 657–69.

Pinker, S. (1999). *How the Mind Works*. W. W. Norton and Company: New York.

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

Prinz, J. (2006). "Is the Mind Really Modular?" in *Contemporary Debates in Cognitive Science*, R. Stainton (ed.), Malden, MA: Blackwell, 22–36.

Pritchett, B. (1992). *Grammatical Competence and Parsing Performance*, University of Chicago Press.

Putnam, H. (1988). *Representation and Reality*. Cambridge: MIT Press.

Rattan, G. (2002). "Tacit knowledge of grammar: a reply to Knowles." *Philosophical Psychology,* 15(2), 135–54.

Rayner, K., (1998). "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin*, Vol. 124, No. 3, pp. 372–422.

Rayner, K., Carlston, M., and Frazier, L. (1983). "The Interaction of Syntax and Semantics During Sentence Processing: Eye Movements in the Analysis of Semantically Biased Sentences," *Journal of Verbal Learning and Verbal Behavior*, 22, 358–374.

Rey, G. (2006). "Conventions, Intuitions and Linguistic Inexistents: A Reply to Devitt." *Croatian Journal of Philosophy*, 6, 549–69.

Rey, G. (2008). "In Defense of Folieism," *Croatian Journal of Philosophy*, 8, 23, 177–202.

Rohde, D.L.T. (2002). *A Connectionist Model of Sentence Comprehension and Production* Ph.D. Dissertation, Department of Computer Science, Carnegie Mellon University.

Rosenthal, D. M. (2005). *Consciousness and Mind*. Oxford: Clarendon Press.

Schank, R. and Birnbaum, L. (1984). "Memory, Meaning, and Syntax," in *Talking Minds: The Study of Language in Cognitive Science* Bever, T. G., Carroll, J.M., and Miller, L.A., eds., Cambridge: MIT Press, 209–251.

Sellars, W. (1963a). "Empiricism and the Philosophy of Mind," in *Science Perception and Reality*. Atascadero, CA: Ridgeview Publishing, 127–196.

Sellars, W. (1963b). "Some Reflections on Language Games," in *Science, Perception, and Reality*. Atascadero, CA: Ridgeview Publishing, 321–358.

Sellars, W. (1974). "Meaning as Functional Classification," *Synthese*, 27, 417–37.

Shieber, S. M., Schabes, Y., and Pereira, F. C. N. (1993). Principles and implementation of deductive parsing, Technical Report CRCT TR–11–94, Computer Science Department, Harvard University, Cambridge, Massachusetts. Available at: http://arXiv.org/.

Smolensky, P. and Legendre, G., eds. (2006). *The Harmonic Mind*. Cambridge: MIT Press.

Stabler, E. P., Jr. (1983). "How are Grammars Represented?" *Behavioral and Brain Sciences*, 6, 391–402.

Stevenson, S. (1999). "Bridging the Symbolic-Connectionist Gap in Language Comprehension." in Lepore, E., ed., *What is Cognitive Science?*, Oxford: Blackwell Publishers, 336–355.

Stevenson, S. and Smolensky, P. (2006). "Optimality in Sentence Processing," in *The Harmonic Mind*, Smolensky, P. and Legendre, G., eds., Cambridge: MIT Press, 307–337.

Stich, S. P. (1978). "Beliefs and Subdoxastic States," *Philosophy of Science*, 45, 499–518.

Sturt, P., Pickering, M. J., Scheepers, C. and Crocker, M.W. (2001). "The Preservation of Structure in Language Comprehension: Is Reanalysis the Last Resort?" *Journal of Memory and Language* 45, 283–307.

Tabor, W. and Tanenhaus, M. K. (1999). "Dynamical Models of Sentence Processing," *Cognitive Science*, 23 (4), 491–515.

Weinberg, A. (1999). "A Minimalist Theory of Human Sentence Processing," in Epstein and Hornstein (eds.), *Working Minimalism*, Cambridge: MIT Press, 283–315.

Yang, C. and Berwick, R. C. (1996). "Principle-Based Parsing for Chinese," in *Language, Information and Computation* (PACLIC 11), 363–371.