# 17 THE LANGUAGE OF THOUGHT

*Susan Schneider*

According to the language-of-thought (or LOT) hypothesis, conceptual thinking occurs in an internal language-like representational medium. However, this internal language is not equivalent to one's spoken language(s). Instead, LOT is supposed to be the format in which the mind represents concepts, rather than merely the natural-language words for the concepts themselves. The LOT hypothesis holds that the mind has numerous internal "words" (called "symbols") which combine into mental sentences according to the grammatical principles of the language. Conceptual thinking has a computational nature: thinking is the processing of these strings of mental symbols according to algorithms. The LOT program and the connectionist program are often viewed as competing theories of the format, or representational medium, of thought (see Sharkey and Sharkey, this volume; and Ryder, this volume).

Why believe in LOT? As we shall see, many of the most well received motivations arise from the following crucial and pervasive feature of conceptual thought: thought is essentially combinatorial. Consider the thoughts *the cappuccino in Florence is better than in New Brunswick* and *Surprisingly, Bush thought about Einstein*. You have probably not had any of these thoughts before, but you were able to understand these sentences. The key is that the thoughts are built out of familiar constituents, and combined according to rules. It is the *combinatorial* nature of thought that allows us to understand/produce these sentences on the basis of our antecedent knowledge of the grammar and atomic constituents (e.g., *Einstein*, *Italy*). Clearly, explaining the combinatorial nature of thought should be a central goal of any theory of the cognitive mind. For, as Gary Marcus puts it, "what is the mind such that it can entertain an infinity of thoughts?" (Marcus 2001: 1). LOT purports to be the only explanation for this important feature of thought.

In this overview of the LOT program, I shall begin by laying out its three central claims, as well as stressing the key philosophical issues which the LOT project is supposed to inform. I then discuss the main motivations for believing that there is a LOT. Finally, I close by exploring some "skeletons" in the LOT closet: relatively ignored issues that the success of LOT depends upon.

## What is LOT?

The idea that there is a LOT was developed by Jerry Fodor, who defended this hypothesis in an influential book, *The Language of Thought* (1975). As Fodor has emphasized, the LOT hypothesis was inspired by the ideas of Alan Turing, who defined computation in terms of the formal manipulation of uninterpreted symbols according to algorithms (Turing 1950; Fodor 1994). In his "Computing Machinery and Intelligence," Turing had introduced the idea that symbol-processing devices can think, a view which many in cognitive science are sympathetic to, yet which has also been the focus of great controversy (e.g., Searle 1980; Dreyfus 1992). The symbol-processing view of cognition was very much in the air during the time in which the LOT hypothesis was developed. Around the same time that *The Language of Thought* came out, Allen Newell and Herbert Simon suggested that psychological states could be understood in terms of an internal architecture that was like a digital computer (Newell and Simon 1972). Human psychological processes were said to consist of a system of discrete inner states (symbols) which are manipulated by a central processing unit (or CPU). Sensory states served as inputs to the system, providing the "data" for processing according to the rules, and motor operations served as outputs. This view, called "classicism," was the paradigm in the fields of artificial intelligence, computer science, and information-processing psychology until the 1980s, when the competing connectionist view also gained support. LOT, as a species of classicism, grew out of this general trend in information-processing psychology to see the mind as a symbol-processing device.

Now let us turn to a more detailed discussion of the LOT hypothesis. In essence, the LOT position consists in the following claims:

1. Cognitive processes consist in causal sequences of tokenings of internal representations in the brain.

This claim has enormous significance, for it provides at least a first approximation of an answer to the age old question, "how can rational thought ultimately be grounded in the brain?" At first pass, rational thought is a matter of the causal sequencing of tokens (i.e., patterns of matter and energy) of representations which are ultimately realized in the brain. Rational thought is thereby describable as a physical process, and further, as we shall see below, both a computational and semantic process as well.

In addition,

2. These internal representations have a combinatorial syntax and semantics, and further, the symbol manipulations preserve the semantic properties of the thoughts (Fodor 1975; Fodor and Pylyshyn 1988).

This claim has three components:

2a. Combinatorial syntax.

As noted, complex representations in LOT (e.g., #take the cat outside#) are built out of atomic symbols (e.g., #cat#, #outside#), together with the grammar of the LOT.

> 2b. Combinatorial semantics.

The meaning or content of a sentence in the LOT is a function of the meanings of the atomic symbols, together with their grammar.

> 2c. Thinking, as a species of symbol manipulation, preserves the semantic properties of the thoughts involved (Fodor 1975; Fodor and Pylyshyn 1988).

To better grasp (2c), consider the mental processing of an instance of *modus ponens*. The internal processing is purely syntactic; nonetheless, it respects semantic constraints. Given true premises, the application of the rule will result in further truths. The rules are *truth preserving*. John Haugeland employs the following motto to capture this phenomenon:

> Formalist motto – "If you take care of the syntax of a representational system, the semantics will take care of itself" (Haugeland 1989: 106).

And lastly,

> 3. Mental operations on internal representations are causally sensitive to the syntactic structure of the symbol (Fodor and Pylyshyn 1988).

Computational operations operate upon any symbol/symbol string satisfying a certain structural description, transforming it into another symbol/symbol string that satisfies another structural description. For example, consider an operation in which the system recognizes any operation of the form "(*P*&*Q*)" and transforms it into a symbol of the form "(*P*)." Further, the underlying physical structures on to which the symbol structures are mapped are the very properties that cause the system to behave in the way it does (Fodor and Pylyshyn 1988: 99). (It turns out that this feature of classical systems – that the constituents of mental representations are causally efficacious in computations – plays a significant role in the LOT-connectionism debate. For in contrast to symbolic systems, connectionist systems do not operate on mental representations in a manner that is sensitive to their form (for discussion see Fodor and Pylyshyn 1988; Macdonald 1995: Ch. 1; Marcus 2001: Ch. 4; Smolensky 1988, 1995).

(1)–(3) combine in a rather elegant way. For they generate a view which is closely related to the LOT hypothesis, called "the computational theory of mind" (or put simply, CTM). CTM holds the following:

> CTM – Thinking is a computational process involving the manipulation of semantically interpretable strings of symbols which are processed according to algorithms (Newell and Simon 1972; Fodor 1994; Pinker 1999; Rey 1997).

Stephen Pinker captures the gist of the manner in which (1)–(3) give rise to CTM:

> arrangements of matter … have both representational and causal properties, that is … [they] simultaneously carry information about something and take part in a chain of physical events. Those events make up a computation, because the machinery was crafted so that if the interpretation of the symbols that trigger the machine is a true statement, then the interpretation of the symbols created by the machine is also a true statement. The Computational Theory of Mind is the hypothesis that intelligence is computation in this sense. (Pinker 1999: 76)

This statement aptly connects the CTM hypothesis to the aforementioned age old question, "How can rational thought be grounded in the brain?" We've already noted that on the present view, rational thought is a matter of the causal sequencing of symbol tokens which are ultimately realized in the brain (thesis [1]). To this we add: these symbols, which are ultimately just patterns of matter and energy, have both representational (thesis [2b]) and causal properties (thesis [3]). Further, the semantics mirrors the syntax (thesis [2c]). This leaves us with the following picture of the nature of rational thought: thinking is a process of symbol manipulation in which the symbols have an appropriate syntax and semantics (roughly, natural interpretations in which the symbols systematically map to states in the world).

This account of the nature of rational thought has been summoned to solve an important puzzle about intentional phenomena. By "intentional phenomena" what is meant is a thought's "aboutness" or "directedness," that it represents the world as being a certain way. It has long been suspected that thought is somehow categorically distinct from the physical world, being outside the realm that science investigates. For how is it that a thought (e.g., the belief that the cat is outside, the desire to eat pizza), which, as we now know, arises from states of the brain, can be about, or directed at, something in the world? The LOT-CTM framework has been summoned to answer to this question. In essence, the proponent of LOT approaches this question in a "naturalistic" way, trying to ground intentionality in the world which science investigates. Now, we've already noted that symbols have a computational nature. As such, they are clearly part of the domain that science investigates. But the proponent of LOT has a naturalistic story about the aboutness, or intentionality, of symbols as well. Symbols refer to, or pick out, entities in the world, in virtue of their standing in a certain causal or nomic relationship that exists between the symbols and property tokens/individuals in the world. Simply put, the symbols are "locked on to" properties or individuals of a certain sort in virtue of standing in a certain nomic or causal relationship specified by a theory of meaning or mental content (for further discussion see Ryder, this volume). So the intentionality of a thought, e.g., *the espresso is strong*, is a matter of a causal, and ultimately physical, relationship between symbolic computational states and entities in the world (e.g., espresso).

This, then, is the gist of the LOT picture. At least at first blush, the LOT project seems to be a coherent naturalistic picture of the way the cognitive mind might be.

But, importantly, is it true? That is, is the cognitive mind really a symbol-manipulating device? Let us turn to the major reasons that one might have for suspecting that it is.

## The key arguments for LOT

The most important rationale for LOT derives from the following observation: any empirically adequate cognitive theory must hold that cognitive operations are sensitive to the constituent structure of complex sentence-like representations (Fodor 1975; Fodor and Pylyshyn 1988). This observation has been regarded as being strong evidence for a LOT architecture. To develop this matter in more detail, there are the following closely related features of cognition that seem to require that any theory of cognition appeal to structure-sensitive representations: productivity, systematicity, and inferential coherence.

### *The productivity of thought*

Consider the sentence, "The nearest star to Alpha Centauri is dying." As noted earlier, despite the fact that you've never heard a novel thought before, you are capable of understanding it. Thought is productive: in principle, you can entertain and produce an infinite number of distinct representations. How can you do this? Our brains have a limited storage capacity, so it cannot be that we possess a mental phrase book in which the meaning of each sentence is encoded. Instead, there must be a system with a combinatorial syntax. This allows for the construction of potentially infinitely many thoughts given a finite stock of primitive expressions (Fodor 1975: 31; Fodor and Pylyshyn 1988: 116; Fodor 1985, 1987).

### *The systematicity of thought*

A representational system is systematic when the ability of the system to entertain/produce certain representations is intrinsically related to the ability to entertain/produce other representations (Fodor and Pylyshyn 1988). Conceptual thought seems to be systematic; e.g., one doesn't find normal adult speakers who understand, "Mary loves John" without also being able to produce/understand "John loves Mary." How can this fact be explained? Intuitively, "Mary loves John" is systematically related to "John loves Mary" because they have a common constituent structure. Once one knows how to generate a particular sentence out of primitive expressions, one can also generate many others that have the same primitives (Fodor 1987; Fodor and Pylyshyn 1988; Fodor and McLaughlin 1990).

### *Inferential coherence*

As Fodor and Pylyshyn have observed, we do not encounter normal human minds which are always prepared to infer from *P&Q&R* to *P* but not infer from *P&Q* to

*P* (1995: 129). Thought is inferentially coherent: given that a system can draw a particular inference that is an instance of a certain logical rule, the system can draw any inferences that are instances of the rule. And again, this has to be due to the fact that mental operations on representations are sensitive to their form (Fodor and Pylyshyn 1988).

In sum, these three features of thought all seem to arise from the fact that mental representations have constituent structure. As noted, they have been regarded as providing significant motivation for LOT. It is currently a source of great controversy whether connectionist systems can explain these important features of thought (see, e.g., Calvo and Colunga 2003; Fodor and Pylyshyn 1988; Fodor and McLaughlin 1990; Elman 1998; van Gelder 1990; Marcus 2001; Smolensky 1988, 1995). Connectionist models are networks of simple parallel computing elements, with each element carrying a numerical activation value which the network computes given the values of neighboring elements, or units, in the network, employing a formula (see infra, Sharkey and Sharkey, this volume). In very broad strokes, critics claim that a holistic pattern of activation doesn't seem to have the needed internal structure to account for these features of thought. Critics have argued that, at best, connectionist systems would provide models of how symbol structures are implemented in the brain, and would not really represent genuine alternatives to the LOT picture (Fodor and Pylyshyn 1988). There is currently a lively debate between this "implementationalist" position and radical connectionism, a position which rejects the view that connectionism, at best, merely implements LOT, advancing connectionism as a genuine alternative to the LOT hypothesis.

In addition to arguments for LOT based on the combinatorial structure of thought, the following two arguments are well-known arguments as well.

Fodor advances the first argument as the central argument of his 1975 book. The rough argument is as follows: ($P_1$) The only plausible psychological models of decision making, concept learning and perception all treat mental processes as computational. ($P_2$) Computation presupposes a medium of computation – a representational system in which computations are carried out. ($P_3$) Remotely plausible theories are better than nothing. (C) Therefore, we must take seriously the view that the mind has a LOT (Fodor 1975: 27). Much of Fodor's defense of the argument is devoted to exploring the basic form of information-processing models of learning, decision-making, and perception (Fodor 1975).

It is important to bear in mind that the argument, which dates back to 1975, preceded the rise in popularity of connectionism. LOT is no longer "the only game in town" (as Fodor used to boast) (Fodor 1975). While the view that contemporary cognitive science is computational is still very well received, nowadays, a computationalist need not be a classicist; she can be a connectionist instead. These issues are subtle: As mentioned, "implementational connectionists" actually believe in LOT, holding that connectionist networks merely implement LOT. It is likely that they would agree with something like the above argument. Radical connectionists, by contrast, would likely object that the conclusion does not follow from the premises;

($P_1$) and ($P_2$) are compatible with connectionism as well. Whether the connectionist response is effective depends upon nuances of the LOT-connectionism debate which we cannot delve into here (for a helpful introduction to these issues see Macdonald and Macdonald [1995]). Suffice it to say that the proponent of LOT, armed with arguments along the lines of (1)–(3), would likely charge that any connectionist model of psychological phenomena that purports to be a genuine alternative to (rather than mere implementation of) LOT will not satisfy the demands of these arguments.

A second challenge to this argument was raised, ironically, by Fodor himself, who, after publishing *The Language of Thought*, has expressed doubts about the plausibility of computational explanation of decision-making, and conceptual thought more generally, and has offered arguments which can be viewed as an attack on $P_1$ (Fodor 2000). This important issue will be discussed below.

Finally, a fifth argument for LOT is Fodor's well-known argument for nativism (Fodor 1975, 1981). Because Fodor's (1975) emphasized this argument, and because Fodor himself has been associated with extreme concept nativism, extreme concept nativism has become unduly wedded to the LOT program. Indeed, many assume that if there's a LOT, then vocabulary items in the LOT must be innate. But notice that nativism is not entailed by theses (1)–(3); nor is it invoked in any of the aforementioned motivations for LOT.

In very broad strokes, Fodor's nativist argument for LOT is along the following lines. Since concept learning is a form of hypothesis formation and confirmation, it requires a system of mental representations in which the formation and confirmation of hypotheses is to be carried out. But then one must already possess the concepts in one's LOT in which the hypotheses are couched. So we must already have the innate symbolic resources to express the concepts being learned (Fodor 1975: 79–97, 1981).

The above argument, as it stands, is open to the possibility that many lexical concepts are constructed from more basic, unstructured, concepts. These lexical concepts can be learned concepts because they have internal structure, being assembled from more basic, innate concepts. These lexical concepts are thus not innate. So, strictly speaking, the above argument does not entail the extreme concept nativism associated with Fodor's project. However, Fodor famously rejects the view that lexical concepts are structured, arguing in his (1981) that lexical concepts do not have internal structure, as the leading theories of conceptual structure are highly problematic. If Fodor is correct, we are left with a huge stock of lexical primitives (Fodor 1981). And, according to Fodor, primitive concepts are innate. If this is correct, then the above rough argument presents a case for radical concept nativism.

Critics and proponents of LOT uniformly reject radical concept nativism (including Fodor himself in his [1998]). After all, it is hard to see how concepts that our evolutionary ancestors had no need for, such as [carburetor] and [photon], could be innate. Of course, proponents of LOT generally believe that LOT will turn out to have some empirically motivated nativist commitments invoking both certain innate modules and primitive symbols. However, it is important that LOT be able to accommodate any well-grounded empirically based view of the nature of concepts that cognitive science develops, even one in which few or no concepts are innate. Nonetheless, Fodor's argument and concerns about conceptual structure are intriguing, for they raise some very important questions: What is wrong with the argument? Can primitive (unstructured) concepts be learned? Are many lexical concepts structured?

While I've stressed that LOT shouldn't require the truth of radical concept nativism, it should be mentioned that there is a nativist commitment that seems reasonable to wed to the LOT program. LOT can be regarded as an innate cognitive capacity, because, according to the proponent of LOT, any sophisticated language-like computational system requires an internal language that has primitive vocabulary items obeying rules enabling the language to be systematic, productive, and compositional. But this sort of nativism is distinguishable from concept nativism; for this innate capacity can exist, while the stock of symbols in each person's inner vocabulary may differ. In such a scenario, we each have a cognitive system which satisfies (1)–(3), but some, or even all, of the primitive vocabulary items differ.

## Some important qualifications

Needless to say, with such a bold view of the nature of thought, numerous qualifications are in order. First caveat: I have thus far said nothing about the nature of consciousness. Even philosophers who are sympathetic to computational accounts of the mind suspect that computational theories may fall short as explanations of the essential nature of consciousness (Block 1991; Chalmers 1995). LOT does not aspire to be a theory of consciousness or to answer the hard problem of consciousness; instead, it is a theory of the nature of language-like mental processing that underlies higher cognitive function, and more specifically it is designed to account for the aforementioned combinatorial features of thought, issues which are, of course, important in their own right.

Indeed, it is important to bear in mind that the scope of the LOT hypothesis is itself a matter of significant controversy. LOT is not primarily concerned with the nature of mental phenomena such as perceptual pattern recognition, mental imagery, sensation, visual imagination, dreaming, hallucination, and so on. While a LOT theorist may hold views that explain such phenomena by something similar to LOT, it is likely that even if LOT is correct, it does not apply to all the above domains. Indeed, it may turn out that certain connectionist models better explain some of these phenomena (e.g., pattern recognition) while the symbol-processing view offers a superior account of cognition. Such a "hybrid" view is sympathetic to a connectionist picture of sensory processes, while claiming that when it comes to explaining conceptual thought, the symbol-processing account is required (Wermter and Sun 2000). Fodor himself rejects hybrid models, suggesting instead that modular input systems have their own LOT (1983); they do not have a full-blown LOT, but for Fodor, it is a matter of degree.

Second qualification: although the LOT hypothesis holds that the mind is computational, this view should not be conflated with the view that the mind is like a commercially available computer, having a CPU in which nearly every operation is executed. Although symbolicists in the 1970s and '80s seem to have construed classicism in this way, this view is outdated. As Stephen Pinker notes, LOT is implausible when it is aligned with the view that the mind has a CPU in which every operation is executed in a serial fashion (Pinker 2005). Although introspectively our thoughts seem to be sequential, introspection only reveals a portion of the workings of the mind; it is uncontroversial that the brain has multiple non-conscious processes that operate in a massively parallel manner. Classicism and LOT merely require the weaker view that the brain has a "central system." On Fodor's view, the central system is a non-modular subsystem in the brain in which information from the various sense modalities is integrated, deliberation occurs, and behavior is planned (Fodor 1983). Crucially, a central system need not be a CPU; for it is not the case that every operation needs to be executed by a central system, as it does with a CPU. Instead, the central system may only be involved in higher cognitive tasks, e.g., planning, deliberation, categorization, not in mental operations that do not involve consciousness or reasoning.

Indeed, it is well worth getting clear on the nature of the central systems. For, as we shall see below, when we consider some issues requiring further development by the proponent of LOT, *inter alia*, the proponent of LOT seems to owe us a positive account of the nature of the central systems. Let us now turn to these outstanding issues.

## Looking ahead: issues in need of future development

Back in 1975 Fodor has noted that characterizing the LOT, "is a good part of what a theory of mind needs to do" (Fodor 1975: 33). Unfortunately, even today, certain key features of the LOT program remain unexplained. Herein, I shall consider two important problems that threaten the success of the LOT program. The first issue concerns the notion of a symbol in LOT. While the notion of a symbol is clearly key to the LOT program, unfortunately, the program lacks a well-conceived notion of the symbolic mental states that are supposed to be the very basis of cognition. Second, as noted, Fodor himself has expressed doubts about the plausibility of computational explanation. More specifically, he suspects that the central systems will defy computational explanation and has offered two arguments in support of this pessimistic view (Fodor 2000). It is rather important whether these two arguments are correct; if Fodor is correct, then we should surely reject the LOT hypothesis. For, as noted, the central system is supposed to be the system in which deliberation and planning occur. So it is reasonable to regard it as the primary domain which LOT characterizes. But LOT is obviously a computational theory, so how can it correctly characterize the central systems if they are not, in fact, computational to begin with? Further, if LOT fails to characterize the central systems, it is difficult to see why we should even believe that it applies to the modules.

### *Symbols*

Let us first consider where LOT stands concerning the nature of mental symbols. To provide a theory of the nature of symbols, one needs to locate features of symbols according to which the symbols should be taxonomized, or classified. For instance, should two symbol tokens be regarded as being of the same type when they have the same semantic content? Or perhaps, instead, symbols should be type-individuated by computational properties, such as computational roles? If so, what properties or roles? For the proponent of LOT the stakes are high: without a plausible theory of primitive symbols, there cannot be a complete understanding of what the LOT hypothesis is supposed to be. After all, without a theory of symbol natures, it remains unclear how patterns of neural activity could be, at some higher level of abstraction, accurately described as being symbol manipulations. For what is it that is being manipulated? Further, without an adequate theory of symbol natures, related philosophical projects that draw from the LOT approach are undermined. First, the aforementioned attempt to naturalize intentionality will be weakened, for such accounts will lack an account of the nature of the internal mental states that are appealed to as the computational basis of intentionality. For according to the proponent of LOT, these mental states are the symbols themselves. Second, as noted, those who are interested in LOT frequently say that meaning is determined by some sort of external relation between symbols and properties or individuals in the world. Unfortunately, since symbols are the internal mental states, or "vehicles" that the meanings lock on to, such theories of mental content will be radically incomplete.

Existing theories of the nature of symbols include individuation by (externalist) semantic content and individuation by the role that the symbol plays in the computational system, where the notion of "computational role" is fleshed out in various ways. Concerning semantic proposals, it has been objected that a semantic manner of typing LOT expressions ruins the prospects for naturalism (Pessin 1995). For the externalist hopes to naturalize intentionality by taking the intentionality of thought to be a matter of a symbol bearing some sort of external relationship (e.g., historical, informational) to a property or thing in the world. But if the intentionality of thought is supposed to reduce to a physical relation between the symbol and the world, the symbol itself cannot be typed semantically. For this is an intentional phenomenon, and in this case, the intentionality of thought couldn't reduce to the physical (Pessin 1995).

Computational-role proposals also seem to be problematic. The "computational role" of a symbol is the role that the symbol plays in computation. As mentioned, there are different ways that computational role can be construed. Proposals can be distinguished by whether they consider all, or merely some, elements of a given symbol's role as being built into the nature of a symbol. A "molecularist" claims that in defining the nature of a symbol, only a certain privileged few computational relations are required for a given symbol to be of a certain type. To consider a tinker toy example, a molecularist view could hold that to have a token of the symbol, [cat], the system in question must have thoughts such as [furry] and [feline], but the system need not have others, e.g., [likes cat treats], [black cats are timid]. The advantage of molecularism is that because only some elements of a symbol's computational role constitute the symbol's nature, many individuals will have common symbols, so groups of individuals can figure in psychological explanations in virtue of the symbols they have. For there will be equivalence classes of systems which, when tokening a given symbol and in common conditions, will behave in similar ways.

Although this would surely be a virtue of the molecularist theory, many would say that molecularism faces insurmountable obstacles. For consider the related molecularist theories of narrow content. In the context of debates over the nature of mental content, molecularist views attempted to identify certain conceptual or inferential roles as being constitutive of narrow content. Such views were criticized because, according to the critics, there is no principled way to distinguish those elements of conceptual or inferential role that are meaning constitutive from those which are not (Fodor and LePore 1992; Segal 1999; Prinz 2002). Unfortunately, similar issues seem to emerge for molecularism about symbol types, although the issues do not concern meaning; instead, the issue concerns whether there can be a select few symbol-constitutive computational relations (Aydede 2000; Schneider n.d.). A natural reaction is to embrace the view that all of the computational relations individuate the symbol. But if a symbolic state is individuated by all the computational relations it participates in, a natural concern is that symbolic states will not be shared from person to person (Aydede 2000; Prinz 2002; Schneider n.d.).

In sum, the nature of symbols is very much an open question.

### *The computational nature of the central systems*

A second major challenge to the LOT approach stems, ironically, from Fodor's aforementioned view that the cognitive mind is likely to *not* be computational. His first argument involves what he calls "global properties," features that a sentence in the LOT have which depend on how the sentence interacts with a larger plan (i.e., set of LOT sentences), rather than merely depending upon the nature of the LOT sentence itself. For example, the addition of a new LOT sentence to an existing plan can complicate (or alternately, simplify) a plan. Since the added simplicity/complexity varies according to the context, that is, according to the nature of the plan the new sentence is added to, simplicity/complexity seems to be a global property of the mental sentence (Fodor 2000). Global properties, according to Fodor, give rise to the following problem for CTM:

> The thought that there will be no wind tomorrow significantly complicates your arrangements if you had intended to sail to Chicago, but not if your plan was to fly, drive or walk there. But, of course, the syntax of the mental representation that expresses the thought #no wind tomorrow# is the same whichever plan you add it to. The long and short is: the complexity of a thought is not intrinsic; it depends on the context. But the syntax of a representation is one of its essential properties and so doesn't change when the representation is transported from one context to another. So how could the simplicity of a thought supervene on its syntax? As please recall, CTM requires it to do. (2000: 26)

In a bit more detail, Fodor's argument is the following: cognition is sensitive to global properties. But CTM holds that cognition, being computational, is only sensitive to the "syntax" of mental representations. That is to say that cognition is sensitive to the type identity of the primitive symbols, the way the symbols are strung together into well-formed sentences, and the algorithms that the brain computes. And these "syntactic" properties are *context insensitive* properties of a mental representation. That is, what a mental representation's syntactic properties are does not depend on what the other mental representations in a plan are: it depends only on the type identity of the LOT sentence. But whether a given mental representation has the global properties that it has will typically depend upon the *context* of the other representations in a plan (that is, it depends upon the nature of the other LOT sentences in the relevant group, as in Fodor's example involving it being windy). So it seems that cognition then cannot be wholly explained in terms of computations defined over syntactic properties (Fodor 2000; Ludwig and Schneider 2008; Schneider 2007).

The second problem concerns what has been called, "The Relevance Problem." According to Fodor, this is the problem of whether and how humans determine what is relevant in a computational manner. Fodor suspects that if one wanted to get a machine to determine what is relevant, the machine would need to walk through virtually every item in its database, to see whether a given item is relevant or not. This is a huge computational task, and it could not be accomplished quickly enough for a system to act in real time. However, humans make quick decisions about relevance all the time. Hence, it looks like human domain general thought (i.e., the processing of the central systems) is not computational (Fodor 2000).

Elsewhere, Kirk Ludwig and I have argued that the problem that Fodor believes global properties pose for CTM is a non-problem (Ludwig and Schneider 2008; Schneider 2007). And concerning the relevance problem, elsewhere I've argued that while the relevance problem is a serious research issue, it does not justify the overly pessimistic view that cognitive science, and CTM in particular, will likely fail to explain cognition (Schneider 2007). Although we do not have time to consider all of these issues, I will quickly raise one problem with each of Fodor's two concerns. Both problems rely on a common example. Before entertaining this example, let us try to answer an important question: suppose that both problems can exist in the context of uncontroversially computational processes. What would this fact show? The following answer seems plausible: It would mean that the presence of a globality or relevance problem does *not* entail that the system in question is non-computational.

Now, bearing this in mind, notice that each of Fodor's arguments maintains that, as a result of the given problem, the central systems are non-computational. However, I shall now proceed to show that both problems exist in uncontroversially computational systems. We are now ready to consider our example. Consider a chess-playing program. Suppose that a human opponent makes the first move of the game, moving a certain pawn one square forward. Now, the program needs to decide, given the information of what the previous move was, which future move to execute. Even in an uncontroversially computational system like this one, we can quickly see that Fodor's globality problem emerges. Let us suppose that there are two game strategies/plans in the program's database and the program needs to select one plan, given information about what the first move is. Let one plan involve taking the bishop out early in the game, while the other plan involves taking the rook out early in the game (where "early" means, say, within four turns). Now, it is important to notice that the impact that the addition of the information about what the opponent's first move was on the simplicity of each of the two plans does not supervene on the type identity of the string of symbols that encodes the information about the opponent's first move. Instead, the impact of the addition of the string of symbols on the simplicity of each plan depends on the way that the string interacts with the other sentences in the plan. Thus (our new globality argument continues), the processing of the chess program is not syntactic, and thus, not computational. Hence, it appears that a globality problem emerges in the context of highly domain-specific computing (Schneider 2007).

Using this same simple example, we can also quickly see that a relevance problem emerges. Notice that skillful chess playing involves being able to select a move based on the projected outcome of the move as far into the future of the game as possible. So chess programmers deal with a massive combinatorial explosion all the time, and in order to quickly determine the best move, clever heuristics must be employed. This is precisely the issue of locating algorithms that best allow for the quick selection of a future move from the greatest possible projection of potential future configurations of the board (Marsland and Schaeffer 1990). And this is just the relevance problem as it has been articulated by Fodor and other philosophers (Schneider 2007). The upshot: both problems emerge at the level of relatively simple, modular, and uncontroversially computational processes. But if both problems can occur in the context of uncontroversially computational processes, the presence of a globality or relevance problem does not entail the conclusion that the relevant system is non-computational. And this is the conclusion that was needed to undermine the possibility that the central systems are computational.

Perhaps Fodor could say that the relevance problem, as it presents itself to the central systems, is somehow different. And moreover, it is different in a way that suggests that relevance determination in the central systems is non-computational. An obvious point of difference is that unlike modular processing, central processing is supposed to be domain general. However, this point of difference doesn't seem to warrant the extreme view that the processes in question would be non-computational. For one thing, there are already programs that carry out domain-general searches over immense databases. For consider your own routine Google searches. In about 200 milliseconds you can receive an answer to a search query involving two apparently unrelated words that involved searching a database of over a billion Web pages (Schneider 2007). Second, Fodor's relevance problem concerned how the brain could sift through massive amounts of data given the constraints of real time, and domain generality entails nothing about the size of a database that a relevance search draws from. A database that records the mass of every mass-bearing particle in the universe would be topic specific, yet still be of a much greater size than a human's memory (Schneider 2007).

Now, in contrast to the globality problem, which I suspect is merely a non-problem (Ludwig and Schneider 2008), the relevance program does present a challenge to programmers. The challenge for programmers is to find judicious algorithms which maximize the amount of information subject to the constraints of real time. However, if my above argument concerning relevance is correct, it is implausible to claim that a relevance problem entails that the system in question is non-computational. Yet it is natural to ask whether there are better ways of formulating the problem that relevance presents for CTM. Elsewhere, I discuss and rule out different formulations (Schneider 2007). But for now, let me suggest that a very different way to proceed with respect to the relevance problem is to assume that the presence of a human relevance problem is not terribly different from relevance problems existing for other *computational* systems. But, in the human case, the "solution" is a matter of empirical investigation of the underlying brain mechanisms involving human searches. This alternative approach assumes that evolution has provided *Homo sapiens* with algorithms that enable quick determination of what is relevant, and further, it is the job of cognitive science to discover the algorithms. On this view, Fodor's injunction that research in cognitive science rest at the modules must be resisted (Fodor 2000; Schneider 2007). Proponents of LOT should instead seek to provide detail concerning the nature of the central systems, in order to understand the nature of symbolic processing, including, especially, what the algorithms are that symbolic systems compute. An additional bonus of this more optimistic approach is that locating a computational account of the central systems could help solve the problem of symbol individuation, for once algorithms that the central systems compute are well understood, it is possible that they can be summoned to individuate symbols by their computational roles in the central systems.

## Conclusion

Well, where does all this leave us? I still have not answered the question I posed earlier, "Is LOT true?" But doing so would be premature: for cognitive science is now only in its infancy. As cognitive science develops, we will learn more and more about the various representational formats in the brain, sharpening our sense of whether LOT is a realistic theory and how the different representational formats of the brain interrelate. And, in the course of our investigations, it is likely that new and intriguing issues will come to the fore. In addition, we've canvassed a number of existing controversies still awaiting resolution. *Inter alia*, we've noted that many individuals are drawn to the symbol-processing program because it provides insight into the mind's combinatorial nature. But, as discussed, LOT is no longer the "only game in town," and it still remains to be seen whether connectionist models will be capable of explaining the combinatorial nature of cognition in a way that supports a genuine alternative to LOT. We've also discussed two other pressing issues which currently require resolution: first, the LOT/symbol-processing view requires a plausible account of the nature of symbols; and second, as discussed, there are the well-known worries about the limits of computational explanation of the cognitive mind which were posed by Fodor himself.

It is also worth mentioning that our discussion of presently known issues requiring more development is not intended to be exhaustive. Also of key import, for instance, are issues involving the modularity of mind and the nature of mental content (these issues are canvassed in the present volume, Chapters 17 and 18). For instance, any proponent of LOT would be interested in Peter Carruthers' recent book on modularity, which has recently developed the LOT approach within a modularist view of the central systems (Carruthers 2006). And intriguingly, Jesse Prinz has recently challenged the very idea that the mind is modular (Prinz 2006). And, as mentioned, the proponent of LOT will need a plausible theory of mental content in order to provide a complete account of the nature of intentionality. Yet the debate over mental content rages on. In sum, thirty years after the publication of Fodor's seminal book, *The Language of Thought*, there are still many areas to investigate. So in lieu of a firm answer to our question, we can at least acknowledge the following: the LOT program, while no longer the only game in town, is an important and intriguing proposal concerning the nature of conceptual thought.

## Acknowledgements

Thanks very much to Mark Bickhard, Paco Calvo and John Symons for helpful comments and suggestions on an earlier draft of this essay.

## References

Aydede, Murat (2000) "On the Type/Token Relation of Mental Representations," *Facta Philosophica* 7 (March): 23–49.

Block, Ned (1991) "Troubles with Functionalism," in David M. Rosenthal (ed.), *The Nature of Mind*, Oxford: Oxford University Press, chapter 23, pp. 211–28.

Calvo, Paco, and Colunga, Eliana (2003) "The Statistical Brain: Reply to Marcus' *The Algebraic Mind*," *Proceedings of the American Twenty-Fifth Annual Conference of the Cognitive Science Society*, London: Routledge, pp. 210–15.

Carruthers, Peter (2006) *The Architecture of the Mind*, Oxford: Oxford University Press.

Chalmers, David (1995) "Facing Up to the Hard Problem of Consciousness," *Journal of Consciousness Studies* 2, no. 3: 200–19.

Dreyfus, Hubert (1992) *What Computers Can't Do: A Critique of Artificial Reason*, New York: Harper.

Elman, Jeffrey (1998) "Generalization, Simple Recurrent Networks, and the Emergence of Structure," in M. A. Gernsbacher and S. Derry (eds), *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Erlbaum.

Fodor, Jerry A. (1981) "The Present Status of the Innateness Controversy," in J. A. Fodor (ed.), *Representations: Philosophical Essays on the Foundations of Cognitive Science*, Brighton, UK: Harvester Press, pp. 257–316.

—— (1983) *The Modularity of Mind*, Cambridge, MA: MIT Press.

—— (1985) "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum," *Mind* 94: 76–100; also in J.A. Fodor (1990).

—— (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, MA: MIT Press.

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

—— (1994) *The Elm and the Expert*, Cambridge, MA: MIT Press.

—— (1998) *Concepts: Where Cognitive Science Went Wrong,* Oxford: Oxford University Press.

—— (2000) *The Mind Doesn't Work That Way*, Cambridge, MA: MIT Press.

Fodor, Jerry A., and LePore, E. (1992) *Holism: A Shopper's Guide*, Oxford: Blackwell Publishing.

Fodor, Jerry A., and McLaughlin, B. (1990) "Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work," *Cognition* 35: 183–204.

Fodor, Jerry A., and Pylyshyn, Zenon W. (1988) "Connectionism and Cognitive Architecture: A Critical Analysis," in S. Pinker and J. Mehler (eds), *Connections and Symbols* (*Cognition* special issue), Cambridge, MA: MIT Press; also in Macdonald and Macdonald (1995).
Haugeland, J. (1989) *Artificial Intelligence: The Very Idea,* Boston: MIT Press.
Ludwig, Kirk, and Schneider, Susan (2008) "Fodor's Challenge to the Classical Computational Theory of Mind," *Mind and Language*, 23: 123–43.
Macdonald, Cynthia, and Macdonald, Graham (1995) *Connectionism: Debates on Psychological Explanation*, vol. 2, Oxford, Basil Blackwell.
Marcus, Gary (2001) *The Algebraic Mind*, Boston, MIT Press.
Marsland, T. Anthony, and Schaeffer, Jonathan (ed.) (1990) *Computers, Chess, and Cognition*, New York: Springer-Verlag.
Newell, Allen (1980) "Physical Symbol Systems," *Cognitive Science* 4: 135–83.
Newell, A., and Simon, H. A. (1972) *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.
Pessin, A. (1995) "Mentalese Syntax: Between a Rock and Two Hard Places," *Philosophical Studies* 78: 33–53.
Pinker, S. (1999) *How the Mind Works*, New York: W. W. Norton.
—— (2005) "So How *Does* the Mind Work?" (Review and reply to Jerry Fodor [2000]), *Mind and Language* 20: 1–24.
Prinz, J. (2002) *Furnishing the Mind: Concepts and Their Perceptual Basis*, Cambridge, MA: MIT Press.
—— 2006. "Is the Mind Really Modular?" In Bob Stainton (ed.), *Contemporary Debates in Cognitive Science*, New York: Blackwell.
Pylyshyn, Z. (1986) *Computation and Cognition*, London: MIT Press.
Rey, Georges (1997) *Contemporary Philosophy of Mind*, New York: Blackwell.
Schneider, S. (2005) "Direct Reference, Psychological Explanation, and Frege Cases," *Mind and Language* 20, no. 4: 423–47.
—— (2007) "Yes, It Does: A Diatribe on Jerry Fodor's *The Mind Doesn't Work That Way*," *Psyche* 13, no. 1: 1–15.
—— (n.d.) "The Nature of Primitive Symbols in the Language of Thought: A Theory," unpublished manuscript.
Searle, John R. (1980) "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, no. 3: 417–24.
Segal, Gabriel (1999) *A Slim Book on Narrow Content*, Cambridge, MA: MIT Press.
Smolensky, Paul (1988) "On the Proper Treatment of Connectionism," *Behavioral and Brain Sciences*, 11: 1–23.
—— (1995) "Reply: Constituent Structure and Explanation in an Integrated Connectionist/Symbolic Cognitive Architecture," In Macdonald and Macdonald (1995).
Turing, A. M. (1950) "Computing Machinery and Intelligence," *Mind* 59: 433–60.
van Gelder, Tim (1990) "Why Distributed Representation Is Inherently Non-Symbolic," in G. Dorffner (ed.), *Konnektionismus in Artificial Intelligence und Kognitionsforschung*, Berlin: Springer-Verlag, pp. 58–66.
Wermter, Stephan, and Sun, Ron (2000) "An Overview of Hybrid Neural Systems," in Stephan Wermter and Ron Sun (eds), *Hybrid Neural Systems*, Heidelberg: Springer, pp. 1–13.