Philosophy and Phenomenological Research
Vol. XLIX, No. 4, June 1989

Does Semantics Run the Psyche?

RADU J. BOGDAN
Tulane University

I. Introduction

If there is a dogma in the contemporary philosophy of the cognitive mind, it must be the notion that cognition is semantic causation or, differently put, that it is semantics that runs the psyche. This is what the notion of psychosemantics and (often) intentionality are all about. Another dogma, less widespread than the first but almost equally potent, is that common sense psychology is the implicit theory of psychosemantics. The two dogmas are jointly encapsulated in the following axiom. Mental attitudes such as beliefs and desires have essentially semantic contents, or are semantically evaluable. (This is why they are called *propositional* attitudes.) Mental attitudes have causal powers in virtue of their semantic properties. The content of an attitude has causal powers *qua* semantic, or more exactly in virtue of its syntactic structure which reflects relevant semantic properties and relations. (Propositions attitudinized cause in virtue of their semantically sensitive syntax.) It is the fact that mental attitudes cause in virtue of being semantic that explains why the cognitive mind is essentially semantic and why common sense psychology is implicitly true of the semantic mind.

Nobody has made and exploited this case better than Jerry Fodor. His latest book, *Psychosemantics*,[1] further develops and fortifies the two dogmas with a rich supply of bold, intricate and imaginative arguments. I have questioned elsewhere the second dogma to the effect that common sense psychology is an implicit theory of the semantic mind.[2] Here I want to wrestle with the first, which is about the semantic mind itself. The ques-

[1] The MIT Press/Bradford Books, Cambridge, MA, 1987.

[2] See my "Fodor's Representations," *Cognition and Brain Theory* 6 (1983): 237-49; and "Mental Attitudes and Common Sense Psychology: The Case Against Elimination," *Noûs* 22 (1988): 369-98.

tions I want to ask are: Is a cognitive state causal in virtue of having semantic properties? Is semantic information (structurally, not quantitatively) *information enough* to causally run cognition and behavior? Does semantics in general explain mental causation?

Very often, and certainly in most of the dogmatically standard examples, Fodor's included, the answers to these questions are positive. The mind often runs on semantic fuel only. If I think that P, and also that P entails Q, then I causally come to think that Q, because the former two thoughts cause the latter. If, at the moment, this is all I am thinking about, then the causation which drives my thinking not only respects, but is essentially shaped by, the semantic properties and relations involved. This is semantic causation. If *all* my cognition went like this, then my mental states would indeed cause in virtue of their semantic character, essentially. If Fodor were right about this and his semantic story of the mind definitive, the study of cognition would be neatly regimented and simplified, although cognizing itself would not be much fun any longer.

I see, however, two related problems with Fodor's psychosemanticism. Their discussion, inevitably incomplete and superficial, will take up most of this essay. One problem is whether cognition is essentially semantic. The other is whether the naturalization of the psychosemantics (which is a major novelty of Fodor's book) can be causal. The problems are related because, if Fodor's solution to the second problem is right, then his solution to the first may also turn out right. (We can't let that happen, can we, Granny?)

II. Is Cognition Essentially Semantic?

The first problem first. I begin with a schematic but familiar example, then theorize a bit. I hear some utterances exchanged in a conversation. The context is such that only a part of their joint content conveys new information which matters to me and moves me to action. The *selection* of the new information need not, and most often is not, made on just semantic grounds. If the information is new and important to me, it is not just because I know the semantic values of its (linguistic) constituents. I also have memories, expectations and plans *relative to which* the incoming information can be new and important. This is one thing: the structure of the information I ultimately extract from a variety of sources (input as well as memory) need not reflect only facts of aboutness or reference, that is, semantic facts. It very often reflects my current aggregate cognitive and conative states as they feed into my readiness to action. This latter reflection is constrained therefore by discriminations which are not exclusively world bound, hence semantic, but also memory and desire and plan and action bound, hence pragmatic. The novelty, importance and rele-

vance of information are pragmatic properties defined over this latter sort of discriminations.

Which brings us to the other thing. It is *as* new, important or relevant, hence as pragmatic, that information *causally* contributes to my action. If the information were not new, important or relevant, I would not have acted on it, even though I would have valued it semantically the same way in either case. The point I am making is not confined to conversation. The same semantic underdetermination of the causally efficacious information is also present in inference, decision making, explanation and memory recall. It is a very pervasive feature of our mental life.[3] Fodor has been publicly observed not to like this sort of examples, perhaps because they are semantically messy, but then it wasn't I who painted that apocalyptic and anarchic portrait of central cognition in the *Modularity of Mind*,[4] a portrait which reflects precisely the informational turbulence of the pragmatic sort sampled by my example.

Let me motivate all this with a bit of theory. We can think of information in general in the form of structures shaped by causal interactions under principles of organization or types and laws. I will bring types and laws under the notion of constraints. Let us call *semantic* those properties of an information structure, intrinsic as well as relational, in virtue of which it is about something. And let us say, with Fodor, that *aboutness* itself is defined by the systematic correlation between an information structure and (types of) items in the world. To simplify enormously, let us also say that the key contributors to semantic aboutness are concepts, inputs and their distal causes. On this simplification, an information structure tokened in the brain is semantic because it reflects in its organization the interaction of an input with concepts. This input-concepts interaction allows the information structure it tokens to covary with, and hence be about, distal items in the world. Finally, let us call *mental* those properties of information in virtue of which it is causally potent in central cognition (thinking, planning, and the like) and behavior. My suggestion then is that mental information is rarely just semantic, and hence that mental causation itself is rarely just semantic, because generally the properties in virtue of which an information structure is causally efficacious in cognition and behavior are rarely type identical with the properties in virtue of which such a structure is semantic.

---

[3] For details, I refer the reader to two solid books on the pragmatics of cognition: D. Sperber and D. Wilson, *Relevance* (Harvard University Press, 1987); and John Holland et al., *Induction* (The MIT Press, 1986). See also my "Manufacture of Belief" in Bogdan (ed), *Belief* (Oxford University Press, 1986).

[4] Fodor's earlier book, with The MIT Press/Bradford Books, 1983.

The psychosemanticist can make two moves at this point. One concerns semanticity, the other the mental causation. The first move consists in bringing pragmatically mental properties under semantic properties by defining the latter in such a way (say, as 'nonsyntactic', or as 'not purely intrasymbolic', as Fodor suggests in correspondence) as to accommodate the former. One can make this move, but what's the point? We started with a notion of semanticity which, rough as it is, is positive, concrete and has Fodor's blessing: it is about aboutness and says that aboutness is systematic mind-world covariation. As a result, semantic properties cannot be merely coextensive with nonsyntactic properties, for so are most physical or chemical properties; moreover, most of the nonsyntactic properties have nothing to do with mind-world covariation. Nor can semantic properties be coextensive with pragmatically mental properties, for the latter were just said to reflect novelty or relevance relations to memory, conative states and action which are not constrained only by mind-world covariations. In either case, then, the move entails abandoning a specific notion of semanticity (mind-world covariation) in favor of a negative, hence open-ended and uninformative alternative.

The other move of the psychosemanticist is to argue from (what we may call) downward causal implementation. To be causally efficacious in cognition, an information structure must have physical, biochemical and (very likely) syntactic properties. Yet none of these types of properties can, by themselves or together, explain the causal efficacy of information unless we also consider its semantic and mental properties. Think of this overlay of property types as follows. The ultimate or lowest level implementers of any causal process in nature are always physical. If the implementers are organized only under physical types and laws, or constraints, then their causal powers (i.e. the property types in virtue of which they cause) are Physical. If the physical implementers are *also* organized biochemically (say, as genes), that is, under biochemical constraints, then their causal powers are no longer Physical. They are now physical and Biochemical, for they reflect the genetic form of organization of physical matter. And so on to Syntactic, Semantic and the Rest.

It does no good to argue that causation reduces downwards to its physical bottom. The carriers or implementers of causation do (for they always are elementary particles) but the *form* of causation does not. Causes cause in virtue of their organization or form. The latter is always constrained (typed and law-governed) at a certain level of complexity. Genes, for example, do not (typewise) cause what they do in virtue of just being elementary particles. The same is true of the semantic structures in the mind. Surely, semantic structures (in humans) are syntactic, as they are physical,

and indeed are causally efficacious because syntactic and physical. But, I point out, semantic structures are *further* constrained when they have to do a mentally pragmatic job, in which case the form of their causal efficacy is no longer just syntactic, although it respects all the syntactic constraints. (Likewise, a syntactic structure causes not just in virtue of being a physical structure, although its causal efficacy must respect all the physical constraints). This is why the causal powers of the information pragmatically constrained reflect a new form of organization, the Mental form.

To see why all this matters, we need to plug the point into a larger argument. Fodor himself provides some key premises in his excellent chapter 2. After reminding us that Putnam's Twin examples show that the common sense relational individuation of cognitive states violates supervenience, he makes the very reasonable and often forgotten point (which I enter as a first premise) that the neural, molecular and subatomic states, underlying the cognitive ones, can also be individuated relationally (31). This is obvious when we read these sorts of states informationally (which we should, when discussing cognition), for information can only be understood relationally. Two questions then arise. Why do we say of cognitive states, but not of neural or subatomic states, that they encode *semantic* information? And why do we also say that the semantic information internally encoded by the cognitive states of the Twins does *not* distinguish between H_2O and XYZ and hence is not relational *to that detailed or even complete extent*?

Now we have a bit of a handle on the answer. Semantic information states reflect a type of organization, and hence have causal powers, that neural and subatomic states by themselves do not have. (For example, only the former, but not the latter, are organized by concepts and cause in virtue of their conceptual organization.) The structure and causal efficacy of semantic information is designed to neglect lots of input differences, among which is that between H_2O and XYZ. Yet the latter difference is one which the underlying neural and subatomic states are not likely to neglect.

This sort of answer seems to be in the spirit of Fodor's next major point, which I enter as a second premise in my argument. Methodological individualism, he writes, instructs that cognitive states be individuated "with respect to their causal powers" and that "no property of [cognitive] states, relationally or otherwise, counts taxonomically unless it affects causal powers" (42). As a result, when cognitive states cause in virtue of the semantic information they encode, the Semantic causation does not capture the difference between H_2O and XYZ. (The difference would show

up only if the causation were Chemical.) Methodological individualism thus urges that we taxonomize the cognitive in terms of the *types* of states which have appropriate causal powers. And so I do. The conversational example and the argument built around it indicate that the mind deploys types of structures whose causally efficacious information is fixed and operates under constraints (such as recognition of new information, relevance to explanation or decision, guidance of action, etc.) which are not merely semantic. Since those types of structures have causal powers in virtue of properties which are other than semantic, methodological individualism requires that we taxonomize the information encoded as Mental, not Semantic.

Recall now that the first premise of our argument (supplied by Fodor) says that various types of internal states involved in cognition can be individuated relationally. At various levels of organization, such types of states encode information from the environment. The difference between, say, the neural and semantic forms of information originates in the constraints (types and laws) on their encoding. The semantic form is constrained to abstract (or conceptualize) away facts of the environment that the neural form is not. If we ask why there are such differences, we must appeal to the design of cognition which counts on specific forms of information to run specific mechanisms and functions.

If this much is granted, then it is not clear why the causal powers of cognition ought to reside in its semantic properties only — or, equivalently, why the semantic properties of information ought to be the closest to the causal trigger. Since the evidence suggests that the causal mechanisms involved in cognition align internal states not only to the way the world is (a semantic task) but also to memory, plans, focus of interest and action, and so on (a pragmatic task), it ought to follow that the information constrained to drive these mechanisms must have a form of encoding appropriate to its pragmatic tasks. That form is Mental.

Are, therefore, semantic properties and causal power attributable to the very same things (contents), as Fodor and so many others claim? Not always and not essentially. Intriguingly, further confirmation of the tension between semantic content and causal potency in mentation emerges from Fodor's splendid criticism of the doctrine which has tried the hardest to reconcile the two, functional role semantics. I joyfully agree with Fodor's criticism but I read it as making my point, not his. Which brings us to our second major problem.

III. The Naturalization of Psychosemantics

How should we type individuate the semantic aboutness or content of a cognitive state? By attending to its functional role, hence to its causal

interactions with other states, says one popular answer, meaning holism, best defended by functional role semantics. Functional role semantics holds that function determines content. For this to work there must be a theory of how function shapes content nonsemantically, or else the enterprise is unmotivated and circular. Fodor shows convincingly that the functional role enterprise fails on both counts. In so doing he provides an interesting dialectical opening to our discussion.

Nothing is more dialectical than a dilemma. Functional role semantics faces a fat one. If it is serious about *functional role*, then it must be serious about mental causation, because a functional role is a causal role redescribed. But then, what plays a functional/causal role cannot normally be just a semantic content, since (by my previous argument) the latter by itself does not have causal power. Thus, a functional role semantics which is serious about mental causation cannot be serious about semantics. If functional role semantics is serious about *semantics*, then it cannot be serious about mental causation (hence about functional role) because, again, semantic contents do not have causal powers, mentally. In either case, functional role does not fit with semantics. Let us take a closer look at this diagnosis.

The functional horn first. The causally potent cognitive states are said to take the form of beliefs, intentions and the like. If their functional roles determine anything, it is their *attitudinal* types (believing, intending, etc.), as Fodor points out, rather than their contents. We do talk of the functional role of a belief content, and mean it to normally individuate the conceptual associations which shape, along the pragmatic lines suggested earlier, the mental information which causally drives cognition and behavior. What is functional here is a mental, not a semantic role. If functional role were to determine information, it should be *mental*, not semantic information. Functional role appears to determine semantic content only when the causation involved obeys constraints no higher than Semantic. My introductory example of a modus ponens was a case in point. In normal, unregimented, full bodied mentation, that is the exception, not the rule.

There is, however, a further problem. When functional role (normally) determines mental content or (rarely) just semantic content, what it determines in fact are associations among concepts (in the form of propositions), not the concepts themselves. The concepts are antecedently determined. Since concepts are the key internal shapers of aboutness, functional role semantics must assume, and therefore cannot account for, semantic aboutness.

The only hope for functional role semantics is to retreat to the ur-contents, the concepts. This brings us to the semantic horn. The trouble with ur is that it never causes anything. This lands the project in a further minidilemma. On the one hand, by themselves concepts are causally inert; if we want to specify their aboutness by listing their potential causal roles (good luck!), we may end up with indefinitely long and useless lists. When, on the other hand, concepts become causally active, it is by way of animating thoughts and beliefs, which takes us back to the first and major functional horn.

There must be some other way to get at what concepts and generally cognitive states are about. Fodor's, in chapter 4, is the denotational way. The aboutness of a cognitive state should be understood in terms of its denotation (concepts in terms of the properties they apply to, thoughts in terms of the facts making them true, and so on). The tough question for a philosopher of mind is of course, How do cognitive states *get* to be about their denotations? Fodor wants a naturalist answer. He wants "a theory that articulates, in nonsemantic and nonintentional terms, sufficient conditions for one bit of the world to *be about* another bit." (98) No question, that's the right destination to arrive at. But is Fodor's way the way to get there?

What could the nature of those sufficient conditions be? CAUSAL, more or less, says Fodor. The naturalization of psychosemantics must be the principled output of a sophisticated causal theory. Why sophisticated? Because any crude causation (say, in the form of a nomological correlation of the sort 'A' means A iff all and only A tokens cause 'A' tokens) will not do. This is because there is error (a not-A can token an 'A') and hence disjunction (either A or B can token an 'A') and the crude theory cannot tell the difference. Neither Dretske (learning the reliable correlation A to 'A' first, mistokening 'A' afterwards) nor teleology (thought to guarantee optimal tokening of mental symbols) can help much, so what's left is to upgrade the good old causation. I am not going to go into the logical details of Fodor's bold enterprise, which are many and ingenious. The net result is an interesting asymmetric relation between good and disjunctive (or error prone) tokening of mental symbols. The good wins by being independently the first to call the semantic shots. How does it do it? How does the semantic game get started in the first place? Now this is a really beefy question.

It brings us back to the notion of the form of causation and invites a related notion as well, that of frame of causation. The notion of form of causation, we recall, says that a token A causes a token B, among other things, in virtue of A's type of internal structural properties. One can leave

it at that if both A and B belong to the same *frame* of causation, meaning that the same constraints apply to both A and B. A frame includes all causes and effects of the same form. Molecules pushing other molecules around form one big causal frame. Change the frame, by changing the constraints on the effect, and we have causation *across* frames. Light impacting the retina is an example of causation across frames, since light waves are not constrained by the same types and laws as retinal patterns are. (Transduction in general is causation across frames.) The result of causation now is the joint product of the constraints operative in the two frames: the retinal information tokened by the light input reflects (in its organization) both the retinal constraints *and* the impact of the light. If, as a result, the retinal pattern is said to be about something, that aboutness cannot be due to the light input alone, for it must also reflect the constraints on the retinal organization. The causation from one frame (light) cannot by itself *type* the right information structures in another frame (retina).

We can understand why learning cannot be a simple, one frame causation. Learning is causation (of concepts) across frames. When 'star' means star because the two are causally connected, the causation in question is across frames. The 'star' in the head is not subject to the same types and laws as the 'star'-star connection. The constraints operative in the 'star' frame must figure in the account the 'star'-star covariance. If the conditions on covariance miss the 'star' frame, they are not going to deliver the required naturalization. The external stimulus (from stars) cannot type a concept ('star') in an organism unless there are appropriate internal constraints ready to oblige. This is how learning goes, and Fodor (possibly the wildest innatist alive) knows and loves this truth more than anybody. This does not mean that learning does not generate lots of semantic types (concepts). It surely does but always by exploiting prior types and available typing procedures. Even the psychophysical correlations which come closest to satisfying the crude causal theory do presuppose some prior semantic typing. It is a hardware sort of typing: I am *built* in such a way that red tokens out there reliably cause tokens of 'red' up here. The hardware typing is not itself semantic, which is a good starting point. My hardware design constrains the tokening of my mental color symbols. As Fodor notes, psychophysics does (noncircularly) naturalize the semantics of the crudest registrations.

The psychophysical naturalization of semantic content is only the first leg of Fodor's journey (112-20). I think it is a right (although too short) leg to start with. The psychophysical gambit is not *causal*: the color types are already there, hardtyped, though not necessarily as mental symbols. The

causal interaction with redness simply tokens the existing hardtype 'red'. As Fodor knows, the psychophysical move does not take us too far. But he thinks that it will take us to some psychophysical concepts more complex than red (say, looking horsey, or looking protony in the cloud chamber). The argument is that, although 'horses' and 'protons' are not psychophysical concepts, their tokens (real horses and protons) are responsible for, and reliably covary with, tokens of related psychophysical properties. We have now a firm and nonsemantically defined relation between psychophysical concepts and tokens of real concepts. So far so granted. The question is, How do we get to the latter?

At this point I must digress to confess. Fodor's psychophysical move quenches my naturalist thirst, for it attempts to *explain* semantic typing from (nonsemantic) hardware constraints and hardware-world correlations, and does so by being sensitive to causation across frames. But Fodor's next moves fooled me. (I am not sure I am fully recovered.) True, Fodor warns us. He says (118) that the story he is going to tell is old fashioned since it connects having concepts with having experiences, and knowing meanings with knowing what would count as evidence. True, he also warns of some honest cheating (120). And indeed we are told the old fashioned story of how physicists *know about* the correlations between psychophysical concepts (looking protony) and the tokens of theoretical concepts (protons). That theoretical knowledge bridges the gap to the semantic types or concepts. The grand correlation we are after, from proton to its concept, is implemented by psychophysics and theory internalized as knowledge. "True theories, when internalized, correlate the state of the head with the state of the world" (121).

It does not matter what concrete mechanisms allow worldly tokens to exert causal control over concepts; a naturalist semantics does not need to specify *all that*. All that matters is that causal control through a reliable correlation does obtain and that it can be characterized nonsemantically (121). So it is now official: "For purposes of semantic naturalization, *it's the existence of a reliable mind/world correlation that counts, not the mechanisms by which that correlation is effected*" (122).

Now you can begin to understand why I was fooled. This is not what I took the naturalization of psychosemantics to be. *Mea* culpa. I applauded Fodor's first naturalization move to psychophysics precisely because it allows reliable causal chains to token protosemantic types which are themselves *explained* in terms of the *design* of the hardware. I thought that is the sort of understanding the naturalizer of cognition craves for and that Fodor is going to deliver. The hardware frame was as instrumental as the stimulus frame in explaining what sort of information is being

shaped psychophysically. Fodor's later moves turned out to be too thin for my taste. And then it dawned on the conceptual analysis side of my philosophical brain (which went to deep sleep in the last few years) that what Fodor is doing is indeed the thin, metaphysical naturalization of conceptual analysis, of the sort found in analyses of seeing and knowledge (Chisholm, Armstrong, Dretske, Goldman, others). There is nothing wrong with these analyses; they are important. It is just that they represent a different project.

I deliberately choose seeing and knowledge as points of comparison to emphasize the *success evaluation* task of their conceptual analyses. Seeing and knowledge are concepts which evaluate the success of vision and cognition. One need not be told what the *nature* (or design) of vision and cognition is in order to evaluate their success. One need only be told of the success conditions. The naturalist analyses use causal or nomological correlations to specify those success conditions, while assuming that the nature of the relata thus correlated (in particular, the nature of vision and cognition) is antecedently and independently understood. The analysis of seeing or knowledge is not an account of vision or cognition.

I see the same contrast between an analysis of semantic success (say, of concept application) and an account of semantic typing (say, having concepts, or having this as opposed to that concept). The analysis of semantic success tells us in what conditions a semantic type (a concept) applies or tokens successfully; it does not tell us what makes the type in question *semantic* (why, for example, it is a concept type and not a neural type), nor does it tell us, more specifically, why it is *that* semantic type and not another (the concept of X and not the concept of Y). Analyses of success conditions are analyses of conditions of semantic tokening, not typing. We have to understand the type (concept) before we worry in what conditions it tokens successfully — or unsuccessfully, for that matter.

With this diagnosis in hand, let us look at what Fodor says he is doing. He says (121/122) that in order for the concept 'proton' to be about token protons out there in the world unspecified internal mechanisms must *successfully* track or correlate with the tokens. That is all. The view assumes but does not explain the presence of concepts. Elsewhere Fodor says, ". . . what we wanted for semantics was naturalistically specifiable conditions under which instantiations of *proton* are guaranteed to affect what is in the belief box" (119). So there is *already* a belief box full of concepts by the time Fodor's naturalization goes into action. Fodor's angle of analysis takes in only the route *from* concepts to their instances (124). In all these cases, Fodor is talking about the (sufficient) naturalized conditions for *semantic success* (such as concept application), NOT about the

antecedent conditions of semantic typing (of having a concept in the first place). I would have thought that the ur-problem of psychosemantics is the latter, not the former. I would have thought that the question a naturalizer of psychosemantics must ask is, What makes *brain states* be *about* protons or horses or anything at all? — and not, GIVEN that brain states are about protons or whatever, when do they *succeed to be about* protons or whatever? The latter is a question about job *performance*, the former about the *job* itself.

Why can't the performance account for the job? It can but only with the assistance of some philosophical trick. The success of concept tokening is evidence for the concept. When does the evidence count as characterizing the nature of what it evidences? Remember good old clean positivist times? This is when. But didn't I say that Fodor said that his naturalization story is going to be sort of old fashioned and that Granny grinned and rocked and said "I told you so" (118, then 122)? A few pages later (125/126), Fodor acknowledges that his account is verificationist *minus* the specification of the route and the mechanisms from concepts to their tokens: "The moral is that *the route doesn't matter (much)*; what makes 'star' mean *star* is *that* the two are connected, not *how* the two are connected. It is the covariance that counts."

I am not urging that the naturalization of psychosemantics entail a historical account of concept formation, nor that it entail an account of the mechanisms and causal routes involved in concept application. The naturalization question is not about how we acquire a concept, nor about what it takes causally to apply a concept successfully. The question is what it is to have concepts or intentionality generally, that is, types of states with aboutness. It is as much a question of empirical metaphysics (not of descriptive psychology and physics), as is that of semantic success.[5] It is a question about the nature of intentionality, or if you like, about the nature of the cognitive contribution to semantic success. It is a question that needs answers in the form of constraints on the organization of information processing, irrespective of its hardware and causal implementation. Neural or any other sort of states would not be semantic, and hence successful in covarying with the world, unless they were internally constrained to do so. I expect the naturalization of psychosemantics to account for the internal constraints on aboutness.[6]

---

[5] 'Empirical metaphysics' is Fodor's term for the nature of his account in this book (in correspondence).

[6] I develop these points in "Information and Semantic Cognition," *Mind and Language* 3 (1988): 81-122, and "Guidance to Goal: The Roots of Teleosemantics," *forthcoming*.

The *how*, then, does not matter, except in one fundamental respect: not any piece of worldly ware *can* be about another. Covariation may tell us that and when one is about the other (the evidence) but it does not tell why or in virtue of what (the nature). The latter are *design* questions about how a piece of ware types or organizes the information from the world in formats which allow that information to be about aspects of the world. The design part of the *how* question (revealing the intentional nature of cognition) matters enormously, or else there would be nothing out there in the world to covary with something up here in the head. The intentional design is implemented by, but is not the same as, mechanisms and causal routes. We know that much from functionalism. If I want to know why something is a car engine, I have to understand its design or principles of organization and operation; I do not have to understand the particulars of the implementing mechanisms and causal routes; I do not have to look under the hood. (I never do.) Fodor is right to exclude from the naturalization of psychosemantics the how of the mechanisms and causal routes but wrong, I think, to also exclude the how of the very design of the semantic psyche.

IV. Concluding Caveats

Does this mean that Fodor's naturalization project is off the track? I can't and don't want to say that. To see why, let us make a temporary terminological convention. I will call the external relations between symbols, other cognitive states, etc. AND the world *denotational*, and the internal properties in virtue of which symbols and other cognitive states are denotational *intentional*. Thus, semantic = denotational + intentional. Throughout this discussion I took the notion of psychosemantics to overlap considerably with that of intentionality (as just defined). This is why I was fooled by Fodor's analysis. I thought that he wants to naturalize intentionality (the psychological part of the psychosemantics), whereas in fact his project has been all along to *assume* intentionality and naturalize the conditions of its semantic, indeed denotational success.

Yet there may be a more troublesome reading of Fodor's project. It may well be that his external (denotation as covariation) angle on the naturalization of psychosemantics is *all* we can hope for. To understand intentionality (or *psycho* semantics) is to understand its external or *denotational semantics*. There may be no deeper questions about intentionality than those about the naturalized conditions of its denotational success. I resist this pessimism. Yet if Fodor is right about the naturalization of psychosemantics in terms of the success conditions for the application of cognitive symbols and concepts, then one can see how (as anticipated at the

end of the introduction) he may also turn out to be right about the causal potency of cognitive states qua Semantic.

If the mind is essentially semantic (as Fodor assumes), and if what we need to and can know about the semantic mind (as he attempts to show in this book) are the naturalized conditions of its denotational success, then knowing those conditions amounts to knowing what the mental causation must be like. The mental causation must simply reflect and secure, through syntax, the conditions of its denotational success. This is another way of saying that (i) syntax reflects, respects and tracks semantic relations, (ii) mental causation takes Syntactic form and (iii) that's it. Then we can confidently conclude that what causally moves the mind is precisely what makes it semantically successful. If that is true, if (internal) intentionality is *fully* reflected in the conditions of its (external) denotational success, then an analysis of the latter may amount to an account of the former. The rest is hardware, mechanisms, causal routes and other inessential implementers that we can joyfully quantify over. The (denotational) semantics runs the psyche.

Is that true? I doubt it, for I doubt Fodor's first thesis (that the mind is essentially Semantic) and the thesis established in this book (that naturalizing semantic success amounts to naturalizing *psycho* semantics and explaining its causal potency). But I have no doubt that if anybody *can* make semantics run the psyche, that's got to be Jerry Fodor.[7]

[7] I want to thank Jerry Fodor and Graeme Forbes for good, perceptive and helpful comments on an earlier draft.