

## A Guide to Naturalizing Semantics

### I

Semantic predicates - *is true, refers, is about, has the truth conditional content that p* and so forth- apply to various items, most centrally to natural language expressions and to mental states and events (types and tokens). For example, both the sentence "The cat is crying" and the belief that the cat is crying are about the cat and possess the truth conditional content that the cat is crying. It is generally (and correctly) held that the semantic properties of natural language expressions (and other non-mental representations) are derived from the semantic properties of mental states.<sup>1</sup> According to one version of this view natural language expressions obtain their semantic properties from conventions governing their use. These conventions are themselves explained in terms of regularities involving the beliefs, intentions, and so forth of speakers of the language.<sup>2</sup> In the following I will assume that some such view is correct and focus on the semantic properties of mental states and events.<sup>3</sup>

The semantic properties of mental states are what makes them *intentional states*. Thus the intentional content of e.g. the thought that the cat is on the mat is the truth conditions of the thought. The topic of this paper is the question: In virtue of what do intentional mental states/events possess *their* semantic properties? For example, what makes it the case that a particular thought is about the cat and has the truth conditions that the cat is crying? The answer cannot be the same as for natural language expressions since the conventions that ground the latter's semantic properties are explained in terms of the semantic properties of mental states. If there is an answer, that is, if semantic properties are real (and really instantiated) and are not fundamental, then it appears that they must be instantiated in virtue of the instantiation of certain non-semantic properties. Recently a number of philosophers, I will call them "Semantic Naturalizers," have attempted to answer this question in a way that they take to be compatible with *Naturalism*. Semantic Naturalism's central contentions are a) that semantic properties, laws, causal relations involving them obtain *in virtue of* the obtaining of facts constituted entirely by naturalistic properties etc. and b) that semantic properties are kinds of the sort suitable for investigation by the methods of the natural sciences. The first contention is usually glossed as a *supervenience* claim to the effect that

---

<sup>1</sup>. Proponents of this view include Grice (1957), Lewis (1968), and Fodor (1975). For a contrary view see Davidson (1984) who holds that mental and public language semantic properties are interdependent and that neither is metaphysically prior to the other.

<sup>2</sup>. The program of accounting for the semantic properties of natural language in terms of those of mental states is identified with Paul Grice (1957) and Stephen Schiffer (1972). A detailed account in terms of conventions can be found in Lewis (1969). Also see Avramides (199?).

<sup>3</sup>. So in the following "semantic property" means semantic property of an intentional mental state or event.

semantic/intentional facts metaphysically supervene on naturalistic facts including naturalistic laws and causal relations. Or, to put it metaphorically, Naturalism involves the thesis that for God to create our world He needed only to have created the pattern of naturalistic property instantiations and the natural laws. Everything else- and specifically tpatterns of instantiations of semantic/intentional predicates- and the laws an causal relations involving them follow from these.<sup>4</sup> The second contention involves the idea that semantic/intentional properties can figure in causal explanations and in laws. They provide a suitable taxonomy for a science of the mind. It is plausible that semantic/intentional properties are scientific in this sense only if the first contention is true. But, of course, it is possible that mental predicates or properties supervene on naturalistic ones even though they don't figure in any science like the natural sciences.

Exactly what properties qualify as naturalistic? The best way to answer this question shifts the discussion to predicates. Naturalistic predicates include those that occur in the natural sciences, physics, chemistry, biology etc. It excludes intentional and semantic predicates but it is usually taken to include certain *modal* notions, for example, *causality, law, information, chance*, counterfactuals and dispositions. Naturalistic properties are then those that are the referents of predicates that can be defined in terms of naturalistic notions.

Semantic Naturalism is a *metaphysical* doctrine about the status of semantic properties.<sup>5</sup> The Semantic Naturalizers I will discuss also seem to accept an *epistemic* thesis that I will call "perspicuous semantic naturalism." It is the view that, at least in some cases, the metaphysical connections between naturalistic and semantic properties are sufficiently systematic and transparent to allow us to understand how it is that certain naturalistic conditions are metaphysically sufficient for certain semantic properties. If Semantic Naturalizers can find naturalistic conditions that are metaphysically sufficient for semantic properties and understand why they are sufficient they would show how semantic naturalism can be true and thus place the semantic within the natural order. If the relationships between the semantic properties of mental states and naturalistic (e.g. biological) properties is sufficiently systematic to count as a reduction then they may hope to secure the scientific status of intentional properties.

Although Naturalism in something like the above sense is widely endorsed in contemporary philosophy there is also a tradition from Brentano through Wittgenstein, Davidson, Hornsby, Putnam, McDowell, Searle and Chalmers that is inhospitable to it. These authors reject what I have called "perspicuous semantic naturalism" and perhaps even the supervenience of the

---

<sup>4</sup>. The proposition that Fx is metaphysically entailed by conditions K just in case K together with a characterization of the nature of F logically imply Fx. The best understood example of this is the realization of a functional property F by lower level property instantiations. In this case it logically follows from the functional nature of F, the nature of the Ps and causal relations among the Ps that whenever the Ps are instantiated M is also instantiated.

<sup>5</sup>. Although it is a metaphysical doctrine it is also contingent since its truth doesn't rule out possible worlds in which some properties are instantiated but not in virtue of the instantiations of natural properties.

intentional/semantic on the natural. They also (except perhaps Chalmers and Searle) think that intentional properties (or concepts) are not suitable for occurring in scientific laws. Adherents to this tradition think that the intentional and the semantic possess features that are radically different from anything found in the natural science. In their view these differences make for a chasm between the intentional/semantic making semantic naturalism an impossibility. Three lines of thought have been especially influential in this regard.

1. Semantic/intentional properties are essentially normative and specifically involved with assessments of correctness and rationality. For example, it is claimed that it is constitutive of the concept *cat* that it *ought* to be applied only to cats. Or that it is part of the content of *and* that one is mistaken if one infers P and Q from Q. Since, there is nothing correspondingly normative to find in the fundamental properties of the sciences it allegedly follows that intentional and semantic facts cannot be reduced to naturalistic facts.<sup>6</sup>

2. The second line of thought is that the principles that govern the attribution of semantic predicates lead to the indeterminacy of the semantic attributions even given all possible relevant evidence. For example, given all of a person's verbal dispositions (the supposed totality of relevant evidence) principles of attribution license alternative assignments of truth conditions and references to that person's sentences and terms. Similarly the functional roles and causal relations of a person's mental representations are claimed to be insufficient to determine reference and other semantic properties.<sup>7</sup> From this it is a thought to be a short step to the conclusion that semantic facts (if there are any) fail to supervene on naturalistic facts.

3. A third (though less prominent) line of thought is that semantic properties of mental states are metaphysically connected to consciousness properties- i.e. the *what its like* features of experience and thought- and that consciousness properties fail to supervene on natural properties. The idea is that a perfect physical duplicate of a person need not also duplicate its conscious experiences. It is a Zombie. And since intentional states (or certain intentional states) require (specific) conscious experiences the Zombie also lacks all or certain intentional states.<sup>8</sup>

---

<sup>6</sup>. There are two issues that are often mentioned by those who think that normativity considerations derail semantic naturalism. One is that grasping a concept involves being in a mental state that obligates one to applying the concept only to items in its extension. It is difficult to see how any purely natural state (or rather naturalistic description) can entail such an obligation (Kripke 1982, Boghossian 1989 and Hale and Wright 199?). The other consideration is the claim that the attribution of intentional concepts is constrained by normative principles of rationality and charity. Davidson (1981, 1984) starts with this claim and tries to fashion it into an argument against the existence of nomic connections between intentional and non-intentional properties. There is little agreement about exactly what Davidson's argument is or even whether its conclusion conflicts with naturalism. Even so it has been influential and is often cited or repeated by those skeptical of naturalization.

<sup>7</sup>. Quine's (1960) arguments for the indeterminacy of translation and for the inscrutability of reference and Putnam's (1978) so called "model theoretic argument" are instances of this line of thought (also Wright 199? and Hale and Wright 199?).

<sup>8</sup>Chalmers argues like this in "The Epistemology of phenomenal Content" and

It is far from clear how far, if any distance at all, these considerations and the arguments fashioned from them go in undermining semantic naturalism. The arguments I know of (e.g. Davidson, Putnam, Davidson, Kripke, Chalmers) that are supposed to- or have been taken to- establish the conclusion that intentional/semantic facts (if there are such) fail to supervene on natural facts are either very obscure or involve highly controversial premises (e.g. that supervenience requires that naturalistic descriptions *a priori* entail intentional/semantic descriptions). Semantic naturalizers think that the anti-naturalization considerations are sufficiently serious to motivate semantic naturalizers to defend semantic naturalism by producing a naturalization of certain semantic properties. The naturalizers grant that an adequate account of semantic properties will need to account both for whatever normativity content properties possess and for the determinacy of reference and truth conditions. We will see these issues coming up in various ways in our survey of naturalistic theories.

## II

The mental states that have been the focus of naturalization proposals are the propositional attitudes; desire, belief, and perception (perceptual belief). There are two parts to naturalizing a particular kind of propositional attitude. First, is specifying natural facts in virtue of which it is an attitude of that particular type, e.g., a belief or a desire. Second, is specifying the natural facts in virtue of which it has its semantic properties its particular truth conditional content. With regard to the first part the view held by most semantic naturalizers is that the property of being a particular kind of attitude e.g. being a belief, is a *functional* property [Fodor 1987]. Functional properties are higher level properties instantiated by an individual *x* in virtue of *x* (or *x*'s parts) and other entities instantiating lower level properties that are lawfully or causally related to each other in certain specified ways.

Most semantic naturalizers also think that the property of being a belief (same for other propositional attitudes) involves an internal mental representation and this representation bears the state's semantic properties.<sup>9</sup> On this view, for example, the belief that the cat is crying involves a relation to an internal representation that has the truth conditional content that the cat is crying. Some semantic naturalizers further propose that mental representations are elements in a *language* of thought, "Mentalese."<sup>10</sup> On this view complex mental representations are composed of names, predicates, logical particles, etc., arranged in syntactic structures. Naturalizing the semantics of Mentalese consists in specifying the natural facts in virtue of which simple Mentalese expressions possess their semantic

---

Searle pursues a similar line though he thinks that conscious experience is a natural phenomenon that is "caused by and realized in" the brain.

<sup>9</sup>. Proponents of this view usually distinguish between explicit and implicit propositional attitudes. Only the former involve relations to mental representations. The latter are dispositions to produce explicit attitudes (Fodor 1987 ch 1.)

<sup>10</sup>, Field (1978) and Fodor (1975, 1987) are important sources of this view. Fodor proposes it as an empirical hypothesis that provides the best explanation of certain features of human thought, specifically systematicity and the capacity to engage in logical reasoning.

properties and then showing how the semantic properties of complex expressions are determined by their structure and the semantic properties of their constituents (Field 1972, 1978). While not every Semantic Naturalizer buys the language of thought hypothesis it will often be convenient to presuppose it in what follows.

There are two conceptions of semantic content that have figured in recent discussions of naturalizing content called "broad content" and "narrow content." "Broad content" refers to the usual truth conditional content of intentional mental states. Hilary Putnam (1975) posed thought experiments that have been taken to show that the usual truth conditional content of certain thoughts fails to supervene on the thinker's intrinsic physical properties. Putnam imagined two people, Oscar and twin-Oscar, who are identical with respect to their intrinsic neurophysiological properties but differ in the following ways. Oscar lives on Earth and speaks English. Twin-Oscar lives on a twin-Earth and speaks twin-English. The primary difference between earth and twin-earth is that on the latter planet the liquid that fills the oceans, that quenches thirst etc. is not H<sub>2</sub>O but XYZ, a chemical compound indistinguishable from H<sub>2</sub>O without chemical analysis. Putnam claims that Oscar's and twin-Oscar's utterances of "water is ..." and the thoughts that each expresses with the sentence differ in their truth conditions. Oscar's thought is true iff H<sub>2</sub>O is...and twin-Oscar's thought is true iff XYZ is ... If this is correct then intentional properties (at least in some cases) do not supervene on intrinsic neurophysiological properties or any properties that supervene on them (e.g. computational or syntactic properties). This view, *semantic externalism*, is now widely held for thoughts that involve natural kind concepts like *water*.

"Narrow content" is a term introduced to designate content properties that do supervene on intrinsic neurophysiological properties (Fodor 1981, 1987). While Oscar and twin-Oscar's thoughts differ in broad content they agree in narrow content. Some philosophers (Fodor 1987) have argued that only narrow content properties are implicated in intentional causation and for this reason are required by an intentional science but there is little agreement concerning exactly how to characterize it or even whether there are such properties (Stalnaker 1991). In any case, most of the naturalization proposals concern broad properties, specifically reference and truth conditional content, so that will be our focus here.

What naturalistic facts are plausible candidates to serve as metaphysically sufficient for the semantic properties of mental representations? Putnam's twin-Earth thought experiments and Kripke's well known theory of proper names (Kripke 1972) both suggest that causal relations are involved in determining the references of predicates and names. Their considerations seem to carry over to mental representations corresponding to predicates and names. It is plausible that Oscar's mental representation "water" refers to H<sub>2</sub>O partly in virtue of the fact that H<sub>2</sub>O has caused or is apt to cause Oscar to think water thoughts. And it is also plausible that part of the account of what makes a person's mental representation "Aristotle" refer to Aristotle is that it possesses a causal history that originates with a baptism of Aristotle. Neither Putnam nor Kripke are sympathetic to the naturalization project but their work is often taken as the starting point for naturalistic proposals. Causation and kindred notions like law, counterfactuals, and probability seem to be the "right stuff," if there is right stuff, out of which to try to build naturalistic accounts of

intentionality.<sup>11</sup>

### The Crude Causal Theory

I will begin our survey of specific naturalization proposals with the crude causal theory (CCT) for the reference of mentalese predicates @. No one has ever held the CCT but it will be useful to describe it and note its most obvious defects since these are the problems that more sophisticated accounts are designed to solve.

CCT) It is metaphysically necessary that (if tokens of @ are caused by and only by instances of the property F then @ refers to F).

The obvious problem with the CCT is that it doesn't allow for the possibility of tokening @ or a sentence containing @ that is not caused by F. This is called "the problem of error" since if @ occurs as part of the perceptual belief that x is a @ then since @ is caused by F it follows that the belief is true. But of course a perceptual belief, for example, the belief that x is a cat may be caused by a small dog not by a cat. The problem of error is a special case of the disjunction problem. Since the CCT implies that whether or not @ is a component of a belief the disjunction of all the causes of @'s tokens are the reference of @. So if @ is caused by cats, small dogs, utterances of "cat" etc. then CCT says that @ refers to the property of being a cat or a small dog or an utterance of "cat" etc. Clearly many of the causes of @ need not among its reference. A naturalist successor to the CCT will need to find some way of naturalistically distinguishing the reference constituting causes from the others.

A second problem is that semantic relations are apparently more fine-grained than causal relations. This is the "fine-grainedness problem." @ may refer to F and not G even though F and G are metaphysically or nomologically co-instantiated. For example, the properties of being triangular and the being trilateral are apparently distinct but necessarily co-instantiated. Triangular things cause tokens of @ just in case trilateral things do but a predicate can refer to one property but not the other. Quine (1960) pointed out a pervasive type of property co-instantiation. When and only when the property of being a rabbit is instantiated so is the property of being an undetached rabbit part. When one of these properties is causally linked to @ so is the other. This makes it quite difficult to see how a causal theory can account for the difference between thinking that there goes a rabbit and

---

<sup>11</sup>. Causation, laws, counterfactuals, etc. are not themselves items mentioned in physics and it is controversial whether they supervene on physical facts. Even so, Fodor and other naturalizers would consider it a successful naturalization if they could show that intentional properties supervene on these properties. However, Putnam (1993) has complained that notions of law and causation presuppose intentional notions. While this may be true on some accounts of these notions it is not true on others. For example, on some accounts probabilities are rational degrees of belief. Obviously explaining semantic properties of beliefs in terms of degrees of belief would not contribute to naturalization. On other accounts probabilities are objective mind independent features of the world. In this case there seems to be no danger of circularity though one may wonder at employing so metaphysical a notion in the cause of naturalism. However, these issues are too complicated to develop here.

thinking that there goes an undetached rabbit part.

### Dretske's Information-theoretic Account

Fred Dretske (1981) proposed a close relative of the CCT that identifies the truth conditional contents of beliefs with part of the information that states of that the belief carries under certain circumstances. The notion of information can be defined this way: state type T carries information of type p iff there is a nomological or counterfactual regularity (perhaps a *ceteris paribus* law) to the effect that if a T occurs p obtains.<sup>12</sup> So, for example, the height of mercury in a thermometer carries information about the ambient temperature. Dretske's idea is to construct the content of beliefs out of the information that they carry under certain circumstances. An initial and crude formulation of the theory is:

DRET) It is metaphysically necessary that (if B carries the information that p then B has the truth condition that p).<sup>13</sup>

Versions of both the fine-grainedness and the error problem cause trouble for DRET. If B carries the information p and p implies q then it also carries the information that q. But, of course, one can believe that P without believing that q even though p implies q. Dretske responds to this problem by identifying the content of a belief with the *maximal* information that it carries under certain circumstances. This is a little progress but it leaves untouched the problem that if p and q are nomologically or metaphysically co-occurring then any state that carries information that p carries the information that q. So according to DRET no belief can have the exact content that there is a rabbit since any state that carries the information that there is a rabbit also carries the information that there is an undetached rabbit part. Notice that it is of no avail to protest that a given believer might not even have the concept *undetached* since that doesn't effect the fact that his belief state still carries the information that there is an undetached rabbit part. Dretske's attempts to handle this problem are not successful.<sup>14</sup>

The error problem arises for DRET in this way. According to DRET the belief that p always carries the information that p and that means that whenever the belief is tokened it is true. Dretske's proposal for solving the error problem is to identify a subclass of the actual tokenings of B as the bearers of the information that constitutes B's truth conditions. Tokens of B outside of this class have the same truth conditional content as those within the class although they may not carry the same information. This permits (but doesn't obligate) the latter tokens to be false. Dretske's initial specification of the class of tokens of the belief state that fix its truth conditional content is the class of tokens that occur and are reinforced during what he calls "the learning period." His idea is that during this

---

<sup>12</sup>. Dretske (1981) characterizes information in terms of probabilistic relations. There are numerous problems with his account that are avoided by the characterization used here. Also see Loewer (1987).

<sup>13</sup>. Dretske characterizes belief functionally as states that guide behavior in certain ways. He doesn't commit himself to a language of thought account of beliefs.

<sup>14</sup>. This is forcefully argued in Gates (1995).

period a type of mental state becomes a reliable indicator of *p* and so comes to have the content that *p*. So Dretske's official account is

DRET\*] N(if the maximal information carried by B during the learning period is *p* then any instance of B has the truth condition *p*.)

DRET\* allows for errors but its naturalistic credentials are questionable. The trouble is that *learning* seems to be a semantic notion. Dretske may think that it is possible to characterize the learning period non-semantically but he can't just take this for granted. In any case, even if the learning period could be characterized naturalistically the account is implausible, at least for some beliefs. There are some beliefs that are learned in circumstances in which the information they carry is not the belief's content. For example, when a child learns to token a belief with a content about tigers by seeing pictures of tigers her belief states carry information about pictures although their content is about tigers. Dretske's account will end up assigning the wrong truth conditional contents to these beliefs.<sup>15</sup>

#### Optimal Conditions Accounts

A different way of specifying a belief's content is in terms of the information it would carry under epistemically optimal conditions (Stampe 1979, Stalnaker 1984, and Fodor 1990). The core idea of this approach is that there is a class of beliefs for which there are conditions- the epistemically optimal conditions- under which a person has the belief just in case it is true.

OPT) N(if B is a belief of kind K then there are epistemically optimal conditions C such that if B nomologically covaries with *p* then B given C then B's truth condition is *p*.)

So, for example, if for subject A there is a belief state B that under optimal conditions covaries with the presence of a red ball located in front of her then B's content is that there is a red ball in front of A. In this case appropriate optimal conditions are that A's eyes are open, she is attending to what she sees, the lighting is good, and so on.

OPT allows for errors since tokens of B that don't occur in epistemically optimal conditions need not be true. It also seems to supply truth conditions with normative force, at least if epistemically optimal is normative. But like Dretske's theory its specification of the meaning constituting conditions is not naturalistic. "Epistemically optimal" is clearly an intentional predicate. It is not at all clear that epistemically optimal conditions can be specified without reference to semantic notions. Different conditions are "optimal" for different beliefs. For example, epistemically

---

<sup>15</sup>. Dretske (1988) suggests a teleological characterization of the state tokens whose information fixes the beliefs content. His basic idea is that those instances of the belief state that produces behavior that is reinforced are the ones whose informational content fixes the belief's semantic content. While this is a naturalistic characterization of the class it is questionable whether it assigns appropriate contents. It is easy to imagine situations in which a false token of a belief produces behavior that is reinforced. For further discussion of Dretske's theory see Loewer (1985) and McLaughlin (1993).

optimal conditions for the perceptual belief that there is a red ball in the room include good lighting. But optimal conditions for the belief that there is a firefly in the room is that the lights are off. This example makes it obvious that the optimal conditions for acquiring true beliefs depends on the belief's content. Of course the naturalizer cannot appeal to the content of a belief in characterizing optimal conditions.

Not only are epistemically optimal conditions for a belief sensitive to the belief's content but for most beliefs, if they possess optimal conditions at all, these conditions involve other beliefs. Whether or not a person's belief state reliably covaries with a state of affairs depends on what other beliefs that person has. For example, a person who fails to believe that fossils are derived from once living organisms or who believes that the earth is six thousand years old will not reliably form beliefs about the age of a fossil. If there are optimal conditions for forming beliefs concerning the age of fossils those conditions will involve having certain beliefs and not having certain other beliefs. To assume that optimal conditions can be characterized naturalistically looks like it is to beg the naturalization problem rather than solve it.<sup>16</sup>

### Teleological Theories

Teleological theories propose to explain the truth conditional content of mental states, especially certain desires and beliefs, in terms of their biological functions. A crude teleological theory (CTT) for belief is:

CTT) N(if O is an organism and B is one of its belief states and it is B's biological function to carry the information that p then B has the truth conditions that p).<sup>17</sup>

The concept of a biological function is defined in terms of natural selection (Wright 1973, Neander 1991) along the following lines: it is the function of biological system S in members of species \$ to F iff S was selected by natural selection because it Fs.<sup>18</sup> S was selected by natural selection because it Fs just in case S would not have been present (to the extent it is) among members of \$ had it not increased fitness (i.e. the capacity to produce progeny) in the ancestors of members of \$.<sup>19</sup> So CTT says that if B was selected because it carried the information that p then B has

---

<sup>16</sup>. This point is developed in Loewer (1987) and more thoroughly in Boghossian (1991).

<sup>17</sup>. Some teleological accounts employ a more general characterization of informational account. S carries the information that p iff  $P(p/S \text{ occurs}) > P(p/S \text{ doesn't occur})$ .

<sup>18</sup>. Selection by conditioning (i.e. by reinforcement) also figures in accounts of function devised by some teleosemanticists (Dretske 1988).

<sup>19</sup>. For example, the biological function of the heart is to pump blood (not to make a thumping sound) since it is that property of pumping blood (not making a thumping sound) that accounts via natural selection for the presence of hearts. Notice that something may have the function to F even if it doesn't F or seldom Fs. It should be noted that it doesn't follow that every biological system that does something useful has that as its function (or that it has any function). Only those things that a system does that lead to an increase in fitness are its functions. So, for example, it is not obvious that e.g. certain cognitive abilities are the product of any function.

the truth condition that p.

CTT is naturalistic and allows for error. In fact it is compatible with almost all tokens of B being false since all that is required is that B was selected because it carried the information that constitutes its content and that could be so even if most past and all present tokens of B are not true. It also seems to supply truth conditional content with normativity. Just as a heart ought to pump blood B ought to be tokened only if it carries the information that p. There are however a number of problems with CTT. One is that it directly applies only to beliefs composed out of innate concepts since only beliefs involving innate concepts could possess a biological function. Perhaps the notion of biological function can be extended beyond features selected by natural selection; e.g. to features selected by operant conditioning and so enlarge the scope of CTT but that remains to be seen. A second problem is that it either fails to assign determinate contents or assigns contents that are much too thick-grained to be the truth conditions of beliefs. This problem has been discussed mostly with respect to the belief, or proto-belief of animals, especially a frog's (the hope being that extension to human's will come when the bugs are worked out).

Suppose that B is an internal state of a frog that is responsive to stimuli and that controls the frog's snapping behavior. Tokens of the state B in the frog's ancestors generally carried a great deal of information including, that flies are present, that small moving black things are present, that food is present, and so on. Further, since these various conditions were reliably co-instantiated in the environment in which the frog evolved they are all equally good candidates to be the information that it is the function of B to carry. So CTT implies either that B's content is indeterminate among components of the package or that its content is the whole package of information.

It is not clear whether this is an objection to teleological accounts since it is not clear what beliefs or desires, if any, frog's have. But it is an objection if teleological accounts are incapable of delivering more fine-grained contents than the one it apparently attributes to the frog. More elaborate theories of content that promise to solve this problem are due to Millikan (1984, 1986, 1989) and Papineau (1993). Both accounts, especially Millikan's, are rather elaborate. Here I will just briefly sketch Papineau's approach.

PAPB) If D is a desire and B a belief and p is the (minimal?) state of affairs whose obtaining guarantees that actions based on B and D satisfy D then B has the truth condition p.

If we suppose that the frog desires to catch a fly and that this desire together with B lead to its snapping then B's truth conditional content is the minimal state of affairs that will guarantee that snapping will result in catching a fly. In this case it is a belief with something like the content *if I snap then I will catch a fly*. Of course, PAPB is not naturalistic since it appeals to the concept of satisfying a desire and that is a semantic concept. Papineau attempts to remedy this by providing a naturalistic account of the contents of desires.

PAPD) If q is the minimal state of affairs such that it is the biological function of D is to operate in concert with beliefs to bring about q then D is the desire that q.

Papineau's idea is that if the desires of type D was selected because it contributed by acting in concert with beliefs to bringing about q then q is D's content. Let's suppose that the content of A's desire D is that she eats an apple. On a particular occasion D (together with beliefs) may cause the moving of A's hand, A's eating an apple, A's eating a fruit, and A's being nourished. Papineau suggests that the moving of the hand (to grasp the apple) isn't among D's functions since there are occasions when D was selected (A's ancestors who possessed D had increased fitness or D was reinforced in A) even though D didn't cause their hands to move. On the other hand, Papineau supposes that whenever D was selected D ate an apple, ate a fruit was nourished, etc. He suggests that the most specific of these features of the behavior which led to D's being selected is D's content; i.e. eating an apple.

There are a number of worries that one might have concerning Papineau's account. One is that it applies, at best, only to certain beliefs and desires. PAPB provides contents only to means-ends beliefs (although Papineau suggests how the account can be extended to other beliefs). Many desires could not have been selected for by natural selection since they are desires that possess impossible satisfaction conditions or desires for situations that have never obtained or have obtained too recently to be selected. It is hard to see how the desire to not have any children (or the desire that no one has any children) could have been selected for on the basis of bringing about its content. Perhaps these objections are not all that damaging if PAPD is intended just as a sufficient condition that applies to a certain class of desires. But then we will need a naturalistic specification of that class of desires. More damaging to PAPD is that possessing the function of bringing about x is not a sufficient condition for D's being the desire to bring about x. Suppose that D is the desire to eat an apple. It is compatible with this that there have been occasions when D lead not to apple eating but pear eating (some ancestors of A mistook pears for apples). It is plausible that eating pears (pears being as nutritious as apples) lead to increased fitness in which case D's function is to cause (together with beliefs) eating apples or pears. PAPD yield the result, contrary to our assumption, that D is the desire to eat apples or pears. There seems to be no reason why a desire could not have as its function causing, together with beliefs, some situation that differs from its content. If PAPD is incorrect then PAPB, even if it is correct, is no longer adequate as a naturalization of belief.

It is plausible that the human cognitive system contains subsystems that have the functions of producing states that bring about certain effects and producing other states that carry certain information (and work in concert with the first kind of state to produce effects). But there is no reason to suppose that these states are individuated exactly in the same way that beliefs and desires are. Truth conditional content seems much more determinate and fine-grained than anything that teleology is capable of delivering. This is made obvious by considering that there cannot be any selectional advantage for creatures whose beliefs are about rabbits over those whose beliefs are about undetached rabbit parts yet our contents are so fine-grained as to distinguish these beliefs. states.

#### Fodor's Asymmetric Dependence Theory

Fodor (1990b) proposed a variant of the causal (or informational)

account that is intended to be a naturalization of the reference of a simple Mentalese predicates. It appeals to the idea that the meaning constituting causes are those which, in a sense to be soon explained, resilient. It will simplify exposition of his theory to define two technical notions. The law  $Q \rightarrow C$  ( $Qs$  cause  $Cs$ ) *asymmetrically depends* on the law  $P \rightarrow C$  just in case if  $Ps$  didn't cause  $Cs$  then  $Qs$  would not cause  $Cs$  but if  $Qs$  didn't cause  $Cs$  then  $Ps$  would still cause  $Cs$ .  $C$  *locks onto*  $P$  just in case (i) it is a law that  $Ps$  cause  $Cs$ , (ii) there are  $Qs$  ( $=Ps$ ) that cause  $Cs$  and (iii) for any  $Q \neq P$  if  $Qs$  cause  $Cs$  then  $Qs$  causing  $Cs$  asymmetrically depends on  $Ps$  causing  $Cs$ .<sup>20</sup>  $P \rightarrow C$  is instantiated. This is supposed to give the result that Oscar's Mentalese "water" refers to  $H_2O$  and twin-Oscar's Mentalese "water" refers to  $XYZ$ . However, this addition may not be needed if the dependency relations concerning laws involving Oscar's and twin-Oscar's mental expressions are different. If  $C$  locks onto  $P$  then  $P \rightarrow C$  is resilient in that it survives the breaking of  $Q \rightarrow C$  for  $Qs$  other than  $P$ . Fodor's proposal then is:

ADT)  $N(\text{if } C \text{ locks onto } P \text{ then } C \text{ refers to } P)$

Suppose that it's a law that cows cause "Cow"s (or rather the word's mentalese counterpart), that other things also cause "Cow"s and that such causal relations asymmetrically depend on the cow  $\rightarrow$  "Cow" law. Then, according to ADT, "Cow" refers to cow. ADT handles the error and disjunction problems this way. Horses on a dark night can cause "Cow"s even though the horses on dark nights are not among the reference constituting causes of "Cow;" i.e. horses on a dark night  $\rightarrow$  "Cow"s depends on cows  $\rightarrow$  "Cow"s. If a horse caused "Cow" is a constituent of the belief "There is a cow" then the belief is false. Of course, this account of error is correct only if ADT is correct. If ADT is not correct then it may count some erroneous beliefs as true or some true beliefs as erroneous.

Fodor provides some commentary along with the theory. One point is that the law connecting a property to a predicate that refers to it is a *ceteris paribus* law. That is, it holds only as long as certain unspecified conditions obtain. Presumably this means that only under certain kinds of circumstances do cows actually cause A's mental representation "Cow." Presumably these conditions are that cows are perceptually salient to A, A's perceptual system is in good working order etc. A second point involves the dependence relation between causal laws. Sometimes Fodor says that it is a basic relation among laws that cannot be explained in other terms. But sometimes he explains it in terms of counterfactuals;  $Q \rightarrow C$  depends on  $P \rightarrow C$  just in case if  $P \rightarrow C$  had not obtained then neither would  $Q \rightarrow C$  have obtained. Fodor insists that the counterfactual be understood *synchronically* not *diachronically*. If A learned to recognize cows on the basis of pictures of cows then it may be that cow  $\rightarrow$  "Cow" depends diachronically on cow picture  $\rightarrow$  "Cow." That is, it's true that if there hadn't been causal connection between pictures of cows and A's "Cow"s there wouldn't be a connection between cows and A's "Cow"s. But Fodor thinks that synchronic dependence goes in the opposite direction. Once A has acquired "Cow" then cow  $\rightarrow$  "Cow" is more resilient than cow-picture  $\rightarrow$  "Cow." A third point is that the account of reference is *atomic*. By this is meant that it is metaphysically possible for A's Mentalese predicate  $C$  to lock onto  $P$  even if  $C$  bears no inferential or causal relations to any of A's other symbols or even if A's Mentalese vocabulary contains only the predicate  $C$ .

---

<sup>20</sup>. Fodor sometimes also adds the requirement that the law

Fodor welcomes this surprising feature of his account since he thinks that there are reasons to hold that inferential or causal relations among thoughts are not constitutive of the thought's semantic properties (Fodor and Lepore 1992).

There are two questions that need answers to evaluate Fodor's theory. Is it genuinely naturalistic? Is C locking onto P really a sufficient condition for C's referring to P? Answering these questions is made difficult by the fact that the central notions in Fodor's account - *ceteris paribus* laws and asymmetric dependence between laws - are technical notions that are not clearly defined.

There are two places to worry whether ADT is genuinely naturalistic. First, supposing that it is a law that  $P \rightarrow C$  then it is reasonable to believe that its *ceteris paribus* conditions include having and not having certain other intentional states. We noticed a similar point in our discussion of optimal conditions theories. Does this make  $P \rightarrow C$  non-naturalistic? Not necessarily. If the fact that  $P \rightarrow C$  is a law is naturalistically reducible then it too is a naturalistic fact. But do we have any reason other than the belief that semantic naturalism is true to think that  $P \rightarrow C$  is naturalistically reducible?

Second, and more worrying is whether the dependency relations that Fodor requires are naturalistic. These dependency relations are not themselves the subject of any natural science so Fodor cannot claim, as the teleo-semanticist does, that he is explaining a semantic notion in terms of a scientifically respectable notion; i.e. biological function. Further, it is not obvious that the synchronic counterfactuals that Fodor appeals to when explaining asymmetric dependence have truth conditions that can be specified non-intentionally. Why is Fodor so certain that the counterfactual (synchronically construed) *if cow  $\rightarrow$  "Cow" were broken then cow-picture  $\rightarrow$  "Cow" would also be broken* is true? Perhaps if the first law were to fail "Cow" would change its reference to cow-picture and so the second law would still obtain. If so, then while "Cow" refers to cow ADT would say that it refers to cow-picture.<sup>21</sup> Fodor cannot respond by saying that in understanding asymmetric dependence the counterfactual should be understood as holding the actual reference of "Cow" fixed since that would be introducing a semantic concept into the explanation of asymmetric dependence. I do not think that these points show that ADT is not naturalistic but rather to show that the burden is on Fodor to argue for the naturalistic credentials of the dependency relation. Fodor sometimes seems tempted to just take the dependency relation to be metaphysically primitive and declare that it is part of the natural order (Fodor 1991). One could see some irony in calling on such elaborate metaphysical notions to defend scientific naturalism.

Is the fact that C locks onto P sufficient for C to refer to P? It is difficult to answer this question without having a clear characterization of asymmetric dependence. The intrepid philosopher who thinks that she has devised a counter example to ADT runs the risk of being told by its inventor

---

<sup>21</sup>. Boghossian (1991) argues that either locking on is not sufficient for reference or is not naturalistic. His argument shows that to get the counterfactuals that underlie the locking on relation to come out right the similarity relation relative to which they are evaluated must take into account *semantic similarities*.

that she has gotten the dependency relations wrong. There are a number of such putative counter examples in the literature (Baker 1991, Boghossian 1991, Adams and Aizawa 1994, Gates 1995) and answers to the counter examples by Fodor (1991, 1994).<sup>22</sup> Instead of going into the details of these objections I will sketch two general worries about the account.

We attribute propositional attitudes to one another on the basis of folk psychological generalizations and general information about what people tend to believe, desire and so forth under certain circumstances. So, for example if A is a normal human being looking at a cow a hundred feet away then we expect A to believe that there is a cow in front of her. If in fact there is not a cow but a cleverly made cardboard cow facade then we expect A to at first believe that there is a cow but that when she moves closer to the cardboard cow and examines it she will cease to believe that there is a cow. Our ability to attribute beliefs, desires etc. to each other depends, at least in part, on generalizations like these. When testing a theory of intentionality we appeal to such generalizations. We ask whether it is possible for the putative naturalistic sufficient condition for A's believing that p to be satisfied while our folk psychological generalizations give the result that A doesn't believe that p? The problem I see with ADT is not that there are clear cases in which C locks onto P but C fails to refer to P. It is rather that, as far as I can see, ADT doesn't engage folk psychology. For all we know an assignment of beliefs to A employing ADT and an assignment employing the usual folk psychological principles may diverge radically. I am not arguing that they must or do diverge but that Fodor has provided no reason to think they don't. The worry isn't an idle one since it is not at all clear what asymmetric dependence has to do with our folk psychological principles of belief attribution. If ADT is to carry conviction we need some account of why it is that the contents it assigns will match those assigned by folk psychology.<sup>23</sup>

The second problem is the familiar one of the inscrutability of reference that seems to bedevil all naturalistic theories. If cow  $\rightarrow$  "Cow"s is a law then so is undetached cow part  $\rightarrow$  "Cow" (and laws involving various other properties metaphysically co-instantiated with cow (Quine 1960)). Neither one of these putative laws asymmetrically depend on the other since they hold in exactly the same possible worlds. So it looks like if a predicate locks onto any property it either locks onto all those properties that are metaphysically co-instantiated or onto the disjunction of all these properties (Gates 1995).

One response to the problem is to declare that properties like undetached cow part and temporal state of a cow etc. are not eligible to enter into laws and causal relations. Without a naturalistic justification of

---

<sup>22</sup>. One of Boghossian's counter examples to Fodor's theory is particularly persuasive. He imagines a natural kind concept K and laws  $X \rightarrow K$  and  $Y \rightarrow K$  where X and Y are different substances that are nomologically indistinguishable by us (they behave differently only in black holes. It may then be that neither of these laws asymmetrically depend on the other. Fodor's theory would have the consequence that K refers to the disjunction  $X \vee Y$ . But surely in the imagined situation K might refer only to X in virtue of the role it plays in physical theory.

<sup>23</sup>. This point is developed at length in different ways by Carl Gillett and Andrew Milne in dissertations at Rutgers.

this claim the response is another instance of borrowing from metaphysics to buy naturalism. Fodor, to his credit, has not taken this route but suggested an addition to ADT to cope with the problem (Fodor 1994). He argues that the inferential relations among sentences containing the predicate "Cow" will differ (for a thinker whose mentalese contains the truth functional connectives) depending on whether "Cow" refers to cow or to undetached cow part. By adding further conditions on the inferential relations borne by sentences to each other, he proposes to specify sufficient conditions for "Cow" to refer to cow (and no other property). The account is too complex to deal with in details here. I will just say that at best Fodor's proposal excludes some properties from being the references of "Cow" but fails to single out cow as the unique reference.

### Causal Role Semantics

Causal role (aka "conceptual role" and "inferential role") semantics (CRS) is another approach to naturalizing semantics that deserves mention, albeit only a brief one here. The mention is brief because although causal role semantics has been in the air for some time (Sellars 1974, Harman 1982, Field 1978, Loar 1981, Block 1986) no one has actually proposed a CRS that is naturalistic and assigns specific truth conditions to mental states or representations. The basic idea of CRS is that the semantic properties of a mental representation are partially constituted by certain causal or inferential relations between that and other mental representations. If only causal relations among mental representations are taken into account then at best CRS is an account of narrow content. To turn it into an account of broad content causal relations between mental representations and external items need to be added.

CRS should be distinguished from theories of interpretation like Davidson's (1984) that also ground truth conditions in causal relations among mental representations (or natural language representations) and external events. Davidson's theory of radical interpretation places constraints on the contents of a person's propositional attitudes. The most important one is that a correct theory of interpretation should assign mostly true beliefs. But the account is not a naturalization since the semantic concept *truth* is used in formulating the constraint.

CRS seems plausible as an account of what makes it the case that logical expressions possess their meanings. For example, it is plausible that dispositions to infer S from S#R and to infer S#R from S and from R are relevant to making it the case that "#" is conjunction. But elaborating this into a naturalistic sufficient condition of "#" to be conjunction is not completely straight forward. The most obvious difficulty is characterizing those causal relations that count as *inferences* without appealing to *truth*.

But as an account of the semantics non-logical expressions CRS seems inadequate. The immediate difficulty with CRS is that most of the actual causal roles of a person's sentence A do not seem necessary for it to possess its truth conditions. For example, A's Mentalese sentences "There is a cat" and "There is an animal" might have their usual truth conditions even though A has no disposition to infer the latter from the former. Given externalism, CRS cannot adequately specify sufficient conditions for a sentence to possess particular truth conditions solely in terms of its causal connections to other sentences. It will also need to invoke causal connections with external items. But this brings it back to the problem of specifying exactly which

causal connections are content constituting. CRS has made no distinctive contribution to answering this question naturalistically. The prospects for a naturalized CRS do not look good (Fodor and LePore 1992).

### Conclusion

None of naturalization proposals currently on offer are successful. We have seen a pattern to their failure. Theories that are clearly naturalistic (e.g. CCT) fail to account for essential features of semantic properties, especially the possibility of error and the fine-grainedness of content. Where these theories are sufficiently explicit we have seen that they are subject to counter-examples. In attempting to avoid counter examples semantic naturalists place restrictions on the reference (or truth condition) constituting causes or information. But in avoiding counter examples these accounts bring in, either obviously or surreptitiously, semantic and intentional notions and so fail to be naturalistic.

What if the semantic naturalist project fails? What if after centuries of searching the best philosophical minds fail to come up with naturalistic sufficient conditions that can be shown to be sufficient for intentional/mental properties? There are various responses. One is that the failure shows, or goes some distance in showing, that the supervenience claim of semantic naturalism is false. There are two reactions to this conclusion. One is that since naturalism is true it follows that there are no semantic properties (or they are never instantiated). This view, Semantic Elimitivism (Churchland 1981, Stich), preserves naturalism at the expense of semantics. The other response is to claim that there are semantic properties but they are metaphysically independent of natural properties. This view, Semantic Dualism (Brentano, Davidson (?) 1982, McDowell(?) 1994, Chalmers 2002), preserves semantics at the expense of naturalism. Neither option is very pretty. Elimitivism strikes some philosophers as self-refuting (Boghossian 1990, Brandom) and others (Fodor 1987) merely as obviously false in light of the success of folk psychological and cognitive science explanations that employ semantic concepts.<sup>24</sup>arises from the worry that

---

<sup>24</sup>. A sophisticated version of elimitivism maintains that *robust semantic properties don't exist (or are uninstantiated) but that* deflationary semantic predicates can be used to specify reference and truth conditions. A robust semantic property is a property that may enter into causal explanations and exists independently of our concepts and definitions. In contrast a deflationary truth predicate, "DT", for a language L is defined by providing a list of the conditions under which the predicate applies; e.g. "Snow is white" is DT iff Snow is white; "Snow is green" is DT iff snow is green; etc. More generally (p)("p" is DT iff p) where the quantifier is substitutional. An important feature of DT is that, unlike robust truth, it applies only to the language for which it is defined. There is no reason to suppose that items in the extension of a deflationary predicate have anything, in particular causal and explanatory powers, in common. It seems to follow that deflationary semantic notions cannot be employed in causal explanations or play an explanatory role in an intentional cognitive science. The attraction

deflationary truth and reference are too thin to do the work that we want done by semantic concepts. For recent discussion see Horwich (1990), Field (1986 and 1994). Semantic dualism raises the specter of semantic/intentional epiphenomenalism. If, as is widely believed, the natural sciences are causally complete then there seems to be no room for causation (of physical effects) in virtue of properties metaphysically independent of natural properties (Papineau 1992, Loewer 1995). So the situation seems to be that while there are reasons to worry that semantic naturalism might be false there are also reasons to doubt the alternatives. The semantic naturalist will resolve this paradox if he can produce a naturalization of semantic properties. That would be enough to quell doubts concerning semantic naturalism since we would then know that the gap between the semantic and the natural can be bridged.

Of course the failure of naturalization proposals to date does not mean that a successful naturalization will not be produced tomorrow. But another possibility, one that philosophers have recently begun to take seriously (e.g. McGinn 1993), is that while semantic naturalism is true we may not be able to discover naturalistic conditions that we can *know* or have reason to believe are sufficient for semantic properties; i.e. perspicuous semantic naturalism may be false. It may be that the naturalistic conditions that are sufficient for semantic properties are too complicated or too unsystematic for us to be able to see that they are sufficient. Or, it may be that there is something about the nature of semantic concepts that blocks a clear view of how the properties they express can be instantiated in virtue of the instantiation of natural properties. The situation can be compared with some other situations in which it has been proposed that there are facts that are beyond our epistemic grasp. I have in mind epistemic views concerning vagueness (Williamson), the view that phenomenal concepts are special direct recognitional concepts (Loar), and certain fundamental physical theories that posit facts that cannot be ascertained (e.g. Bohmian quantum mechanics). Of course these examples are very controversial and differ in various ways from the situation involving intentional/semantic predicates. In particular in each of these cases there are accounts (of varying plausibility) of why the posited realm of facts is epistemically inaccessible. As far as I know there is no clear proposal that explains why we cannot understand how semantic naturalism can be true even though it is true.<sup>25</sup> As of now we don't know whether semantic naturalism is true and, if it is true, we don't know whether we can know of any particular proposed naturalization that it is correct (though, as we have seen, we can know of some that they are incorrect).<sup>26</sup>

---

of deflationism (the view that the only instantiated semantic predicates are deflationary ones) is that it allows us to use semantic predicates for certain purposes (e.g. for infinite conjunction and disjunction) but is compatible with Naturalism. Skepticism concerning deflationism

<sup>25</sup>. Boghossian (1990) argues that belief holism (the fact that which situations are apt to cause one to acquire a particular belief depends on one's other beliefs) prevents us from certifying that any naturalistic condition on content constituting causes or information is correct.

<sup>26</sup>. I am grateful to Paul Boghossian, Jerry Fodor, Gary Gates, Carl Gillett, and Fritz Warfield for helpful discussion and (not always heeded) advice.

Barry Loewer  
Rutgers University

#### Bibliography

- Avramides, A. (199?), "Entry on Convention" in Companion to the Philosophy of Language Cambridge
- Baker, L. (1991), "Has Content Been Naturalized", in Meaning in Mind eds. Loewer, B. and Rey, G. Blackwell, pp17-32.
- Block, N. (1986), "Advertisement for a Semantics for Psychology", in Studies in the Philosophy of Mind, eds. P. French, T. Uehling, and H. Wettstein, Vol. 10 Midwest Studies in Philosophy Vol 10, University of Minnesota Press.
- Boghossian, P. (1989), "The Rule Following Considerations" Mind 83, pp. 507-49.
- Boghossian, P. (1990), "The Status of Content," Philosophical Review 99, 157-84.
- Boghossian, P. (1991) "Naturalizing Content," in Loewer and Rey (1991).
- Churchland, P. (1981), "Eliminative Materialism and the Propositional Attitudes," Journal of Philosophy 78, pp 67-90.
- Davidson, D. (1980), Essays on Actions and Events Clarendon Press
- Davidson, D. (1984), Inquiries into Truth and Interpretation, Clarendon Press.
- Dretske, F. (1981), Knowledge and the Flow of Information MIT Press
- Dretske, F. (1988), Explaining Behavior MIT Press
- Field, H. (1978), "Mental Representation," Erkenntnis 13, pp. 9-61. Field, H. (1972), "Tarski's theory of Truth," Journal of Philosophy, 69, pp. 347-75.
- Field, H. (1977), "Logic, Meaning, and Conceptual Role," Journal of Philosophy, 74 pp. 379-409.
- Field, H., (1986), "The Deflationary Conception of Truth", in Fact, Science, and Value, eds. McDonald, G. and Wright, C. Blackwell pp. 55-117
- Field, H. (1994), "Deflationist Views of Meaning and Content", Mind 103, pp. 249-285.
- Fodor, J. (1987), Psychosemantics: the Problem of Meaning in the Philosophy of Mind MIT Press.
- Fodor, J. (1990a), "Psychosemantics, or Where Do Truth Conditions Come From"

- in Mind and Cognition ed. W. Lycan, Blackwell pp 312-38.
- Fodor, J. (1990b), A Theory of Content and Other Essays MIT Press
- Fodor, J. (1991), "Replies" in Loewer and Rey 1991.
- Fodor, J. (1994), The Elm and the Expert MIT Press.
- Fodor and Lepore (1992), Holism: a Shopper's Guide Oxford: Blackwell.
- Gates, Gary (1995), "The Price of Information", forthcoming Synthese.
- Grice, P. (1957), "Meaning" Philosophical Review, 66, 377-88.
- Hale, B. and Wright, C. (199?), "Entry on Rule Following" in Companion to the Philosophy of Language Cambridge.
- Harman, G. (1982), "Conceptual Role Semantics" Notre Dame Journal of Formal Logic, 23, pp. 242-56.
- Horwich, P. (1990), Truth Blackwell.
- Kim, J. (1993), Supervenience and Mind Cambridge University Press.
- Kripke, S. (1972), Naming and Necessity Harvard University Press
- Kripke, S. (1982), Wittgenstein on Rules and Private Language, Harvard University Press.
- Lewis, D. (1969), Convention Harvard University Press.
- Loar, B. (1981), Mind and Meaning Cambridge: Cambridge University Press.
- Loewer, B. (1987), "From Information to Intentionality," Synthese 70, 287-317. Reprinted in Stich and Warfield (1994).
- Loewer, B. (1995), "An Argument for Strong Supervenience" in New Essays on Supervenience ed. E. Savellos, Cambridge.
- Loewer, B. and Rey, G. (1991), Meaning in Mind: Fodor and his Critics Oxford: Blackwell.
- McDowell, J. (1994), The Mind and the World Harvard University Press.
- McGinn, C. (1982), "The Structure of Content", in Thought and Object, ed. A. Woodfield, Oxford University Press.
- McGinn, C. (1993), Problems in Philosophy: The Limits of Inquiry Blackwell.
- McLaughlin, B. ed (1993), Dretske and his Critics Blackwell.
- Millikan R. (1984), Language, Thought, and Other Biological Categories, MIT Press.
- Millikan, R. (1986), "Thoughts Without Laws; Cognitive Science With Content." Philosophical Review 95, 47-80.
- Millikan, R. (1989), "Biosemantics," Journal of Philosophy, 86 pp 281-97. Reprinted in Stich and Warfield (1994).
- Millikan, R. (1991), White Queen Psychology and Other Essays for Alice MIT Press.
- Neander, K. (1991), "Functions as Selected Effects: The Conceptual Analyst's Defense" Philosophy of Sciences 58.
- Neander, K. (1995), "Misrepresentation and Malfunction"
- Papineau, D. (1993) Philosophical Naturalism Oxford:Blackwell.
- Pietroski, P. (1993), "Intentionality and Teleological Error," Pacific Philosophical Quarterly pp.23-38.
- Putnam, H. (1975) Mind Language and Reality (Philosophical Papers Vol II) Cambridge.
- Putnam, H. (1978), Meaning and the Moral Sciences Rutledge and Keegan Paul.
- Putnam, H. (1992), Renewing Philosophy Harvard University Press.
- Quine, W.V. (1960), Word and Object MIT Press.
- Schiffer, S. (1972), Meaning Oxford University Press.
- Sellars, W. (1974), "Meaning as Functional Classification", Synthese 61 pp. 64-79.
- Stalnaker, R (1984), Inquiry MIT Press.
- Stalnaker, R. (1991), "Semantics for the Language of Thought" in Loewer and Rey (1991) pp. 229-37.
- Stich, S. and Warfield, T. (1994), Mental Representation Oxford:Blackwell.
- Wright, L. (1973), "Functions", Philosophical Review 84.

