

Holism, Conceptual-Role Semantics, and Syntactic Semantics

WILLIAM J. RAPAPORT

*Department of Computer Science and Engineering, Department of Philosophy, and Center for Cognitive Science, State University of New York at Buffalo, Buffalo, NY 14260-2000, USA;
E-mail: rapaport@cse.buffalo.edu*

Abstract. This essay continues my investigation of ‘syntactic semantics’: the theory that, *pace* Searle’s Chinese-Room Argument, syntax *does* suffice for semantics (in particular, for the semantics needed for a computational cognitive theory of natural-language understanding). Here, I argue that syntactic semantics (which is internal and first-person) is what has been called a conceptual-role semantics: The meaning of any expression is the role that it plays in the complete system of expressions. Such a ‘narrow’, conceptual-role semantics is the appropriate sort of semantics to account (from an ‘internal’, or first-person perspective) for how a cognitive agent understands language. Some have argued for the primacy of external, or ‘wide’, semantics, while others have argued for a two-factor analysis. But, although two factors can be specified—one internal and first-person, the other only specifiable in an external, third-person way—only the internal, first-person one is needed for understanding how someone understands. A truth-conditional semantics can still be provided, but only from a third-person perspective.

Who knows only one thing knows not even that. A thing entirely isolated would be unknowable. There would be nothing to say of it, or any language for it. The reason for this has been obvious to people as different as Saint Thomas Aquinas and William James. Thomas said: “the soul is pleased by the comparison of one thing with another, since placing one thing in conjunction with another has an innate affinity with the way the mind acts.”¹ And James said: “the first thing the intellect does with an object is to class it along with something else.”² (Wills, 1991: 18)

The question of how one understands the language one thinks in does seem to be a peculiar one. . . . CRS [Conceptual Role Semantics] clarifies the situation. (Loewer, 1982: 310)

1. Syntactic Semantics

In a series of earlier essays (Rapaport, 1988, 1995, 1999, 2000b), I have set forth a theory of ‘syntactic semantics’ as a way of understanding how computers can think (and how John Searle’s (1980) Chinese-Room Argument objection to the Turing Test can be overcome). In the present essay, I explore the ramifications of this theory with respect to holism and conceptual-role semantics. After briefly rehearsing my arguments for syntactic semantics, I explore and defend conceptual-role semantics, and show that syntactic semantics *is* conceptual-role semantics.



Minds and Machines 12: 3–59, 2002.

© 2002 Kluwer Academic Publishers. Printed in the Netherlands.

Syntactic semantics has three basic theses (detailed explication and argumentation is to be found in the essays cited above):

- SS1. Semantics, considered as the study of relations *between* uninterpreted markers and interpretations of them, can be turned into syntax: a study of relations *among* the markers and their interpretations. This is done by incorporating (or ‘internalizing’) the semantic interpretations along with the markers to form a system of new markers, some of which are the old ones and the others of which are their interpretations. Hence, syntax (i.e., ‘symbol’ manipulation of the new markers) can suffice for the semantical enterprise (contra Searle).
- SS2. Semantics can also be considered as the process of understanding one domain (by modeling it) in terms of another. Call these the ‘syntactic’ and ‘semantic’ domains, respectively. This can be viewed recursively: If we are to understand one thing in terms of another, that other thing must be antecedently understood. Hence, the semantic domain can be treated as a (new) syntactic domain requiring a further semantic domain to understand *it*, in what Brian Cantwell Smith (1987) has called a ‘correspondence continuum’. To prevent an infinite regress, some domain must be understood in terms of itself. This base case of semantic understanding is ‘syntactic understanding’ (Rapaport, 1986): understanding a (syntactic) domain by being conversant with manipulating its markers, as when we understand a deductive system proof-theoretically (or, to anticipate, when we understand the language we think in, as Barry Loewer said).
- SS3. An internal (or ‘narrow’), first-person point of view makes an external (or ‘wide’), third-person point of view otiose for the task of understanding cognition (as opposed to the task of verifying correspondences between cognition and the external world).

2. Comparisons, Patterns, and Roles

Let us begin by exploring SS2 a bit. To understand a syntactic (or ‘new’, not-yet-understood) domain in terms of a semantic (or ‘given’, antecedently-understood) domain, one determines correspondences between them by making *comparisons*. The result of a comparison is a determination that the ‘new’ item ‘plays the same role’ in *its* (syntactic) domain that the corresponding ‘given’ item plays in *its* (semantic) domain. The two items are analogous to each other; a pattern seen in one domain has been matched or recognized in the other. *Each* item—new *and* given—plays a role in its respective domain. These roles are, in their respective domains, *syntactic* roles, that is, roles determined by relationships to other items in the domain. The *semantic* item’s role is its *syntactic* role in the ‘given’ domain. These relationships are not *cross-domain* relationships, but *intra-domain* relationships—that is, syntactic relationships, in Charles Morris’s (1938) sense.

In what sense are these roles ‘the same’? They *correspond* to each other. This means (1) that the two *domains* are both instances of a common pattern (which is understood syntactically) and (2) that the new and given *items* both map to the same item in the common pattern. (This general phenomenon is known as ‘unification’; see Knight, 1989) But then why not say that it’s the common pattern that is the proper *semantic* domain, rather than say that the semantic domain is the ‘given’ domain? Leo Apostel (1961: 2) suggested something like this: “If two theories are without contact with each other we can try to use the one as a model for the other or to introduce a common model interpreting both and thus relating both languages to each other.” Typically, however, one uses as the ‘favored’ semantic domain one that is ‘familiar’. If one *did* take the common pattern as the semantic domain, the question of ‘same role’ would arise again. But this time, there is no *other* common pattern, so there’s no regress. But *now* what counts is the mapping between the two domains—the syntactic domain and either the ‘given’ domain or the common pattern (it doesn’t matter which). That mapping must have certain features, namely, those characterizing semantic interpretation functions, such as being a homomorphism (cf. Rapaport, 1995, §2.2.2).

What is the role of an item in the common pattern? That’s a *syntactic* question, to which I now turn.

3. Conceptual-Role Semantics and Holism

3.1. OVERVIEW

I have just talked of pattern matching as the way to determine correspondences between two domains. When two patterns, *A* and *B*, match, the result is a determination that a part of pattern *A* ‘plays the same role’ in pattern *A* that a corresponding part of pattern *B* plays in pattern *B*. That role, I suggested, was the part’s syntactic role in its own domain—a role determined by the part’s internal relationships to the other parts of the pattern. According to SS2, this is where semantics ‘bottoms out’: in the syntactic understanding of a (syntactically specified) domain, where what counts for a term’s meaning is its syntactic role.

This kind of semantics has come to be called ‘conceptual-role semantics’ or ‘inferential-role semantics’ (on the distinction between these terms, see §§4.2, 6.7, 6.7.2.1). Conceptual-role semantic theories are almost always associated with holistic theories of semantics. Both have lately come under sustained attack from Jerry Fodor and Ernest Lepore (1992), who argue that there are no good arguments for holism. That may be, yet I find holism attractive. I take my task in this essay not so much to argue *for* it (I doubt that I could find an argument stronger than those refuted by Fodor and Lepore) as to paint an attractive picture of holism and conceptual-role semantics and to clarify that picture in the light of the critiques of holism and conceptual-role semantics.

3.2. SNePS AND CASSIE³

For convenience and perspicuousness, I will use as a model a knowledge-representation, reasoning, and acting system that consists of a vast propositional, intensional, semantic network with ways of incorporating sensory images among its nodes. The nodes will represent individual concepts, properties, relations, and propositions, and the connecting arcs will structure atomic concepts into molecular ones (including structured individuals, propositions, and rules).

The specific knowledge-representation and reasoning (KRR) system I will use to help fix our ideas is the SNePS *Semantic Network Processing System* (Shapiro, 1979; Shapiro and Rapaport, 1987, 1992, 1995). But you can think in terms of other such systems, such as (especially) Discourse Representation Theory,⁴ description logics (e.g., the KL-ONE family),⁵ Conceptual Dependency,⁶ or Conceptual Graphs.⁷ Or, if you prefer, you can think in terms of a connectionist system: there is no loss of generality in focussing on a *symbolic* system such as those just mentioned, for a connectionist system is just as computational—as syntactic—as a classical symbolic system (Rapaport, 2000a). So, a connectionist system that passed the Turing Test would make my points about the syntactic nature of understanding equally well.

As a *knowledge-representation system*, SNePS is (1) symbolic (or ‘classical’; as opposed to connectionist), (2) propositional (as opposed to being a taxonomic or ‘inheritance’ hierarchy), and (3) fully intensional (as opposed to (partly) extensional). As a *reasoning system*, it has several types of interrelated inference mechanisms: ‘node-based’ (or ‘conscious’), ‘path-based’ (generalized inheritance, or ‘subconscious’), ‘default’, and belief revision. Finally, it has certain *sensing and effecting mechanisms*, namely: natural-language competence (by which I mean both understanding and generation; see Shapiro and Rapaport, 1991), and the ability to make, reason about, and execute plans. Such, at least, is SNePS in principle. Various implementations of it have more or less of these capabilities, but I will assume the ideal, full system.

SNePS has two kinds of nodes: base and molecular. Base nodes have no arcs emanating from them (e.g., node B1 in Figure 1); molecular nodes do have outgoing arcs (e.g., node M1 in Figure 1; for more on this distinction, see §6.7.2.2.).

One special case of base nodes is a *lex* node: A *lex* node is labelled with an (English) expression, usually a single word, and serves to link the network to the lexicon that is part of the natural-language-competence component. Typically, *lex* nodes have only one incoming arc, labeled ‘*lex*’. Let *m* be a molecular node (technically, a ‘structured individual node’) with a *lex* arc to a *lex* node (technically, a ‘sensory node’) labeled *w*. Then, the ‘meaning’ of *m* is the concept (technically, ‘Meinongian entity’) expressed by uttering *w*. (E.g., in Figure 1, node M1 represents a concept that the system expresses with the English word ‘round’; see Shapiro et al., 1996: 5–6; Rapaport et al., 1997, §3.1.)

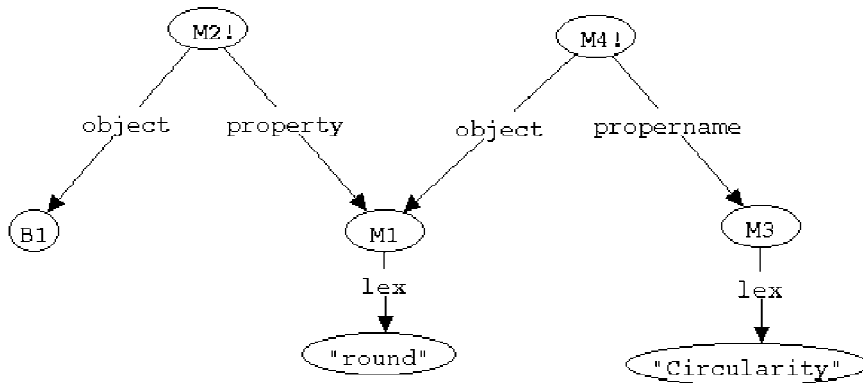


Figure 1. Something is round and the concept of roundness is named ‘Circularity’. M2 = (Cassie’s belief that) B1 is round; M4 = (Cassie’s belief that) M1 is named ‘Circularity’.

One special case of molecular nodes is a rule node. A rule node is the SNePS counterpart of a quantified formula, and is used in rule-based inference (e.g., node M5 in Figure 2). Despite its name, it is *not* a counterpart of a rule of inference, since the latter cannot be expressed in any inference system (cf. Carroll, 1895), whereas rule nodes are expressed. (The SNePS counterpart of rules of inference are embodied in the SNePS Inference Engine. For more details on rule nodes and other matters of SNePS syntax and semantics, see Rapaport et al., 1997, §3.1.)

That SNePS is propositional rather than taxonomic merely means that it represents everything propositionally. Taxonomic hierarchical relationships among individuals and classes are represented propositionally, too. Systems that are, by contrast, primarily taxonomic have automatic inheritance features; in SNePS, this

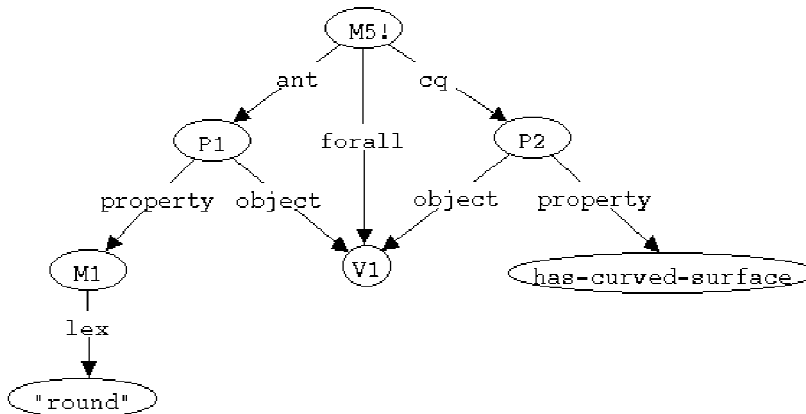


Figure 2. Round things have curved surfaces. M5 = (Cassie’s belief that) $\forall V1$ [Round(V1) \rightarrow Has-Curved-Surface (V1)], where, for the sake of the example, ‘Has-Curved-Surface’ is not—but could be—further analyzed. (Node M1 here is the *same* node as node M1 in Figure 1).

is generalized to path-based inference. Both events and situations can also be represented in SNePS.

But SNePS is intensional, and therein lies a story. To be able to model the mind of a cognitive agent, a KRR system must be able to represent and reason about intensional objects, i.e., objects not substitutable in intensional contexts (such as the morning star and the evening star), indeterminate or incomplete objects (such as fictional objects), non-existent objects (such as a golden mountain), impossible objects (such as a round square), distinct but coextensional objects of thought (such as the sum of 2 and 2, and the sum of 3 and 1), and so on. We think and talk about such objects, and therefore so must any entity that uses natural language.

We use SNePS to model, or implement, the mind of a cognitive agent named ‘Cassie’.⁸ Cassie’s ‘mind’ consists of SNePS nodes and arcs; i.e., SNePS is her language of thought (in the sense of Fodor, 1975). If she is implemented on a Sun workstation, then we might also say that she has a ‘brain’ whose components are the ‘switch-settings’ (the register contents) in the Sun that implements the nodes and arcs of her mind.

We will say that Cassie can represent—or think about—objects (whether existing or not), properties, relations, propositions, events, situations, etc. Molecular nodes that represent Cassie’s beliefs are indicated by an ‘assertion’ operator (‘!’; see node M2 of Figure 1). Thus, all of the things represented in SNePS when it is being used to model Cassie’s mind are objects of Cassie’s thoughts (i.e., Meinongian objects of Cassie’s mental acts); they are, thus, intentional—hence intensional—objects. They are not extensional objects in the external world, though, of course, they may bear relationships to such external objects.

I cannot rehearse here the arguments I and others have made elsewhere for these claims about SNePS and Cassie. I will, however, provide examples of SNePS networks in the sections that follow. (For further examples and argumentation, see, e.g., Maida and Shapiro, 1982; Shapiro and Rapaport, 1987, 1991, 1992, 1995; Rapaport, 1988b, 1991; Rapaport and Shapiro, 1995.)

Does Cassie understand English?⁹ If so, how? Searle, of course, would say that she doesn’t. I say that she does—by manipulating the symbols of her language of thought, viz., SNePS. Let’s turn now to these issues.

3.3. THE MEANING OF A NODE

The Gary Wills quotation that I cited at the beginning nicely expresses the core ideas behind holism and conceptual-role semantics. In SNePS terms, one cannot say anything about an isolated node without thereby enlarging the network and de-isolating the node. As such a process continues, the network grows. This is how holistic conceptual-role semantics begins. Since all that is initially known about the isolated node is now expressed in the rest of the network, the node’s ‘meaning’ is determined by its location or role in that entire network (Quillian, 1967, 1968). Nodes that are very distant from the original one may have little to do directly

with its meaning or role. But they will have something to do with other nodes that, eventually, directly impact on that original node (or are impacted on by it). To use an older terminology, they may be part of that node's 'connotations'. (Hill, 1994, 1995 provides a formal interpretation of this.)

The larger the network, the more meaning its nodes have—that is, the more can be said about them—and the larger their roles are. In *From Folk Psychology to Cognitive Science* (1983), Stephen Stich has argued that a person with a single, isolated 'belief' does not really have any beliefs. I would prefer to say that the more beliefs one has, the more each belief means. Such an isolated belief *is* a belief, but not one that has much of a role to play. (Similarly, as I pointed out in 'Syntactic Semantics' (1988), linguists who build syntactic and semantic theories from studies of isolated sentences would also do better to look at connected discourse.)

Isolation—even a complex network that is isolated from the rest of the network—is a barrier to comprehension. A patient can convey, without understanding it, a message from a doctor to a dentist, both of whom *will* understand it, because the medically ignorant patient cannot link the message to his or her own semantic network, while the medical personnel can link it to theirs (cf. Rapaport, 1988: 126n16). Or consider a fax machine: it takes text, converts it to electronic signals, and reconverts these to text. Yet—like the patient—it has no 'knowledge' of the text. The patient and the fax seem to be in a Chinese Room. But if the fax converted the text to, say, ASCII code, which could then be linked to a knowledge base, we might have an 'intelligent' fax machine, thus escaping this Chinese Room. It is the internal *links* that count; isolation doesn't yield understanding.¹⁰

It is always, of course, a matter of degree. If 'an elephant is so he can have a trunk' (Spencer Brown, personal communication), and that's all we know about elephants or their trunks, then all we know about their trunks is that they can be had by elephants. But as our knowledge of elephants (and their trunks) enlarges, we come to understand more and, no doubt, to express it more informatively, less obviously circularly:

[T]he problem of 'genuine semantics' . . . gets easier, not harder, as the K[nowledge]B[ase] grows. In the case of an enormous KB, such as CYC's, for example, we could rename all the frames and predicates as G001, G002, . . . , and—using our knowledge of the world—reconstruct what each of their names must be. (Lenat and Feigenbaum, 1991: 236.)

Carnap said as much years earlier, in his example of a railroad map (in *The Logical Structure of the World*). There, he showed how to describe any object in a given domain in terms of the other objects, without any external "grounding" (Carnap 1928, §14, pp. 25–27; cf. Rapaport, 1988: 111).¹¹

Let's now look at two of the major conceptual-role semantic theories, the early, influential one of Wilfrid Sellars and the more recent one of Gilbert Harman.

4. Sellars's Theory of Language Games

In a series of papers that became chapters of his *Science, Perception and Reality* (1963), Wilfrid Sellars spelled out a classic theory of conceptual-role semantics.¹²

In 'The Language of Theories' (1959 [1963: 109–113], §§11–18), he distinguishes a variety of kinds of meaning:

Meaning as translation:

- 'round' means *circular*; [I would prefer to say that 'round' means *round*.]
- 'cheval' means *horse*.

Meaning as sense:

- 'round' expresses the concept Circularity; [I would prefer to say that 'round' expresses the concept Roundness.]
- 'cheval' expresses the concept Horsekind.

Meaning as naming:

- 'round' names the concept Circularity; [I would prefer to say that 'round' names the concept Roundness.]
- 'cheval' names Man O'War.

Meaning as connotation:

- 'cheval' connotes the property of having four legs;
- 'Parigi' connotes the property of being the capital of France.

Meaning as denotation:

- 'round' denotes circular things. [I would prefer to say that 'round' denotes round things.]

Conceptual-role semantics is about meaning as translation, though it can make room for all the others (except possibly the last—but see Rapaport, 1995 and Rapaport, 1996, Ch. 3). To see how, and to fix our ideas a bit, let's consider a computational model of this.

4.1. A COMPUTATIONAL MODEL

Suppose Cassie (our computational cognitive agent) hears Oscar say that something 'is round'. As a result, she builds the semantic network dominated by node M2 of Figure 1. (A node *dominates* another node if there is a path of directed arcs from the first node to the second node.) Insofar as Cassie maps Oscar's utterance or use of 'round' to her own 'round' node, she is understanding Oscar by translating

his utterances into her semantic network. (If she has never heard ‘round’ before, she’ll create a new node on which to map Oscar’s utterance; it’s still translation.)

I would say, however, that Cassie’s lex node labeled ‘round’ expresses the *concept* (viz., M1) at the tail of the lex node. Thus, in Figure 1, node M1 is Cassie’s concept of roundness (or circularity, to use Sellars’s somewhat misleading locution). If Cassie wanted to talk about that concept (and to say more than that something (viz., B1) is round), she could name it; node M3 would represent its name, expressed as ‘Circularity’. (Here, I differ a bit from Sellars.)

Connotation can be accounted for, in part, as follows: Suppose Cassie learns that round things have curved surfaces, so she extends her semantic network as in Figure 2 (where node M5 represents what she has just learned). Here, part of the connotation of ‘round’ is given by rule node M5 (as well as, perhaps, by M2 and M4, and so on, throughout the full network).

Denoting, however, is a relation that Cassie cannot deal with for herself. It is an external relation, accessible only to another cognitive agent: Oscar could assert that Cassie’s ‘round’ denotes α . We have the situation shown in Figure 3. According to Sellars, Cassie’s word ‘round_C’ denotes some circular thing, α ; so denotation, for Sellars, is a relation between a word and an external object. As such, it is not accessible to Cassie. (By the way, presumably there are also relations, equally

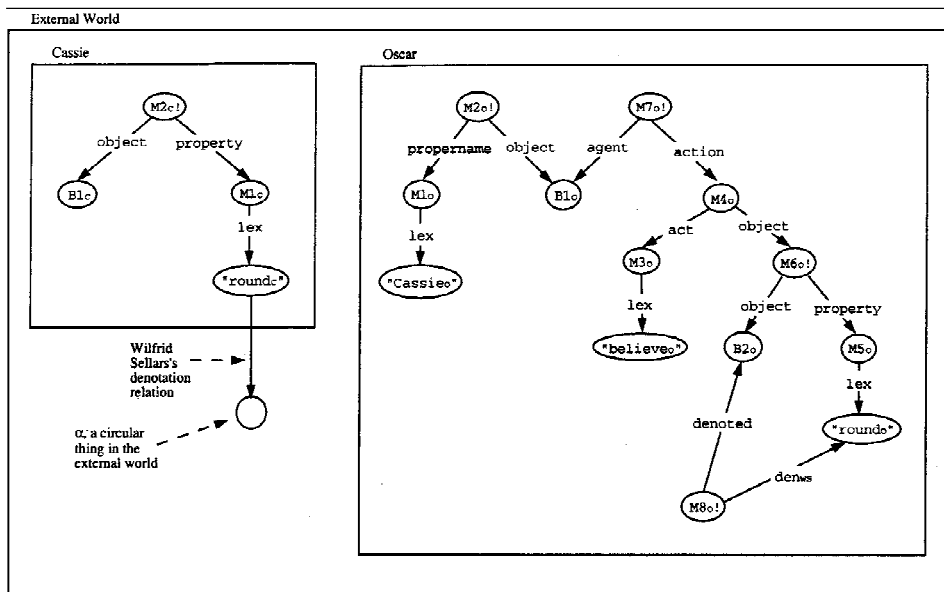


Figure 3. Cassie’s and Oscar’s representations that something is round. In the external world, Cassie’s node “round” denotes-in-Wilfrid-Sellars’s-sense α . In Cassie’s belief space, $M2_C =$ (Cassie’s belief that) $B1_C$ is round. In Oscar’s belief space, $M2_O =$ (Oscar’s belief that) $B1_O$ is named ‘Cassie_O’, $M7_O =$ (Oscar’s belief that) $B1_O$ believes that $M6_O$, $M6_O =$ (Oscar’s belief that) $B2_O$ is round, $M8_O =$ (Oscar’s belief that) ‘round’ denotes-in-Wilfrid-Sellars’s-sense $B2_O$.

inaccessible to Cassie, between α and her *concept* of roundness, viz., $M1_C$, and between α and her concept of α , viz., $B1_C$.) From Oscar's point of view (not much different from *our* point of view with respect to Cassie), Cassie believes that something (which Oscar represents as $B2_O$) is round, and Oscar can believe that Cassie's word 'round' (actually, *Oscar's* representation of her word) denotes (in Sellars's sense) the object (that Oscar believes) that Cassie believes is round, viz., $B2_O$. (Again, presumably, there are relations, equally inaccessible to Oscar, between the following pairs: 'round_O'/'round_C', $B2_O/B1_C$, $B2_O/\alpha$, $M5_O/M1_C$, and $M6_O/M2_C$.)

What can we say about statements like the following?

1. '⌈ x ⌋ means y '.
2. '⌈ φ ⌋ is true'.
3. '⌈ x ⌋ is about y '.

I'd say first that they're missing a parameter or two. The statements should really be, respectively:

1. Cognitive agent C 's use of '⌈ x ⌋ means y for cognitive agent O .
2. Cognitive agent O believes that cognitive agent C 's utterance or belief that '⌈ φ ⌋ is true'.
3. Cognitive agent C 's use of '⌈ x ⌋' is about what cognitive agent O refers to as '⌈ y ⌋'.

So, let me answer the question from Oscar's point of view:

1. For Oscar to say that Cassie's use of '⌈ x ⌋ means y is to say that Cassie's use of '⌈ x ⌋' plays the same role in *her* belief system that '⌈ y ⌋' plays in his (Oscar's).
2. For Oscar to say that Cassie's utterance of '⌈ φ ⌋ is true is to say that he endorses her utterance of '⌈ φ ⌋'; that is, it is to say that he believes it (too); cf. Rapaport et al. (1997). As Sellars puts it,

In general, when I commit myself to

(w) S is a true sentence (of L)

I am committing myself to asserting either S itself (if I am a user of L) or a translation of S into the language I do use. (Sellars 1955 [1963: 354], §78.)

3. For Oscar to say that '⌈ x ⌋ is about y is for him to say that he interprets '⌈ x ⌋' by y , where both are represented by nodes in *his* network: '⌈ x ⌋' is represented by a syntactic, or linguistic, node; y is represented by a semantic, or non-linguistic, node. If Oscar wants to say what his *own* word '⌈ x ⌋' means, he must do it in that way, too: asserting a link between it and some other fragment of his network. (This is the 'internalization' of semantic interpretations mentioned in SS1.)

4.2. REFLECTIONS ON “REFLECTIONS ON LANGUAGE GAMES”.

Sellars’s essay ‘Reflections on Language Games’ (1955/1963) is relevant to syntactic semantics in several ways (see also Rapaport, 1996, §§7.4.2.2, 7.6.2). Here, I want to concentrate on the syntactic nature of his conceptual-role semantic theory.

For Sellars, to use a language is to do certain actions in certain circumstances—presumably, for example, to utter certain expressions in certain circumstances—and this is to be viewed as making ‘moves’ in a ‘language game’ whose ‘positions’ are ‘awareness’ ‘of propositions, properties, relations, demands, etc.’ (pp. 324, 327, §§10, 16). There are three kinds of such moves (p. 328, §§19–23):

1. ‘language-entry transitions’ from observations of the external world to positions in the language game (that is, input, in which the position ‘means’ the observation; cf. p. 329, §22);
2. ‘moves’, or inferences, between positions in the language game (that is, relations among sentences);
3. ‘language-departure transitions’ from ‘ought-to-do’ positions to actions (that is, output, in which the position ‘means’ the action) (cf. p. 329, §23).

In terms of Cassie, language-entry transitions occur when she finds or builds a node in her semantic network as a result of something she hears, reads, or perceives, and language-departure transitions occur when she utters something as a result of an intention to speak or when she performs an action as a result of an intention to act.¹³ The internal, inferential moves correspond to *any and all internal processing* of the semantic network. They need not all be ‘inference’ in any strict logical sense. For this reason, I prefer the term ‘*conceptual-role semantics*’ to ‘*inferential-role semantics*’. (Cf. §6.6). For more on this distinction, see Récanati, 1995: 214). Of course, the input positions could be established in other ways (e.g. by direct manipulation by a ‘computational neuroscientist’, in the case of Cassie, or by Wilder-Penfield-like stimulation, in the case of real brains). For instance, Sellars also allows ‘free’ positions: sentences that are neither the result of internal, inferential moves nor of observations. Roughly, these would be axioms or ‘primitive’ beliefs: sentences taken on faith (p. 330, §25). And the output positions need not result in (successful) action (as long as the system *believes* that it does—cf. the blocks-world robot of Rapaport, 1995, §2.5.1).

To thus “speak of a language as a game with pieces, positions, and moves” is to treat it purely syntactically. “But must we not at some stage recognize that the ‘positions’ in a language *have meaning* . . . ?” (p. 332, §30). This is the key issue. Note, however, that for Sellars it would not be the *pieces* that are to ‘have meaning’, but the *positions*: “As I see it, abstract singular terms such as ‘redness’ . . . and ‘that Chicago is large’ are to be construed, in first approximation, as singular terms for players of linguistic roles . . . ” (Sellars, 1961 [1963: 204]). In Figure 1, the *term* ‘Circularity’ is a proper name for a concept, viz., M1, and it is the concept that *is* the role. What *plays* the role is the term ‘round’. Strictly speaking, then, we could

say that, for Cassie, the meaning of ‘round’ is node M1, whose *role* is specified by its location in the network. For Sellars, syntax suffices: “To say that ‘“rot” means *red*’ is not to describe ‘rot’ as standing ‘in the meaning relation’ to an entity *red*; . . . ” (p. 332, §31). “Positions” do not have an extensional “meaning”:

. . . the German expression ‘Es regnet’ . . . *means* it is raining. . . [I]n saying this . . . , one is not saying that the pattern ‘Es regnet’ plays a certain role in the pattern governed behaviour to be found behind the Rhine. But it would be a mistake to infer from these facts that the semantical statement ‘“es regnet” means *it is raining*’ gives information about the German use of ‘Es regnet’ which would *supplement* a description of the role it plays in the German language game, making a *complete* description of what could otherwise be a partial account of the properties and relations of ‘Es regnet’ as a meaningful German word. (p. 332, §31.)

Although there is a non-syntactic, externally semantic dimension to meaning, it has nothing to do with the language game. Cassie’s (internal) ability to use language is syntactic (and so Searle’s Chinese-Room Argument fails). That is, semantics is *not* a correspondence between language and the world. But semantics *is* a correspondence between two *languages*: between the speaker’s language and the third-person, external observer’s language (and perhaps that observer’s concepts, too): “To say that ‘“rot” means *red*’ . . . is to use . . . the semantical language game . . . for bringing home to a *user* of ‘red’ how Germans use ‘rot’” (p. 332, §31). English-speakers understand a German-speaker’s use of ‘rot’ as their (i.e., the English-speakers’) concept *red* (i.e., as the concept they express with ‘red’). This is semantics in the classic sense: The English-speaker uses a model for interpreting the German-speaker’s utterances. But the model is just the English-speaker’s own language game—a syntactic system.

To say that ‘“rot” means *red*’ . . . conveys no information which could not be formulated in terms of the pieces, positions, moves, and transitions (entry and departure) of the German language game. (p. 332, §31.)

That is, it conveys no information about ‘rot’ that could not be thus formulated. But suppose that an English speaker wonders what ‘rot’ means and is told that it means *red*. The English speaker now has nodes representing the German word ‘rot’ and the concept it expresses; and the English-speaker maps these—internally—to the nodes representing the English word ‘red’ and the concept *it* expresses. Thus, all of the information conveyed by the ‘rot’-means-*red* sentence can “be formulated in terms of the pieces, positions, moves, and transitions . . . of the *English* language game”. In either case, it’s purely syntactic.

As for correspondences between language and the world, Sellars discusses a cousin of the symbol-grounding problem (Harnad, 1990; cf. Rapaport, 1995, §3.2.4) under the rubric ‘prelinguistic concepts’:

Now there appear to be two possible lines that can be taken with respect to such ur-concepts:

- (1) They are interpreted as a structure of symbols and, hence, *in our broader sense*, as a *language*. . . . [A] regress is lurking which can be stopped only by admitting that the meaningfulness of at least one symbolic system is not clarified by the idea of obeying semantical rules.
- (2) As a second alternative, the ur-concepts may be conceived as pre-symbolic abilities to recognize items as belonging to *kinds* (pp. 334–335, §37.)

Possibility (2) is the Lakoff (1987)–Harnad (1990) alternative, which Sellars rejects on the grounds that it commits the homuncular fallacy. Possibility (1) is the purely syntactic view expressed in thesis SS2, above. To clarify the ‘meaningfulness’ of such a symbolic system, we need internal—syntactic—understanding.

Sellars urges a distinction between ‘bishop’ in chess and ‘piece of wood of such and such shape’ (p. 343, §56), and he then elaborates on possibility (1):

. . . I might learn to respond to the move-enjoining sentence ‘Sellars, advance your king’s pawn!’ as I would to ‘Sellars, shove this piece of wood two squares forward!’ (p. 344, §57.)

Compare the Chinese Room: “shoving a piece of wood forward” is the rule-book’s translation of the meaningless squiggle “advance your king’s pawn”. Perhaps, though, shoving that piece forward just *is* advancing one’s pawn, in the same way that talking of certain chemical structures just *is* talking of mathematical lattices (Rapaport, 1995, §2.5.1; I make this sense of “is” more precise in Rapaport, 1999). In any event, Sellars rejects it:

But while this *might* be the description of learning to apply the rule language game . . . , it would make the connection between expressions such as ‘bishop’ . . . in chess language and the expressions in everyday language which we use to describe pieces of wood, shapes, sizes, and arrangements much more ‘external’ than we think it to be. For surely it is more plausible to suppose that the piece, position, and move words of chess are, in the process of learning chess language, built on to everyday language by *moves* relating, for example, ‘x is a bishop’ to ‘x is a ♖-shaped piece of wood’ In other words, chess words gain ‘descriptive meaning’ by virtue of *syntactical relations* to ‘everyday’ words. (p. 344, §58.)

As I have urged with respect to the Chinese-Room Argument (Rapaport, 1988, 1995, 2000b), pulling the semantic rabbit out of the syntactic hat is no trick—it’s all done with *internal* links. My understanding of ‘bishop’ (or Searle-in-the-room’s understanding of a Chinese squiggle) is *not* provided by an external link to a ♖-shaped piece of wood, but by an *internal, syntactic* link to my internal representation of such a ♖-shaped piece of wood.

The fundamental thesis of conceptual-role semantics, as formulated by Sellars, is that

statements of the form

‘ . . . ’ means — (in L)

are incorrectly assimilated to relation statements. . . . [Rather,] they convey . . . the information that ‘ . . . ’ plays the role in L which ‘—’ plays in the language in which the semantical statement occurs. (pp. 354–355, §80.)

Of course, if the semantic language *is* L, the meaning of ‘ . . . ’ would have to be given in terms of the role it plays in L, by specifying its location in the network—its position in the game.

5. Harman’s Theory of Conceptual-Role Semantics

Let’s now have a look at Harman’s variations on Sellars’s theme. In a series of papers, Gilbert Harman has advocated a Sellarsian conceptual-role semantic theory *almost* all of which is congenial to the view I am presenting (Harman, 1974, 1975, esp. pp. 283–284; 1982; 1987; 1988). The issue can be approached by asking whether an internal, conceptual-role semantics based on *translating* one language into another is all that is needed to explain our knowledge of the semantics of language, or whether an external, referential and truth-conditional theory plays a role (if you’ll excuse the expression) (Harman, 1974: 1).

5.1. INTERNAL VERSUS EXTERNAL SEMANTIC THEORIES

I called the latter kind of theory ‘external’, but it is actually both internal and external; that is, it must be a bridge theory that links an internal syntactic domain with an external semantic domain. *Perhaps* such a theory could tell us something about the denotations of terms and the truth values of sentences. But, of course (cf. Rapaport, 1988, 1995, 2000b), since the cognitive agent has no access to the denotations or states of affairs themselves, a theory of truth tells the *agent* nothing. It is simply not available to the agent, who is restricted to the internal point of view (cf. thesis SS3, above). Now, as Harman notes, theories of truth do shed light on meaning—consider possible-worlds model-theoretic semantics for modal logics, clearly a major intellectual achievement. But note, first, that such theories are addressed to professional philosophers and cognitive scientists, who are external observers: Oscar can use such a theory to understand the relation of Cassie’s language to the world, but he doesn’t use the theory when he understands Cassie in everyday conversation. Second, truth theories are correspondences between language and a *model*, not between language and *the world* (see the discussions of the gap between models and the world, in Smith, 1985 and Rapaport, 1995, §2.5). So

they themselves are translations: between the language playing the syntactic role and the language of the model.

There are two other possible roles for truth theories or external links. One, relevant to Sellars's 'entry' and 'departure' rules, we'll come back to shortly (§5.3). The other is the role of truth in logical inference, Sellars's internal 'moves': 'logical implication is a matter of truth and logical form' (Harman, 1974: 11). But here, truth is only a sort of place holder: Logical implication must preserve truth, but no claims are ever made about actual truth *values*, nor need they be. The rules of inference of a syntactic system are themselves purely syntactic (cf. Rapaport, 1995, §2.2). They need not—indeed, *do* not—mention truth. In a given system, some rules might be preferable to others (they can be justified) because they *preserve truth*. That plays a role with respect to which rules to choose, but not in the actual working of the rules. Indeed, that's the whole point of syntactic systems: we devise them in order to talk about truth, so we want them to *represent* truths. The world (together with its objects, relations, states of affairs, and truths) is one thing; the language (with its corresponding terms, relation symbols, wffs, and rules of inference and theorems used to discuss the world) is another. We want language and the world to *correspond*; they don't intersect. (Well, actually they do, of course: The language is part of the world. But that fact is ignored when the language is used to describe (the rest of, or some other part of) the world.)¹⁴

From the internal, first-person point of view, all that we *can* deal with is the syntactic theory. And, if all we're dealing with is the syntactic theory, we don't need truth at all. Or, rather, *Cassie* doesn't need it, and can't have it anyway, and *Oscar* (who studies *Cassie*'s language-use from the external, third-person point of view) has access to truth only as a correspondence among *beliefs* (cf. Harman, 1974: 9): Oscar translates *Cassie*'s utterances into his own semantic network. If he tries to say what *is* true, all he can do is to say what *he* believes: If he didn't believe it, he wouldn't try to claim that it's true. That is, for Oscar to say that φ is true is just for him to say that (he believes that) φ . For Oscar to say that what *Cassie* said is true is also just for him to say that he believes what *Cassie* said (cf. Roberts and Rapaport, 1988; Rapaport et al., 1997).

How do truth conditions provide the meaning of a sentence? 'Snow is white' is true if and only if snow is white; so, 'snow is white' *means* that snow is white. There are two well-known problems with this. First, 'snow is white' is also true if and only if grass is green (at least, this would be so when snow is white if and only if grass is green), but 'snow is white' doesn't *mean* that grass is green. (Although, when it snowed on the first day of Spring the year that I wrote this, I cheered myself up by thinking so!)

Second, although 'All mimsy were the borogoves' is true if and only if all mimsy were the borogoves, to say that 'All mimsy were the borogoves' *means* that all mimsy were the borogoves clarifies little (Harman, 1974: 6; this is the circular dictionary problem, with a circle of radius 0—cf. Rapaport, 1995, §3.2.4). What's missing is knowledge of what 'mimsy' and 'borogove' mean. How could

we find out? We could find the denotations, but that's solipsistically impossible. Alternatively, we could find our mental representations (of the denotations) (cf. Harman, 1974: 6), or we could give a definition of the terms: both of these are purely internal and syntactic, however. Or we could define one in terms of the other (as suggested above in §3 and in Rapaport, 1995, §3.2.2; cf. Rapaport and Ehrlich, 2000): Borogoves are things that can be mimsy, or else being mimsy is something that borogoves can be. Again, this tells us little by itself (more context is needed). In any case, it is still purely syntactic.

Consider both the white-snow and the mimsy-borogoves cases from Cassie's point of view. She hears 'snow is white', and she understands it by mapping 'snow' onto her concept of snow, 'white' onto her concept of white, and forming the proposition that snow is white. That is, she understands the sentence by constructing that proposition, which is now linked to her semantic network. She *believes* that snow is white if and only if either she already had a mental representation of that proposition ("Oh yes; I already knew that") or she has reason to trust the speaker ("Oh yes? Well, if you say so"). If she hears "all mimsy were the borogoves", she will seek to understand by finding (or building) a mimsy-node and a borogove-node, and finding (or building) the proposition that the borogoves were entirely mimsy. But she won't understand it as well as she understands the proposition that snow is white, since it will *not* be linked to the rest of her network. (Or it will be linked to her representation of the rest of *Jabberwocky*. So, at best, she'll have a skeletal understanding in the context of the poem. Or it may be linked to her representations of the rest of *Through the Looking Glass*, in which Humpty Dumpty explains the sentence. In that case, she'll understand it, because further links will have been made. The more links, the more understanding.)

It may be objected that this is an example from literature, so talk of truth conditions is beside the point. But, as Harman points out, that's *part* of the point: "Speakers violate no linguistic conventions when they . . . tell stories" (Harman, 1974: 10; but cf. Galbraith, 1995: 33ff; Segal, 1995: 12ff). So it is not the case that we must claim that speakers try to say what's true. Rather, at most we only have to claim that they try to say what they *believe*. But they don't even always try to do *that*: Sentences from fiction are, depending on your tastes, either false, truth-valueless, or the sort of thing for which a truth theory would be a category mistake (cf. Ryle, 1949; Parsons, 1975; Searle, 1979; Pavel, 1986; Castañeda, 1979, 1989a; Rapaport, 1991a; Rapaport and Shapiro, 1995). In any case, a truth theory yields strange results when applied to sentences from fiction (though no stranger, perhaps, than when applied to modal sentences that require possible—if not fictional—worlds).

The point is that semantics as correspondence between language and *the world* is of no help in giving a first-person explanation of how a cognitive agent understands language. (And it is certainly of no help in giving a first-person explanation of how a cognitive agent understands *fictional* language.) However, semantics as correspondence between language and the agent's mental representations (or lan-

guage of thought) *can* help: “The meaning of a sentence is determined by the thought with which the sentence is conventionally correlated, that is, the thought which, by convention, speakers would normally intend to communicate to a hearer by using that sentence” (Harman, 1974: 10). Of course, to talk of ‘the’ meaning of a sentence is misleading; context needs to be taken into account. But the broader point holds: Meanings of sentences are provided by thoughts, not by truth conditions.

5.2. HARMAN’S SEMANTIC THEORY

Harman, however, formulates this a bit differently from the way that I see it: There are three parts to his formulation. Here are the first two:

- H1. The meanings of linguistic expressions are determined by the contents of the concepts and thoughts they can be used to express. (Harman, 1982: 242, 1987: 55)
- H2. The contents of concepts and thoughts are determined by their functional role in a person’s psychology. (Harman, 1982: 242)

And, in a later essay, Harman analyzes H2 further:

- H2a. The contents of thoughts are determined by their construction out of concepts. (Harman, 1987: 55, 58)
- H2b. The contents of concepts are determined by their functional role in a person’s psychology. (Harman, 1987: 55)

Now, the picture we get from H1 and H2 is that the meaning (M) of a linguistic expression (E) is determined by the content (C) of the thought (T) that is represented by E; and the functional role (F) of thought T determines content C (this Rube-Goldbergian picture is shown in Figure 4).

But this seems to multiply entities. Now, as a Meinongian, I am not normally bothered by such multiplications. However, I fail to see what ‘content’ contributes here, perhaps because I fail to see what it *is*. Nor do I understand what it means for content (whatever it is) to ‘determine’ meaning. In fact, an earlier formulation of Harman’s theory was more streamlined:

The relevant thoughts are to be identified, not in terms of truth conditions, but rather in terms of their potential role in a speaker’s ‘conceptual scheme’ The meaning of a sentence is determined by the role in a conceptual scheme of the thoughts that the sentence would normally be used to express. (Harman, 1974: 10–11)

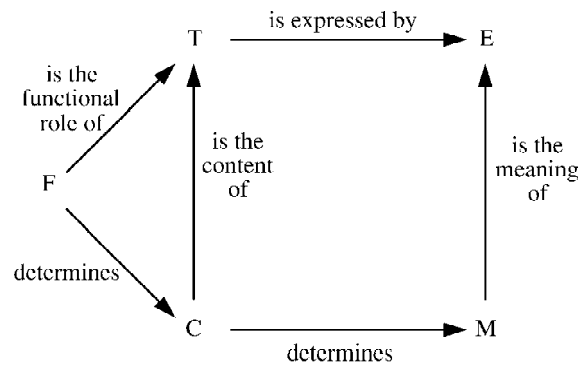


Figure 4. The meaning (M) of a linguistic expression (E) is determined by the content (C) of the thought (T) that is represented by E; the functional role (F) of T determines C.

My view is this:

- R1. The meanings of linguistic expressions *are* the thoughts they express (so ‘meaning’ and ‘expression’ are inverses of each other).
- R2. The content of a thought *is* its functional role.

The SNePS/Cassie picture is this:

- S1. Cassie’s understanding of a linguistic expression is the set of nodes she maps it into (the set of nodes she uses to model the expression).
- S2. Those nodes play a functional role in her entire semantic-network mind.

Presumably, Harman’s ‘concepts’ are SNePS/Cassie’s base nodes¹⁵ (“concepts are treated as symbols in a ‘language of thought’ ” (Harman, 1987: 56)), and Harman’s “thoughts” are SNePS/Cassie’s molecular nodes.

This appears to be consistent with H2a, but H2a is ambiguous: What is it that is constructed out of concepts: Is it *thoughts*? Or is it *contents* of thoughts? On my view, ‘thoughts’ would be constructed out of (or, would be structured by) ‘concepts’ as well as other ‘thoughts’ (for example, Cassie’s thought that Oscar believes that Lucy is rich is constructed out of the thought that Lucy is rich and concepts of Oscar, Lucy, and being rich). And, in contrast to H2b, the ‘meaning’ (in one sense) of thoughts *as well as* of concepts is a function of their location in the entire network of thoughts and concepts.

There is, as I mentioned, a third part to Harman’s theory:

- H3. Functional role is conceived nonsolipsistically as involving relations to things in the world, including things in the past and future. (Harman, 1987: 55; cf. Harman, 1982: 247, 1988)

Now, on the SNePS/Cassie, first-person, internal view, there may indeed be other aspects to the notion of the functional (or conceptual, or inferential) role of a concept or thought. There is, for instance, their role in action (cf. Kumar, 1993a–c, 1994, 1996; Kumar and Shapiro, 1993, 1995), although this role might not be (or contribute) anything over and above the concept's location in the network (and might, in fact, depend entirely upon it). But I part company with Harman on point H3. *Nonsolipsistic* functional role is not something the agent can have access to. Point H3 takes a third-person viewpoint, not a first-person one. I am solely interested in what linguistic expressions mean *to the agent*, not what a third person says that they mean for the agent.

5.3. LANGUAGE IN THOUGHT AND COMMUNICATION

Nevertheless, Sellars's 'entry' and 'departure' rules seem clearly to be links with the external world. They are part and parcel of another issue that Harman raises: the role of language in *thought* as opposed to *communication*. I do not deny that there are "connections between concepts and the external world" (Harman, 1987: 80). I merely deny that such connections tell the *cognitive agent* anything about his or her language or concepts. *Maybe* such connections *do* tell a third person something, but they give no first-person information. (The 'maybe' has to do with the point made in Rapaport, 1995, §§2.5.1, 2.6.2, and in Rapaport, 2000b that, at least, the third person is making connections between his or her *own internal representations* (a) of the other agent's concepts and (b) of his or her own internal model of the world.)

Curiously, the only connections Harman explicitly mentions are those between concepts and *words* and those between concepts and 'normal contexts of functioning' (Harman, 1987: 80). But the link to words is of only *causal* interest. From the SNePS/Cassie point of view, what's important is the *presence* in the internal semantic network of a lex node; *how* it got there is irrelevant. (That's what methodological solipsism is all about; cf. Rapaport, 1996, Ch. 6.) Ditto for normal contexts of functioning: They may give the third person some information, but they avail the first person nothing.

Clearly, it's in the case of 'communication' that these issues come to the fore, not the case of 'thinking'. Harman distinguishes these two uses of language, and finds the latter to be more basic. I agree (to a point), but why then does he care about the external links? Let's look a bit more closely.

The view of language as serving a communicative function *sounds* similar to David Lewis's notion of "language" as

A social phenomenon which is part of the natural history of human beings; a sphere of human action, wherein people utter strings of vocal sounds, or inscribe strings of marks, and wherein people respond by thought or action to the sounds or marks which they observe to have been so produced. (Lewis, 1975: 3.)

But Harman seems to mean something more restrictive, for there can be communication via a syntactic system that is *not* language—for example, Morse code (Harman, 1987: 57).¹⁶

What about the role of language in thought? Harman cites Noam Chomsky (who in turn paraphrases Wilhelm von Humboldt):

[. . .] to have a language is to have a system of concepts [which could be the *meanings* in Lewis's theory of "a language" (Lewis, 1975: 3)]

and it is the place of a concept within this system (which may differ somewhat from speaker to speaker) that, in part, determines the way in which the hearer understands a linguistic expression . . . [T]he concepts so formed are systematically interrelated in an "inner totality", with varying interconnections and structural relations . . . [cf. a semantic network.] This inner totality, formed by the use of language in thought, conception, and expression of feeling, functions as a conceptual world [cf. Dennett's "notional world" (1982)] interposed through the constant activity of the mind between itself and the actual objects, and it is within this system that a word obtains its value

(Harman, 1975: 273; unbracketed ellipses in Harman's text; my interpolations and ellipses in brackets.)

Elsewhere, he calls this use of language 'calculation, as in adding a column of figures' (Harman, 1982: 242, 1987: 56), commenting that conceptual-role semantics 'may be seen as a version of the theory that meaning is use, where the basic use of symbols is taken to be in calculation, not in communication, and where concepts are treated as symbols in a 'language of thought' ' (Harman, 1982: 243). This is clearly a syntactic enterprise.

There is some unclarity, however, when Harman speaks of these two uses of 'language' or of 'symbols' (e.g., Harman, 1987: 56). When he talks of 'symbols', is he talking about external linguistic expressions? Or is he talking about the internal symbols of a language of thought? For SNePS, the nodes are symbols of a language of thought, and they represent propositions, thoughts, and concepts (cf. Shapiro and Rapaport, 1991, Shapiro, 1993; perhaps it would be better to say that they *implement* propositions, thoughts, and concepts). They can be used in 'calculation' (for example, inference) *as well as* in communication (for example, language is generated from them, and they are produced from language). Linguistic expressions are also used in communication. In fact, they are the *vehicles* of communication. *What gets communicated*—what is carried by the vehicle—are thoughts and concepts (that which is represented by the nodes). But linguistic expressions are not normally used in internal calculation (though, of course, they *can* be, as when Cassie wonders what Oscar meant when he *said* 'all mimsy were the borogoves').

My view is that both 'thinking' (or 'calculating') and 'communication' are equally important components. There are spoken and written expressions. And in

Cassie's mind, there are mental concepts in correspondence with them. There are also speakers and hearers, each of whom communicates with others, and each of whom understands the other by means of a semantic interpretation of the other's spoken or written expressions in terms of their own concepts. And, *pace* Harman, thinking *is* communicating with oneself (cf. Harman, 1982: 243): This is Kah-Kyung Cho's point that I mean things by my silent use of language when I talk to myself (Cho, 1992; cf. Rapaport, 1995, §2.5.1), and it works (in part) by the mechanism of 'internal reference' (discussed in Rapaport, 1995, §§2.5.1, 2.6.2, 1996, Ch. 8, §3.1).

Harman and I are, however, not so far apart: "a language, properly so called, is a symbol system that is used both for communication and thought. If one cannot think in a language, one has not yet mastered it" (Harman, 1987: 57). So far, so good. But: "A symbol system used only for communication, like Morse code, is not a language" (Harman, 1987: 57). What, then, about Searle-in-the-room's use of Chinese, for communication only; is that not the use of a language? The answer depends on how much of the story Searle told us. As I have noted elsewhere (Rapaport, 1988, §3.1, 1995, §1, 2000b, §9), he didn't tell us enough. Here's how I see it: Unless the symbols are part of a large network, they have no (or very little) meaning—and, to that extent, maybe Searle has a point. But the more they *are* used for calculation/thinking, the more language-like they are. And, I claim (and I think Harman would agree), they *have* to be part of such a large network, otherwise they could not be used to communicate. They have meaning if and only if, and to the extent that, they're part of a large network. Searle, it seems to me, denies that being part of a large network suffices to provide meaning. What conceptual-role semantics says is that that's the only way to provide it:

... there are two uses of symbols, in communication and speech acts and in calculation and thought. (Nonsolipsistic) conceptual role semantics takes the second use to be the basic one. The ultimate source of meaning or content is the functional role symbols play in thought. (Harman, 1987: 79.)

6. Objections

There have been a large number of objections to conceptual-role semantics. Let's see how powerful they are.

6.1. THE OBJECTION FROM SPEECH-ACT THEORY

Harman raises some potential objections to conceptual-role semantics from speech-act theory (1982: 252–255). But this is not a problem for SNePS/Cassie, since all speech acts have an origination in nodes, hence they do have a conceptual role to play.

Related to this is Harman's discussion of Grice (Harman, 1987: 56–57). There are, at least, three distinct kinds of 'meaning': (1) *natural* meaning (as in: smoke means fire; these are relations between elements entirely within the semantic domain), (2) *non-natural* meaning (as in: 'Feuer' means fire; this seems to be referential meaning, or 'expression meaning'), and (3) non-natural *speaker* meaning ("what a speaker . . . of certain symbols means"; but note that, on my theory—and possibly that of Bruner 1983 (see Rapaport, 1996, Ch. 5, §3)—the speaker could mean one of his or her *concepts or thoughts* rather than something in the world). According to Harman, Grice claims that expression meaning can be analyzed in terms of speaker meaning. This seems reasonable. And, according to Harman, Grice further claims that speaker meaning can be analyzed in terms of the speaker's intentions to communicate. (I have a lot more to say about this in Rapaport, 1996, Ch. 9, §4, where I look at the question of whether non-humans, such as apes and computers, can use language).

But, according to Harman, this last claim

overlook[s] the meaningful use of symbols in calculation. You might invent a special notation in order to work out a certain sort of problem. It would be quite proper to say that by a given symbol you meant so-and-so, even though you have no intentions to use these symbols in any sort of communication. (Harman, 1987: 57)

But you *might* and *could* so use them. So, speaker meaning could, perhaps, be analyzed in terms of the *potential* for communication. Again, *pace* Harman (1987: 56), there seems to be no good reason to deny that "calculation" or thought is internal communication.

Now, Harman has an interesting, but flawed, point to make:

Suppose you use your special notation to work out a specific problem. You formulate the assumptions of the problem in your notation, do some calculating, and end up with a meaningful result in that notation. It would be correct to say of you that, when you write down a particular *assumption* in your notation, you meant such and such by what you wrote: but it would be incorrect to say of you that, when you wrote the *conclusion* you reached in your notation, you *meant* so and so by what you wrote. This seems connected with the fact that, in formulating the *assumption* as you did in your notation, you *intended* to express such and such an assumption; whereas, in writing down the *conclusion* you reached in your notation, your *intention* was *not to express such and such a conclusion* but rather *to reach whatever conclusion in your notation followed from earlier steps by the rules of your calculations*. (p. 57; my italics.)

Harman's point is this: You can't *intend* the *conclusion*, since you haven't reached it yet! Intending to express a thought involves a 'translation' or 'mapping' *from* the thought *to* the notation. After the calculation (which is purely syntactic), you

‘translate’ or ‘map’ *from* the notation *to* the thought; so it can’t have been the case that you *intended* to express that thought. So, you didn’t mean what you wrote when you wrote the conclusion-expressed-in-the-notation.

But that’s quite odd. Consider the old saying that I don’t know what I think until I read what I wrote. We use language to ‘calculate’, to think. Indeed, I *don’t* intend my conclusions *before* I say them—I say them and come to believe them *simultaneously*. But—and this is my point—they mean what they mean in the same way that things I *do* intend to say mean what *they* mean.

Harman continues the previous quotation as follows:

This suggests that you mean so and so in using certain symbols if and only if you use those symbols to express the thought that so and so, *with the intention of expressing such a thought*. (Harman, 1987: 57; my italics.)

But that’s not so. The whole point of symbols and ‘calculation’ is that once I intend a symbol to mean so and so, then that’s what it will always mean (for me), whether or not I intend it at any given time. That’s what enables me to say that the conclusion-expressed-in-the-notation means so and so. It’s what enables me to (inversely) ‘translate’ or ‘map’ from the symbols to meanings (and back again) freely, with or without intentions to communicate.

So: the italicized intention-clause of the right-hand side of the biconditional in the previous quotation has to be modified, perhaps as follows:

Cognitive agent *C* means that so and so in using certain symbols if and only if

1. *C* uses those symbols to express the thought that so and so, and
2. *C* once (or initially) had the intention of expressing such a thought.

(Or perhaps a compositional theory of intending will do the job: Surely, each of the *basic* symbols *in* a thought mean something for me if and only if I use them to express a *concept* with the intention of expressing that concept. Compositionally, a thought-symbol means something for me if and only if I *can* use it to express a thought. Here, no *intentions* to express that thought are needed.)

6.2. THE OBJECTION FROM THE EXISTENCE OF A SHARED EXTERNAL WORLD.

One of the major claims against a conceptual-role semantics is that it ignores the contribution of a truth-functional semantics: the contribution of *reference*, the *fact* that there exists a real world out there that is shared by interlocutors. What *is* the contribution of truth-functional semantics and reference, and what *are* the arguments that (1) they are needed and (2) there exists a shared external world? Let’s look at (2) first.

Clearly, that there is an external world is a fundamental *assumption*. There are, to be sure, G. E. Moore’s arguments for it in ‘Proof of an External World’ (1939), but they amount to little more than a statement of faith or a claim that in fact we

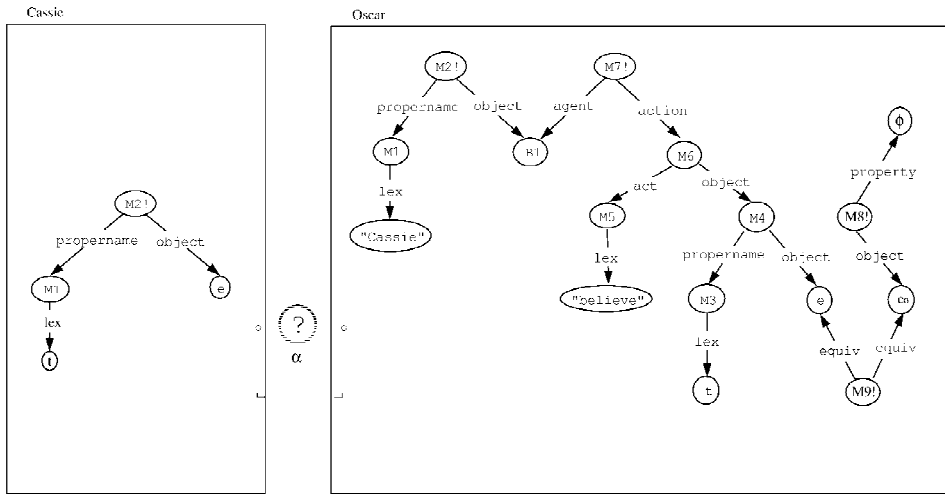


Figure 5. In the external world, α (if it exists) would be the object that Cassie thinks of as e (and refers to by t) as well as the object that Oscar thinks of as e_O . In Cassie's belief space: $M2 =$ (Cassie's belief that) e is named 't'. In Oscar's belief space: $M2 =$ (Oscar's belief that) $B1$ is named 'Cassie', $M7 =$ (Oscar's belief that) Cassie believes that $M4$, $M4 =$ (Cassie's) e is named 't' (by her), $M8 =$ (Oscar's belief that) e_O has property ϕ , $M9 =$ (Oscar's belief that) Cassie's e is the same as (Oscar's) e_O

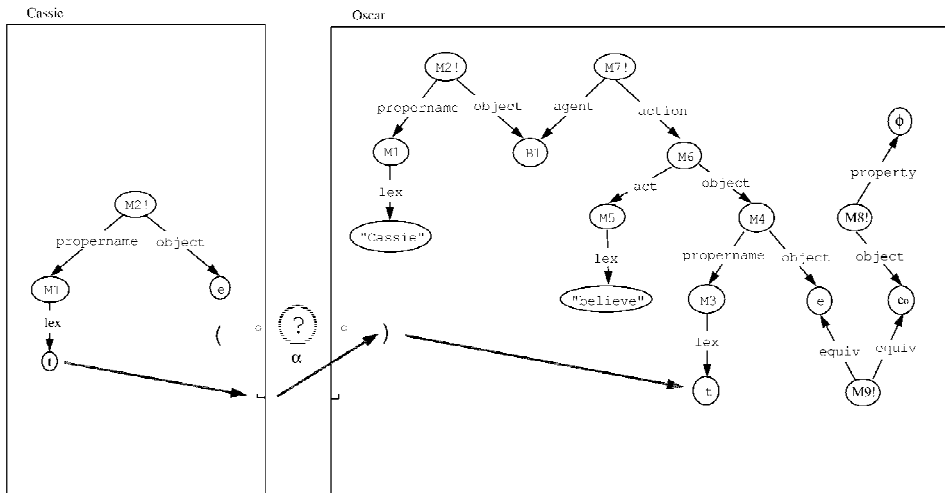


Figure 6. Detail of Figure 5. Cassie utters 't'; Oscar hears 't' and believes that Cassie is thinking of what Oscar thinks of as e_O .

assume that the external world exists or that we behave *as if* it existed. That's consistent with my version of conceptual-role semantics.

What is *reference*, after all? A cognitive agent (for example, Cassie, or I) uses a term t to refer to some entity e in its visual field or in its knowledge base. Consider

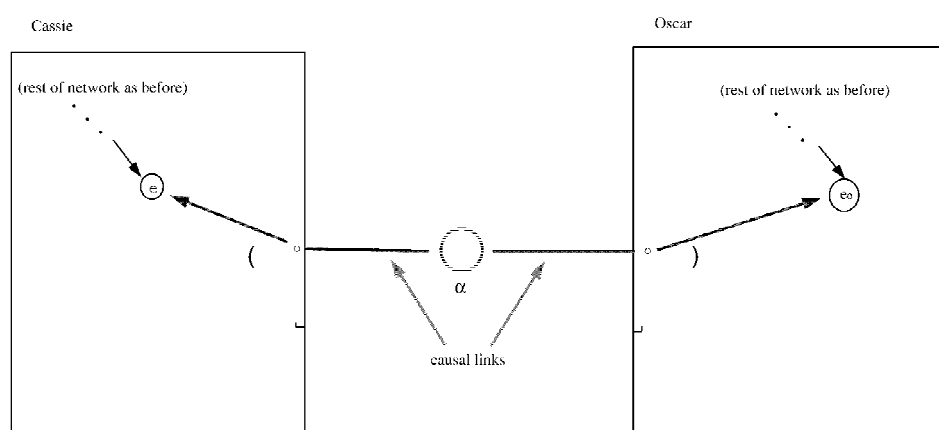


Figure 7. The visual-field case, with causal links.

Figures 5 and 6.¹⁷ The case where t refers to an entity in a knowledge base is purely internal (cf. Rapaport, 1988, §3.4, on deixis). Cassie refers by t to the entity e that she thought of once before. Oscar, hearing Cassie use t , is prompted to think of e_O , which is the object Oscar believes to be equivalent to (or the counterpart of) the one Cassie is thinking of. Whether or not there is an actual object, α , in the external world¹⁸ that corresponds to Cassie's e and Oscar's e_O is irrelevant to explaining the semantics of t . If there is such an α , then there is a correspondence relation between e and α (and an external referential relation between t and α). But that relation is not accessible to *any* mind (except possibly God's, if one wishes to view the external world as (within) God's mind).

In the case where t refers to an entity in one's visual field, t still internally refers to an internal representation, e , this time causally produced (perhaps) by some actual object α . If α exists, then when Oscar hears Cassie use t , Oscar, with luck, will take Cassie to be talking about e_O , which is equivalent to (or a counterpart of) (Oscar's representation of) Cassie's e , as in Figures 7 and 8. Here, that (or whether) α exists is irrelevant to the *semantics* of t , and is not accessible by any (human) mind. If Cassie's and Oscar's communicative negotiations are constrained by the "behavior" of e and e_O (see Rapaport, 1996, Ch. 5), then they might hypothesize the external existence of a noumenal object α , but each of them can only deal with their phenomenal e and e_O , respectively.

Taken together, the knowledge-base and visual-field cases explain why and how a third person can "assign [Cassie's] predicates satisfaction conditions" (Loar, 1982: 274–275). It also takes care of any *argument* that truth and reference are *needed*. Truth and reference, we assume, are there, but inaccessible. Hence, they *couldn't* be *needed*. The contribution of truth and reference is by way of an *attempt* (doomed to failure) to describe what the world is like: They are *metaphysical* notions. Recall that Cassie's claim that Oscar *knows* that φ is really just her claims that Oscar *believes* that φ and that she, too, believes that φ .

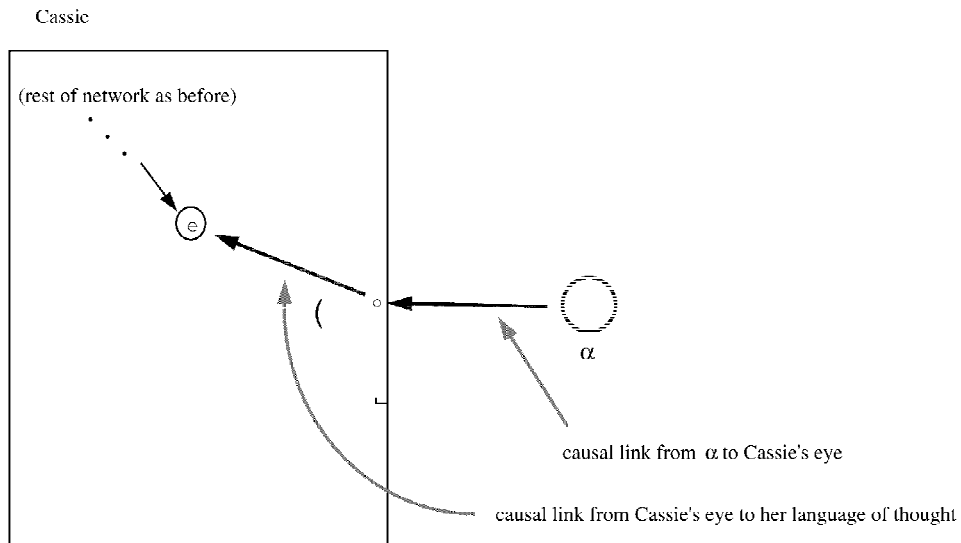


Figure 8. Detail of Figure 7. The causal link from α to e is here analyzed into two links, one from α to Cassie's eye, and another from her eye to her language of thought.

(Similarly, where α_i are 'real-world' objects and $R_{@}$ is a 'real-world' relation, her claim that $\lceil R(x_1, \dots, x_n) \rceil$ is true, in the sense that $(\exists \alpha_1, \dots, \alpha_n, R_{@})[R_{@}(\alpha_1, \dots, \alpha_n)]$, is just her *belief* that $(\exists \alpha_1, \dots, \alpha_n, R_{@})[R_{@}(\alpha_1, \dots, \alpha_n)]$, as in Figure 9. That is, Cassie will have two "mental models": One is her mental model of the actual world; the other is her set of concepts about those things. Perhaps, as is my wont, I am multiplying entities. If so, that just strengthens my internalist perspective: for either R and x would have to go, or $R_{@}$ and α would have to go; what's left is still internal).

6.3. DAVID LEWIS'S OBJECTIONS

David Lewis's 'General Semantics' (1972) is often cited in objections to conceptual-role semantics, or, more specifically, to theories of "semantic interpretation as the assignment to sentences and their constituents of compounds of 'semantic markers' or the like" (p. 169):

Semantic markers are *symbols*: items in the vocabulary of an artificial language we may call *Semantic Markerese*. Semantic interpretation by means of them amounts merely to a translation algorithm from the object language to the auxiliary language Markerese. But we can know the Markerese translation of an English sentence without knowing the first thing about the meaning of the English sentence: namely, the conditions under which it would be true. (p. 169.)

But such a translation algorithm is *all* that Cassie (or any of us) *can* do. For Lewis, however, semantics consists of truth conditions. But how can Cassie come

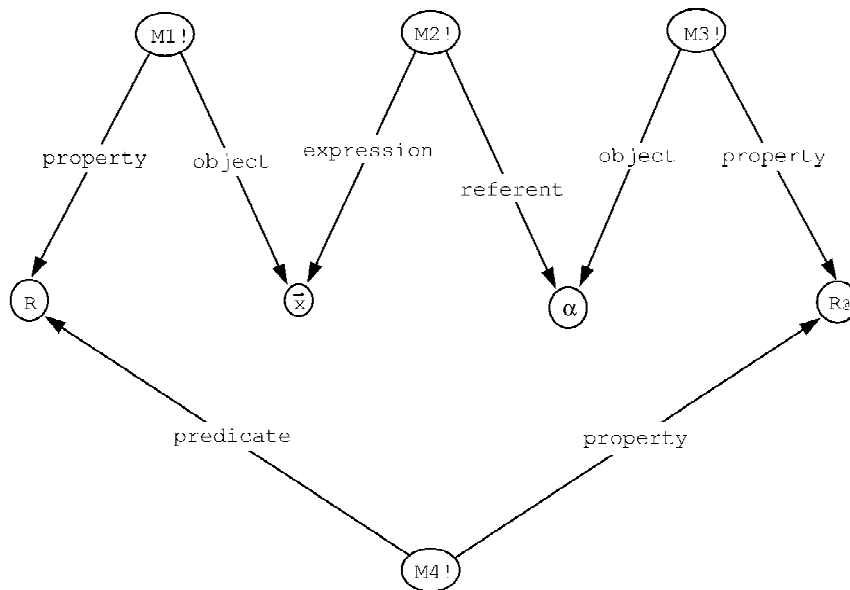


Figure 9. Cassie’s beliefs about properties and predicates. M1 = (Cassie’s belief that) $R\bar{x}$, M3 = (Cassie’s belief that) $R@(\bar{\alpha})$, M2 = (Cassie’s belief that) Expression (i.e., terms) \bar{x} internally refers to $\bar{\alpha}$, M4 = (Cassie’s belief that) Predicate R internally corresponds to property $R@$.

to know those without direct access to the external world? Perhaps she doesn’t need such access. After all, she doesn’t need to know the truth *value* of a sentence, only its truth *conditions*. But that, as we’ve seen, can be handled completely internally. How would Lewis distinguish *that* from Markerese?

Using Markerese is purely syntactic (pp. 169–170). So, ultimately, says Lewis, we need “to do real semantics at least for the one language Markerese” (p. 169). But how? Perhaps via names plus compositionality? If so, then except for the one-time causal production of an internal name by an external object, all is internal and syntactic. And why would we need “to do real semantics”? Perhaps to ground our internal symbols. But that can be done internally (as I argued in Rapaport, 1995, §§3.2.3–3.2.4).

Lewis makes much ado about the finitude of Markerese, which “prevents Markerese semantics from dealing with the relations between symbols and the world of non-symbols” (p. 170). Of course, as Smith (1985) has reminded us (cf. Rapaport, 1995, §2.5.1), semantics in fact does *not* deal with that relation or with “the world of non-symbols”. Lewis’s point is that “meanings may turn out to be . . . infinite entities” (p. 170); our minds, however, are finite (cf. Smith’s (1985) notion of “partiality”, discussed in Rapaport, 1995, §2.5.1). The infinite entities that Lewis takes meanings to be are (roughly) intensions in the Montagovian sense: functions from indices to extensions (cf. p. 176). Presumably, since these take infinite possible worlds among the indices, they are infinite, hence could not be Markerese.

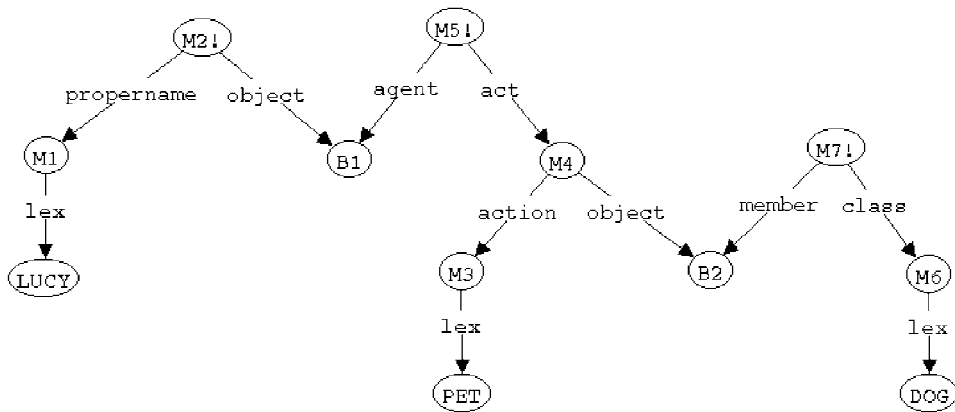


Figure 10. Cassie's belief that Lucy pets a dog: M2 = B1 is named 'Lucy'; M7 = B2 is a dog; M5 = B1 pets B2.

But Markerese symbols could be finite specifications (indeed, algorithms) of such functions, for example, a propositional node (for example, M2 in Figure 10) plus its surrounding network, together with an augmented-transition-network parsing-generating algorithm, which "tells" Cassie how—or provides for her a method—to determine the truth *conditions* of 'Lucy pets a dog'.

'Truth' conditions are, however, a misnomer. Better to call them '*belief*' conditions: Cassie should believe 'Lucy pets a dog' if and only if she believes that B1 represents an entity named 'Lucy', and she believes (*de re*) that B2 represents a member of the class of dogs, and she believes (*de re*) that B1 performs the action of petting B2. (Her believings must be *de re*, since she need not have any beliefs about class membership, and she need not have any beliefs about acts, actions, or their objects as such.)

6.4. TIMOTHY POTTS'S OBJECTIONS

Timothy Potts's essay 'Model Theory and Linguistics' (1973) is instructive, because he agrees with much of what I have had to say yet still locates meaning in the world.

He begins by observing that in model theory, one 'translates' one formal system to another "whose properties are already known [T]he systems thus related to the one under investigation are termed 'models' of it and known properties of the models can then be extrapolated to the new system" (p. 241). This is a clear statement of semantic understanding by general correspondence with an antecedently understood domain; anything, presumably, can be a model of anything else. The problem, as Potts sees it, is that model theory cannot provide a theory of *meaning* for natural language considered as a formal system. His argument is that (1) a theory of *meaning* requires a relation between a language *and the world*, not

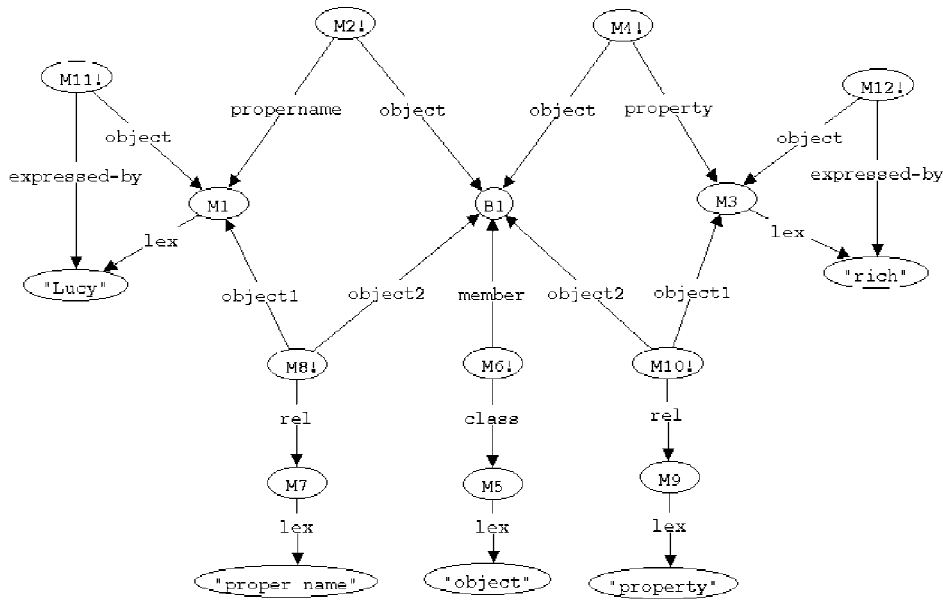


Figure 11. Cognitive-scientist Cassie’s theory of her language of thought. M2 and M4 together constitute her *belief* that Lucy is rich. M6, M8, M10, M11, and M12 together constitute her *theory* about the structure of that belief, where, e.g., M6 is a node (a belief) corresponding to (i.e., that ‘modifies’) the object-arc from M2 to B1. M2 = (Cassie’s belief that) B1 is named ‘Lucy’, M4 = (Cassie’s belief that) B1 is rich, M6 = (Cassie’s belief that) B1 is a (member of the class of) objects, M8 = (Cassie’s belief that) B1 is related by the propername relation to M1, M10 = (Cassie’s belief that) B1 is related by the property relation to M3, M12 = (Cassie’s belief that) M3 is lexically expressed by ‘rich’, M11 = (Cassie’s belief that) M1 is lexically expressed by ‘Lucy’.

between two *languages*, and (2) model theory only gives a relation between two languages. Consistent with my support for conceptual-role semantics, I can accept (2), though I will reject (1). More interestingly, we will see that Potts’s argument for (1) self-destructs! (Another argument Potts has is that natural language isn’t a formal system in the first place. But it *is* a syntactic system, and that’s all that’s needed for the cases I am concerned with.)

First, some preliminary remarks to remind you of the theory I have been adumbrating. Aren’t language-translation manuals theories of meaning of one language in terms of another? As I argued in Rapaport (1995, §3.2.4), a speaker of English would be satisfied if told that the French word ‘*chat*’ means “cat”, while a speaker of French might be satisfied (though I have my doubts!) if told—as one French dictionary has it—that it means “*petit animal domestique, dont il existe aussi plusieurs espèces sauvages*”.¹⁹ But ‘cat’ itself needs to be grounded in a demonstrative definition of the form “that animal over there”. But then we simply have a correspondence continuum (Smith, 1987): ‘*chat*’ means (or is “grounded” in) ‘cat’, which in turn is grounded in the expression “that animal over there”,

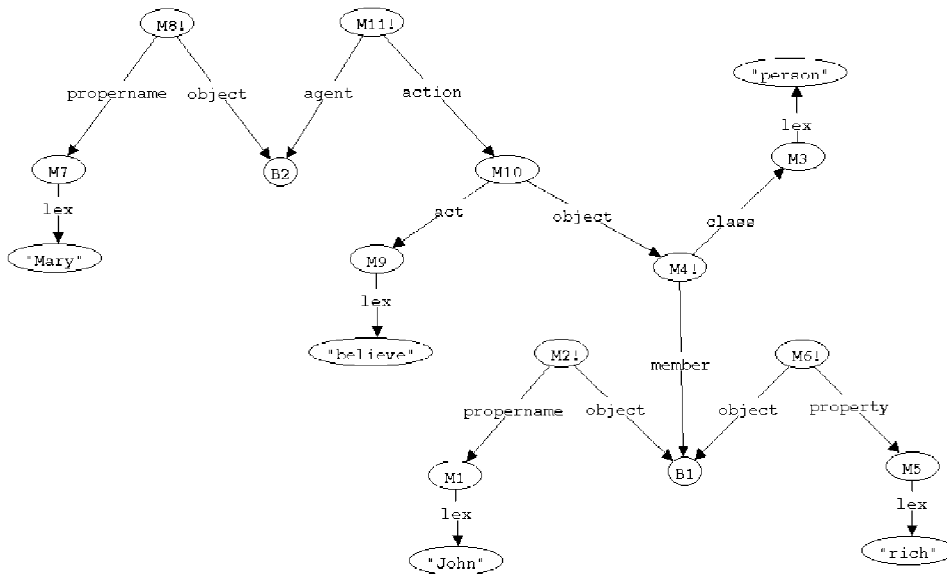


Figure 12. Mary believes (*de re*) of John that he is a rich person. M2 = (Cassie's belief that) B1 is named 'John', M6 = (Cassie's belief that) B1 is rich, M4 = (Cassie's belief that) B1 is a person, M8 = (Cassie's belief that) B2 is named 'Mary', M11 = (Cassie's belief that) B2 believes that M4.

which, finally, is grounded in that animal over there. To learn “the meaning” of ‘*chat*’, one only has to stop at the first antecedently understood domain. And, in any case, the expression “that animal over there” is at best an internal concept. The only “hooks onto the world” (Potts, 1973: 241) are really hooks onto other *internal* nodes. So the expression “that animal over there” is *really* a pointer—not to the world—but to an internal (non-linguistic) representation of the world, as I argued in Rapaport (1995, §§2.5.1 and 2.6.2; cf. Rapaport, 1996, Ch. 8, §3.1, and Perlis, 1991, 1994).

Potts has some useful things to say about models. He first distinguishes

between *being* a structure and *having* a structure: Something *is* a structure if it has distinguishable parts or *elements* which are *inter-related* in a determinate way. . . . [T]wo different things, each of which **is** a structure, can in certain circumstances be said to **have** the *same* structure (p. 244; Potts's italics, my boldface.)

‘Structure’ seems intended as a neutral term; it is, in my terminology, a syntactic notion, since it refers to a system with “elements” that are “interrelated”. To clarify this distinction, Potts discusses the example of a three-dimensional, cardboard model of a house and a two-dimensional blueprint as a model of a (possibly the *same*) house:

Both the drawings and the cardboard model would then qualify as models of the building, each of them having a structure which is also a structure of the

building. But now suppose that we have only the drawings and the cardboard model: the building has not yet been constructed. How can we say that they are models of a building, when there is no building of which they are models? and how can we say that they are models of the *same* building? . . .

These considerations show that the expression *is a model of* is, in logician's parlance, 'intensional'. Accordingly, we cannot say that what makes something which is a structure a model is that there is something else which is also a structure and that both have a structure in common. (p. 245.)

That is, 'is a model of' is intensional in the sense that its second argument need not exist in the external world (cf., e.g., Rapaport, 1985/1986). More to the point, however, is the fact that 'is a model of' is asymmetric. In any case, the common *structure* can be taken as an intentional object (as I argued in Rapaport, 1978), and both the cardboard structure and the blueprint can be taken as models (actually, "implementations") of *it*. Nor does it *follow* from the intensionality of 'is a model of' that the cardboard structure is not a model of the blueprint. Clearly, it *can* be one, as long as the appropriate mappings (correspondences) exist (or can be defined).

Potts provides an argument concerning a gap between the language used to describe a model and the model itself:

In [mathematical] model theory, the structures which are correlated with formal systems are **abstract** structures and thus inaccessible to perception. This is supposed to make no essential difference (p. 247.)

The situation with abstract structures, according to Potts, is that the abstract structure that is the model of the formal system is *not* directly correlated with it. Rather, the only way to access the abstract structure is via an antecedently understood meta-language for it, and it is the correlations between that meta-language and the formal system's object language that do the work:

the abstract structure is a mere beetle in a box. . . . We are not really studying the relations between a formal language and an abstract structure, but between two languages. Model theory is, rather, an exercise in *translation*. We have given meanings to the formulae of our object-language by specifying how they are to be translated into propositions of an **established** language with which it is assumed that we are **already familiar**; to this extent it is true that model theory is concerned with meaning. (p. 248; Potts's italics, my boldface.)

So, Potts has now argued for (2): model theory only gives a relation between two languages. I agree. He still needs to argue for (1): that even though such inter-linguistic translation "is concerned with meaning" to some 'extent', a *real* theory of meaning requires a relation between language and the world, that is, that meaning is *reference*, not sense or conceptual role.

As I see it, of course, it's *primarily* sense or conceptual role. Why do I see it thus? For *de dicto*/intensional reasons: I'm concerned with the *beliefs of a cognitive agent*, not with whether those beliefs are true. Reference enters in two ways.

(a) I explicate sense as a kind of reference to a domain of intensional entities (cf. Rapaport, 1995, §2.6.1). (b) Symbol grounding also requires a kind of reference, but this is a relation between internal nodes, only some of which are perceptually caused (Rapaport, 1995, §3.2.4).

Potts's argument for his claim that model theory doesn't do the job undercuts his claim about (1):

Thus it is just a confusion to suppose that model theory can say anything about the relation of language to the world; it can, at best, only elucidate one language by reference to another. This is all that is needed for its proper, mathematical application, for if the metalanguage is itself a formal language whose properties have already been studied, then the possibility of specifying a translation from the object to the metalanguage allows us to conclude that the object-language has corresponding properties. Talking of a structure in this connection is then quite harmless, though redundant. . . . so the question whether . . . expressions [of the meta-language] have a meaning by denoting [elements of the abstract structure] . . . need not concern us. (pp. 248–249.)

This is astounding! For it can be taken to argue for our purely internal, methodologically solipsistic view by making three substitutions: (i) 'real world' for 'abstract structure' (after all, the real world is supposed to provide the semantic grounding for our language, just as a model is), (ii) 'Cassie's language' for 'meta-language', and (iii) 'Oscar's language' for 'object language'. That is, think of two cognitive agents, Cassie and Oscar, trying to talk about the shared external world by communicating with each other:

[We] can, at best, only elucidate [someone else's] language by reference to [our own]. This is all that is needed for [understanding], for if [Cassie's language] is itself a formal language whose properties have already been studied[—that is, is antecedently understood, *syntactically*—]then the possibility of specifying a translation for [Oscar's language] to [Cassie's language] allows us to conclude that [Oscar understands things as Cassie does]. Talking of [the real world] in this connection is then quite harmless, though redundant. So the question whether [Cassie's language has] a meaning by denoting [things in the real world] need not concern us.

Syntax plus successful communication suffices for semantics. (I explore this theme in Rapaport, 1996, Ch. 5, §3.)

6.5. BARRY LOEWER'S OBJECTIONS

Barry Loewer's essay, "The Role of 'Conceptual Role Semantics'" (1982, cf. Lepore and Loewer, 1981), offers a Davidsonian argument that truth-conditional semantics "will provide the core of an account of the understanding of language used in communication" (p. 307). Here is my reconstruction of his argument. Consider the following reasoning to justify a conclusion that it's snowing:

1. Arabella, a German-speaker, looks out the window and utters “*Es schneit*”.
2. (a) ‘*Es schneit*’ is an indicative sentence.
(b) Arabella is generally reliable.
(c) ∴ Arabella’s utterance of ‘*Es schneit*’ is true.
3. ‘*Es schneit*’ is true if and only if it’s snowing.
4. ∴ It’s snowing.

Now, (4) is supposed to be the conclusion that Arabella’s German-speaking listener, Barbarella, comes to. Here, truth conditions (3) appear to play an essential role in the inference to (4), that is, in Barbarella’s understanding what Arabella said. In contrast, Arabella’s non-German-speaking listener, Esa, does not conclude (4), presumably because he does not know the truth conditions. But let’s consider Barbarella’s and Esa’s cases separately.

Case 1: Barbarella

What is it that Barbarella comes to believe after (1)? Answer: that it is snowing, that is, a belief that she, too, would express as ‘*Es schneit*’. She believes the *proposition*, not the utterance (cf. Shapiro 1993); at least, let’s suppose so, though in this case it doesn’t matter.

But she doesn’t have to arrive at that belief by believing (3). Take her first-person point of view: She hears ‘*Es schneit*’; she processes it as an indicative sentence, and she constructs a mental representation of the proposition it expresses. She believes that proposition because of (2b), *simpliciter*. Thus, neither (2c) nor (3) are needed!

Moreover, (2c) presumably follows from (2a) and (2b) by some rule such as this:

- (i) Indicative sentences uttered by generally reliable people are true.

But (i) is defeasible: Generally reliable people can be mistaken. For instance, Arabella might, without realizing it, be looking at a movie set with fake snow; or Barbarella might not realize that Arabella is acting in the movie and merely uttering her lines! However,

- (ii) Indicative sentences uttered by generally reliable people are believable (or: ought, *ceteris paribus*, to be believed).

seems more reasonable and all that is needed for Barbarella to come to believe that it is snowing. So (3) is not needed at all. And neither, then, is truth-conditional semantics needed to account for the communicative use of language (or, at least, Barbarella’s communicative use).

Case 2: Esa

Loewer ignores Esa, except to say that all Esa comes to believe is that *what Arabella said* (whatever it meant) is probably true. On my view, Esa comes to believe not that but, rather, that he ought to believe what Arabella said (even though he doesn't know what that is). Once again, truth conditions are not needed.

But suppose that Esa, although not a native speaker of German (like Arabella and Barbarella), is *learning* German and can translate 'Es', 'schneit', and N+V sentences into, say, English. Then Esa can reason more or less as follows:

1. Arabella uttered 'Es schneit' (as before).
2. (a) 'Es schneit' is an indicative sentence.
(b) Arabella is generally reliable.
(c) ∴ Arabella's utterance ought to be believed (*ceteris paribus*).
3. 'Es schneit' means (i.e., translates as) "It's snowing".
4. ∴ I ought to believe (*ceteris paribus*) that it's snowing.

Step (3) should be understood, not as saying that the meaning of the German expression 'Es schneit' is the English expression 'It's snowing', but as saying that 'Es schneit' means the same thing as 'It's snowing' (or: plays the same role in German that 'It's snowing' plays in English), where 'It's snowing' *means* (say) M1—where, finally, M1 is a mental representation in Esa's language of thought. Again, there is no need for truth conditions.

Another possibility is that Esa speaks no German at all, but also looks out the window and (somehow) infers or makes an educated guess that 'Es schneit' expresses the weather. Since Esa sees that it's snowing, he infers or makes an educated guess that 'Es schneit' means that it's snowing.²⁰ Again, there is no role for truth conditions to play *in accounting for communicative understanding*. More precisely, there is no role for *external* truth conditions (which is the sort that Davidson, Loewer, et al., are talking about). Arguably, Esa's internal representation of the fact that it's snowing plays the same role *internally* that external truth conditions would play in the Davidsonian/Loewerian story. But this is akin to internal reference. It is all internal, and all syntactic.

Let me conclude my discussion of Loewer with one more lengthy quotation with which I *almost* agree:

The question of how one understands the language one thinks in does seem to be a peculiar one. . . . CRS [conceptual-role semantics] clarifies the situation. It is plausible that understanding a certain concept involves being able to use that concept appropriately. For example, to understand the concept red is, in part, to be able to discriminate red things. According to CRS an expression in *P*'s Mentalese has the content of the concept red just in case it plays the appropriate role in *P*'s psychology, including his [sic] discriminating red things. It follows that if some expression of *P*'s Mentalese is the concept red then *P* automatically understands it. The answer may appear to be a bit trivial—*P* understands the expression of his Mentalese since if he didn't it

wouldn't be his Mentalese—but it is the correct answer. If there are any doubts compare the questions we have been considering with “In virtue of what does a computer ‘understand’ the language it computes in?” Of course the understanding involved in understanding Mentalese is different from the understanding one has of a public language. I argued that understanding the latter involves knowing truth conditions. Not only would knowledge of truth conditions contribute nothing to explaining how we understand Mentalese but, it is clear, we do not know the truth conditions of Mentalese sentences. (Or, for that matter, even the syntax of Mentalese.) If *P* were to encounter a sentence of Mentalese written on the wall (in contrast to its being in just the right place in his brain), he wouldn't have the vaguest idea of what it means because he does not know its truth conditions. (p. 310)

There is much to agree with here—except, of course, that understanding a public language, as *I* have argued, does *not* “involve knowing truth conditions” (except in the sense, which Loewer would not accept, that Esa, above, might have “internal truth conditions”). *P*'s “automatic” understanding of expressions of his Mentalese is just what I have elsewhere called “getting used to” (Rapaport, 1995, §2.1), that is, syntactic understanding.

What about Loewer's last claim, that “If *P* were to encounter a sentence of Mentalese written on the wall . . . he wouldn't have the vaguest idea of what it means because he does not know its truth conditions”? Consider Cassie. She, too, has no knowledge of her language of thought, no knowledge of nodes, arcs, or arc labels. Only if she were a cognitive scientist and had a *theory* of her understanding would she be able to go beyond mere syntax. Even so, it would all be internal: Her theory that her belief that, say, Lucy is rich had a certain structure of, say, nodes and labeled arcs would be expressed in her language of thought. She might, for example, believe (correctly) that her belief that Lucy is rich consisted of two propositions: that someone was named ‘Lucy’ and that that someone was rich. In turn, she might believe (correctly) that the first of these had the structure that an object had a proper name that was lexically expressed by ‘Lucy’ and that the second had the structure that that object had a property lexically expressed by ‘rich’. But her belief that this was so would involve her having *nodes* corresponding to the *arcs* of her actual belief, as in Figure 11. This is all internal, and all syntactic. Now, could she have a theory of the theory of her language of thought? That is, could she talk about the labeled arcs used in that theory? Only by means of “nodifying” them. But there will always be more arc labels about which she cannot talk (and of which, in good Wittgensteinian fashion, she must be silent). Cassie herself can have no semantic understanding of her own language of thought; she can only have a syntactic understanding of it, i.e., she is doomed to only *use* it.

There are further complications. Cassie's theory of her language of thought need not be a theory *about* arcs and nodes. It might (heaven forbid!) be a connectionist theory. Even if her theory *were* about arcs and nodes, and even if her theory of representation matched her actual representations (as opposed, say, to a different node-and-arc representation, e.g., using the alternative SNePS theory of Richard Wyatt, 1989, 1990, 1993, or perhaps a KL-ONE theory), still she would not be able to

supply ‘truth’ conditions, since she would not be able to *mention* (but only *use*) her own representations. Only a third person—a computational neuroscientist—could determine whether her theory were true—that is, could determine whether the representations of her theory corresponded to her actual representations. (And then, of course, this could only be done internal to the computational neuroscientist’s own mind—but I won’t press that point here.)

6.6. WILLIAM G. LYCAN’S OBJECTIONS

William G. Lycan defends the need for truth conditions in his *Logical Form in Natural Language* (1984), arguing that truth plays a role in the translation from utterance to Mentalese:

If a machine or a human hearer understands by translating, how does the translation proceed? Presumably a recursion is required And what property is the translation required to preserve? *Truth together with its syntactic determination* is the obvious candidate. Thus, even if one understands in virtue of translating, one translates in virtue of constructing a recursive truth theory for the target language. (p. 238.)

Now, the translation may *in fact* preserve truth. I don’t deny that there is such a thing as truth (or external reference), only that it’s not needed to account for how we understand language. But the translation algorithm (the semantical procedure of procedural semantics) makes no more explicit *appeal* to truth (to truth *values*) than do rules of inference in logic. Truth can be used to *externally justify* or *certify* the algorithm (or the rule of inference), but the translation (or the inference) goes through anyway, in a purely syntactic fashion.

Negotiation, however, does play a role in adjusting the translation. In fact, the translation might *not* preserve truth. But the process of language understanding is self-correcting.

. . . the assignment of full-fledged truth-conditions to sentences of a natural language helps to explain why a populations’ having that language confers a selectional advantage over otherwise comparable populations that have none (this point is due to Dowty . . .) (p. 240.)

I take this to be part of ‘negotiation’—only here it’s negotiation with the *world*. Is it possible that the claim that truth-conditional semantics plays a role in our understanding of natural language just *is* (1) to accept the existence of (a) others with whom we communicate and (b) the world and (2) the need for negotiation? Sellars and Harman don’t think so:²¹ They allow for language-entry/exit rules. If (1) and (2) *do* amount to the need for truth-conditional semantics, then I suppose we’re just differing on, excuse the expression, semantics, and I probably am taking an intermediary position à la Loewer et al. Still, from the first-person point of view, given that there *is* external input, the rest of the story is all internal. (I explore the issue of negotiation in Rapaport, 1996, Ch. 5.)

6.7. JERRY FODOR AND ERNEST LEPORE'S OBJECTIONS

In 'Why Meaning (Probably) Isn't Conceptual Role' (1991),²² Jerry Fodor and Ernest Lepore argue, not that conceptual-role semantics is *wrong*, but that it is inconsistent with two other principles that normally accompany it: compositionality and the analytic-synthetic distinction (p. 332). Now, personally, I like all three. So am I doomed to inconsistency? I'd like to think not. Let's see.

Fodor and Lepore begin with an assumption (which suggests that the inconsistent triad of conceptual-role semantics, compositionality, and the analytic-synthetic distinction may, rather, be an inconsistent *tetrad*) "that the fact that a word . . . means what it does can't be a brute fact. . . . [S]emantic properties must *supervene on* nonsemantic properties" (p. 329; for more on supervenience, see Rapaport, 1999). This doesn't mean "that semantic properties . . . [are not] irreducibly intentional, or irreducibly epistemological, or irreducibly teleological" (p. 329). It does mean that "It can't be a brute fact . . . that 'dog' means *dog* and not *proton* and that 'proton' means *proton* and not *dog*" (p. 329).

Now, first, why can't that be a brute fact? It's certainly an *arbitrary* fact; for example, 'dog' doesn't *resemble* dogs. So 'dog' *could* have meant 'proton' or even 'cat'. Why *does* 'dog' mean "dog"? The story is, no doubt, buried in pre-etymological history, but one can guess that at some time, someone said 'dog' (or some etymologically-related ancestor) when in the presence of a dog. Isn't that a brute fact? And, if so, it certainly seems to be a *semantic* fact in just about every sense of that term, including that of external correspondence. It is, no doubt, also an intentional (or perhaps epistemological or teleological) fact, but perhaps that's just what it is to *be* a semantic fact.

Now, as it happens, just this story is cited by Fodor and Lepore as an example of a *non-semantic* answer (p. 330). It's one of the versions of what they call "Old Testament" semantics, "according to which the meaning of an expression supervenes on *the expression's relation to things in the world*" (p. 329). Now, I certainly am not an Old Testament semanticist. That is, although I recognize that there was, at some time, a causal link between dogs and 'dog', no doubt mediated by an internal mental representation of a dog, nevertheless that's not, for me, the fundamental meaning of, say, *my* use of 'dog'. For one thing, I might never have seen a dog; I've certainly never seen an aardvark, or a proton, or a unicorn, yet the words for dogs, aardvarks, protons, *and* unicorns are equally *and in the same kind of way* meaningful to me.²³ So their meanings must have to do with something other than (perceptual) experiences of them. But even if I *were* an Old Testament semanticist, I'd consider the dog-'dog' relation to be a semantic one, and brute at that. (For another thing, as Fodor and Lepore point out, there are the Fregean 'morning star'–'evening star' cases, where Old Testament semantics would count these as strictly synonymous, though clearly they are not.)

By contrast, there is "New Testament" semantics, that is, conceptual-role semantics, according to which, semantics supervenes on "intralinguistic relations"

(p. 332). With this, of course—modulo the ‘supervenience’ terminology—I agree. But are such relations “non-semantic properties”? Yes and no: They are the syntactic base case of a recursive conception of semantics; they are syntactic semantics (Rapaport, 1995, 2000b).

Let me note here that Fodor and Lepore see New Testament semantics more as *inferential*-role semantics, albeit broadly construed; cf. pp. 330–331. It is interesting to note that the title of their paper uses ‘conceptual’, yet their arguments are really about *inferential*-role semantics. (This was first pointed out to me by Toshio Morita.) We will see the importance of this distinction later (§6.7.2.1.).

6.7.1. *The Alleged Evils of Holism*

Conceptual-role semantics entails holism, which Fodor and Lepore see as a bad thing (p. 331). I, however, rejoice in the entailment. Why is conceptual-role semantics holistic? Because, by a benign slippery slope, if an expression’s meaning is its conceptual (or inferential) role in the language, it must be its *entire* role in the *entire* language, not some arbitrary subpart of either. Why is holism supposed to be bad? Because it follows

that no two people ever share a belief; that there is no such relation as translation; that no two people ever mean the same thing by what they say; that no two time slices of *the same* person ever mean the same thing by what they say; that no one can ever change his [sic] mind; that no statements, or beliefs, can ever be contradicted . . . ; and so forth. (p. 331.)

Perhaps some of these do follow; but why are they *bad*? Or, rather, can we find the silver lining in this dark cloud? Let’s consider these one by one.

1. *No two people ever share a belief*: This does not follow. If Cassie believes that Lucy is rich, and if Oscar also believes that (the same) Lucy is rich (and if their languages of thought express these beliefs in the same way), then they share that belief. (Their languages of thought may differ, of course, but I take it that that’s not the point Fodor and Lepore are making.) The essential core of the belief (the way it is represented or expressed, its *intrinsic features*) is identifiable independently of its place in the network and is common to its “instantiations” in Cassie and Oscar. Some, like Stich (1983) and probably Fodor and Lepore, might not want to call this a “belief”. But, unlike Stich, I am not here speaking of an *isolated* net consisting only of the nodes and arcs representing “Lucy is rich”. The belief that Cassie and Oscar have in common is indeed embedded in a rich framework of other concepts. But there is a non-arbitrarily identifiable core that they share and which is directly (and solely) responsible for their *utterances* of their beliefs.²⁴ Of course, if Cassie, but not Oscar, believes, in addition, that Lucy is tall, or if Oscar, but not Cassie, believes, in addition, that rich people are snobs, then the (inferential) roles of their beliefs will differ, and, so, the meanings of their utterances that “Lucy is rich” will differ. That is, the *relational properties* of the two “instantiations”

differ, so their roles differ. Hence, by conceptual-role semantics, their meanings differ. So, in a sense, no two people can share a “full” belief; to do that, they would have to have the exact same semantic network, which, if not impossible, is highly unlikely. But Cassie and Oscar can share a belief in a more constrained, yet not arbitrarily constrained, sense. (Cf. (3) and (4), below.)

2. *There is no such relation as translation:* If this means something like literal, word-for-word, expression-for-expression, yet idiomatic translation with no loss of even the slightest connotation, then it indeed follows, but is merely sad, not bad. Languages are just too subtle and complex for that. Literary translation is an art, not a science (cf. e.g., Lourie, 1992). True, ‘*Es schneit*’ or ‘*il neige*’ seem to translate pretty well as ‘it’s snowing’. (Or do they? Would ‘it snows’ be better? Arguably not.) But how about ‘*Pierre a un coeur de pierre*’? “Peter has a heart of stone” misses the pun. The trouble is that the networks of associations for any two languages differ so much that the conceptual roles of its expressions must differ, too. So, translation is out; paraphrases or counterparts are the best we can get. But at least we can get those.
3. *No two people ever mean the same thing by what they say:* This also follows but is not bad. *Your* utterance of ‘Lucy is rich’ does *not* mean what mine does, because of the differing conceptual roles each plays in our network of concepts. Yet we do manage to communicate. How so? Bertrand Russell once observed that if we *did* mean exactly the same things by what we said, there would be no *need* to communicate (1918: 195–196). So *lack* of exact synonymy may be a necessary precondition for communication. If you tell me “Lucy is rich”, I understand you by mapping your utterance into my concepts. Since we speak the same language and live in the same culture, we share a lot of the same concepts, so the mapping is usually pretty good, though never perfect. Witness Cassie and Oscar in (1), above: For Cassie, a tall person is rich (but not necessarily a snob); for Oscar, Lucy is a snob (but not necessarily tall). Though we understand *slightly* different things by what we each say, we understand nonetheless.

Suppose, however, that we don’t understand each other. Suppose I think that ‘punt’ means “kick the ball and see where it lands” (or suppose that I have no idea *what* it means other than in the football *metaphor* “we’ll just have to punt”, uttered, usually, in circumstances where we’re going to try to do something and, if it fails, “we’ll just have to punt”, that is, we’ll have to figure out what to do at that time). (Perhaps it is clear to readers who know more of football than I that I *don’t* understand what it means!) Now suppose that I say to you, “if this plan fails, we’ll just have to punt”, but you *do* understand what it means and take me to be telling you that if what we try fails, then *you’ll* have to find a solution. Clearly, we’ve failed to communicate if that’s not what I intended. Equally clearly, a bit more discussion on our parts can clarify the situation, can help each of us readjust our networks: “Oh, what *you*

meant by ‘punt’ is X”; “Oh, what *you* meant by ‘punt’ is Y, and you know better than I do, since I don’t follow football, so, from now on, that’s what *I’ll* mean by ‘punt’, too”. This permits us to understand each other, *even though we don’t ever mean (exactly) the same thing by what we say.*²⁵

4. *No two time slices of the same person ever mean the same thing by what they say:* This is also true, but not bad. In this very sentence that you are now reading, I don’t mean by ‘mean’ what I meant in the previous sentence in which I used that word, since that sentence was expressed by an earlier time slice of me, who didn’t have *this* sentence that you are now reading as part of his network.²⁶ Indeed, the immediately previous sentence *extends* the conceptual-role-semantic meaning of ‘mean’. Nevertheless, there’s enough of an overlap for communication to succeed. Since this is the first-person case, however, which I’m mostly interested in, let’s consider it a bit further.

One way to clarify the problem is to explicate the conceptual role of an expression E as the *set* of “contexts” containing it. For a concrete instance, in the SNePS case, this could be the set $CR(E)$ of all nodes that dominate or are dominated by the node for the concept expressed by E . (That set may well turn out to be the entire network, not necessarily excluding the nodes for the concept and expression themselves.) Now, suppose Cassie hears a new sentence that uses E . Then E ’s conceptual role changes to a *new* set, $CR'(E) = CR(E) \cup S$, where S is the set of all the nodes newly dominated by and dominating the E -node. Since sets are extensional beasts, $CR(E) \neq CR'(E)$. This, I take it, is the problem that Fodor and Lepore see.

I think there are two ways out of it. One I sketched some time ago in “How to Make the World Fit Our Language” (Rapaport, 1981): As the conceptual role of an expression grows, some parts of it will be seen as more central and, indeed, more stable than others. (Cf. Quine’s “web of belief” (1951, §6). Ehrlich (1995), Ehrlich and Rapaport (1997), and Rapaport and Ehrlich (2000) spell this out in a computational theory of vocabulary acquisition.) Such a central, stable, dictionary-like “definition” of an expression will serve to anchor both *interpersonal* communication and *intrapersonal* meditation. After all, we don’t normally bring to bear *everything* we know about a concept when we hear, use, or think about it. (This can also explain how two people can share a belief.)

The other way out involves using the techniques of non-well-founded set theory to provide a stable identification procedure for nodes in ever-changing (or even circular) networks. The details are spelled out in Hill (1994, 1995).

5. *No one can ever change their mind:* This does not follow. As (4) shows, it’s far from the case that no one can change their mind. Rather, everyone *always* changes their mind (literally, in the case of Cassie). But *that’s* not a problem, for the reasons given in (4).
6. *No statements or beliefs can ever be contradicted:* This either does not follow or else is true but not bad. After all, we reason non-monotonically and are

always, as noted in (5), changing our minds (Martins and Shapiro, 1988; Martins and Cravo, 1991). On the other hand, perhaps this objection holds because (as my colleague Shapiro pointed out) after asserting ‘ P ’, the ‘ P ’ in ‘ $\neg P$ ’ isn’t the same ‘ P ’. In this case, my replies to objections 3 and 4, above, would apply.

6.7.2. *Compositionality and the Analytic-Synthetic Distinction*

So, there’s no reason to reject conceptual-role semantics just because it entails the alleged evils of holism. Is there, then, as Fodor and Lepore want to argue, reason to reject it on the grounds of inconsistency with the hypotheses “that natural languages are compositional, and . . . that the *a/s* [analytic-synthetic] distinction is unprincipled” (in the sense “that there aren’t any expressions that are true or false solely in virtue of what they mean”) (p. 332)?

A preliminary remark before we look at Fodor and Lepore’s argument. For me, truth and falsity are irrelevant, of course. So perhaps I have an easy way out: give up the analytic-synthetic distinction on the grounds of irrelevance. But I suspect that there’s a *doxastic* way to view the analytic-synthetic distinction that can avoid the need to deal with truth values yet still be, potentially, inconsistent with conceptual-role semantics and compositionality: Are there expressions that *ought to be believed* solely in virtue of what they mean? I suspect that the class of such expressions would be identical to the class of analytic expressions as Fodor and Lepore would characterize them. Thus, if ‘bachelors are unmarried’ is supposed to be *true* by virtue of the meanings of ‘bachelor’ and ‘unmarried’ (and ‘are’, plus its syntactic structure), then and only then ought it to be *believed* for that reason. (For the record, I think this warhorse of an example is *not* analytic either way you look at it; see §6.7.3). Likewise, if one ought to *believe* ‘red squares are red’ solely in virtue of the meanings of ‘red’ and ‘square’ (and ‘are’, plus its syntactic structure), then and only then is it *true* in virtue of those meanings. (And, for the record, I think this *is* analytic.)²⁷ In what follows, then, I’ll treat the analytic-synthetic distinction doxastically.

6.7.2.1. *Compositionality and conceptual role.* Consider, first, conceptual-role semantics and *compositionality*. Fodor and Lepore take compositionality to be “non-negotiable”, since it is the only hypothesis that entails “productivity, systematicity and isomorphism”, all of which they take as essential features of natural language (pp. 332–334). Compositionality, of course, only holds for non-idiomatic expressions, as Fodor and Lepore note. To say that, however, is to come dangerously close to circularity. For to say that compositionality only holds for non-idiomatic expressions is to say that it only holds for expressions that can be analyzed, that is, expressions whose meaning *is* determined by the meanings of its parts. So, compositionality only holds for expressions for which it holds. Having said this, however, I should also say that it certainly seems to be a reasonable principle, though I can easily imagine that a sustained effort to understand the semantics of

idioms and metaphors (broadly construed after the fashion of Lakoff, 1987) might undermine it. However, it hasn't, yet. (But cf. Pelletier, 1994a–c, Zadrozny, 1994 for arguments *against* compositionality.)

Productivity certainly seems to be a fact about languages, even *non*-natural ones. A non-compositional language would appear to need an infinite set of primitive terms or an infinite set of formation rules to be productive, and natural languages are clearly finite in both these respects, so finite, non-compositional languages would not be productive.

Systematicity, too, seems a general feature of languages and to follow from compositionality: If the meaning of, say, '*aRb*' were *not* a function of the meanings of '*a*', '*R*', '*b*', and of its formation rule, then there would be no reason to expect '*bRa*' to be well formed or meaningful (though it *might* be).

Isomorphism, however, *seems* a bit more suspect (as even Fodor and Lepore admit, p. 333n2). For one thing, Fodor and Lepore express it in a curiously, albeit apparently harmlessly, one-sided way:

- (I) If a sentence *S* expresses the proposition that *P*, then syntactic constituents of *S* express the constituents of *P*. (p. 333)

What about *vice versa*? Well, if a proposition, *P*, *has* constituents, and if each of them is expressed by (sub-sentential) symbols, then—by compositionality—it does appear that a sentence *S* so structured expresses *P*. But does *P* have to have constituents? What if *propositions* were unanalyzable units? Then the *converse* of (I) would be vacuous, I suppose. But that would play havoc with (I), itself: For *S* might *have* constituents, yet they could not, then, express *P*'s constituents, since *P* wouldn't have any. Here's where compositionality comes to the rescue, I suspect.

What is a proposition, anyway, and what does it have to do with compositionality? Well, compositionality as Fodor and Lepore have it says that *the meaning of a sentence* is a function of its syntactic structural description together with *the meanings of its lexical constituents* (p. 332). The link to propositions must be this: The meaning of a sentence is the proposition it expresses. In that case, lexical meanings must be constituents of propositions. So, compositionality entails that propositions are analyzable. I was willing to grant them that anyway, but I thought it was worthwhile to spell things out.

Here's the first problem (p. 334):

1. Meanings are compositional.
2. Inferential roles are not compositional.
3. ∴ Meanings can't be inferential roles.

We've just accepted (1). Must we accept (2)? Here's the first part of Fodor and Lepore's defense of (2): By compositionality, the *meaning* of, say, 'brown cow' is a function of "the meanings of 'brown' and 'cow' together with its syntax" (p. 334). But, by conceptual-role semantics, the *role* of 'brown cow' is a function of the roles of 'brown' and 'cow' and "*what you happen to believe about brown cows*". So, unlike meaning, inferential role is . . . *not* compositional" (p. 334). I take it that they conclude this because they take the role of 'brown cow' to depend on

something *in addition to* the roles of ‘brown’ and ‘cow’. But that doesn’t seem to be the case: Granted, the role of ‘brown cow’ depends on the roles of ‘brown’ and ‘cow’. What are those roles? Well, they *include* all of my beliefs that involve ‘brown’ and ‘cow’, and *that* includes my beliefs about brown cows. So nothing seems to be added. Now, there *is* a problem—the threat of circularity, viz., that, at bottom, the meaning of ‘brown cow’ will depend on the meaning of ‘brown cow’—but that doesn’t seem to be what Fodor and Lepore are complaining about at this point. Putting that aside for the moment, inferential role *does* seem to be compositional, so it *could* be what meaning is.

Earlier, however, we saw that the *meaning* of ‘brown cow’ has to be a constituent of a proposition—call such a constituent a “concept” for now. So we have two options: (1) *identify* propositions and concepts with roles, or (2) assert that there are two *kinds* of meaning: (a) a sentence means₁ a proposition (and a sub-sentential expression means₁ a concept), and (b) a sentence (or sub-sentential expression) means₂ (or *is*) its role. Now, there’s ample historical precedent for bipartite theories of meaning like (2). We might even think of meaning₁ as a kind of referential meaning. Note that we would then have *three* kinds of referential meaning: classical Fregean *Bedeutung*, internal reference (as discussed in Rapaport, 1995, §§2.5.1 and 2.6.2; 1996, Ch. 8, §3.1) and our new propositional/conceptual sort, which is not unlike a Meinongian theory of meaning (cf. Meinong, 1904; Rapaport, 1976, 1978, 1981, 1985/1986, 1991b, and references therein). Meaning₂—role meaning—would be a kind of *Sinn*. One problem with such a theory is that it doesn’t tell us what propositions or concepts *are*. That’s an advantage to option (1), that of identifying propositions/concepts with roles. I won’t take a stand on this here, though I lean towards the first option, on grounds of simplicity.

Fodor and Lepore’s point is that if I believe that brown cows are dangerous but do not believe that being brown or being a cow is dangerous, then the concept of *dangerous* might be part of the role of ‘brown cow’, yet not be part of the roles of either ‘brown’ or ‘cow’. Here is where Fodor and Lepore’s emphasis on *inferential* role rather than *conceptual* role misleads them. For me, being dangerous might be *inferable* from being a brown cow without being inferable from being brown or being a cow, *simpliciter* (that is, it’s a sort of emergent property or merely contingently but universally true of brown cows). However, if being dangerous is part of the *conceptual* role of ‘brown cow’, it’s also—*ipso facto*—part of the conceptual roles of ‘brown’ and ‘cow’. It can’t *help* but be. If *inferential* role, then, is *not* compositional, but *conceptual* role *is*, then so much the worse for *inferential* role. Inferential role, in any event, is subsumed by the broader notion of conceptual role. At most, then, Fodor and Lepore may have successfully shown why meaning (probably) isn’t *inferential* role. Conceptual role, so far, emerges unscathed, despite Fodor and Lepore’s claim that their argument is “robust . . . [and] doesn’t depend on . . . how . . . inferential role” is construed (p. 335).

6.7.2.2. *Compositionality in SNePS*. Let's look at compositionality from the SNePS viewpoint. Recall that *molecular* nodes have structure, in the sense that they "dominate" other nodes; that is, a molecular node has one or more arcs emanating from it. *Base* nodes, on the other hand, are structureless; that is, they do not dominate any nodes, though they are dominated by other nodes. (An isolated base node would be a "bare particular" (Allaire, 1963, 1965; Baker, 1967; Wolterstorff, 1970) or a "peg" on which to hang properties (Landman, 1986); but SNePS forbids them.) Following William A. Woods (1975), we also distinguish between *structural* and *assertional* information about a node. Roughly, a node's structural information consists of the nodes it dominates; its assertional information consists of the propositional nodes that dominate it.

For example, consider the network of Figure 12, representing the proposition that Mary believes (*de re*) of John that he is a rich person (on the nature of *de re* belief representation, see Rapaport et al., 1997). It contains seven base nodes (B1, B2, "John", "rich", "person", "Mary", "believe") and 11 molecular nodes (M1, . . . , M11).²⁸ Consider B1: As a base node, it has no structure, hence no structural information, but we know assertionally several things about it (or, rather, that which it represents): It is named 'John' (M2), it is rich (M6), and it is a person (M4). Consider M4: Structurally, it is (better: it represents) a proposition that B1 is a person (that is, its constituents are B1 and M3, the latter of which is (or represents) a concept whose only structure is that it is lexicalized as 'person'). Assertionally, we know of M4 that it is believed by Mary. (We also know, since it is an "asserted" node (see §3.2, above), that it is believed by Cassie; this, too, is probably part of its assertional information, even though it has nothing to do with node domination.)

Now, what does M4 mean? *Structurally*, its meaning is determined by the meanings of B1 and M3. For now, let's take the meaning of B1 to be a primitive (or perhaps the node B1 itself). The structural meaning of M3 is determined by the meaning of the "person" node, which, again, we'll assume is either primitive or the node itself. So far, so good for compositionality. However, if meaning is conceptual role in the *entire* network, then we must also consider M4's *assertional* meaning, which is that Mary (and possibly that Cassie) believes it. Is assertional meaning compositional? This may be a matter of legislation. Let's suppose, however, that it *is*. Then the assertional meaning of M4 is determined, let's say, by the assertional meaning of M10 (which is the only node that directly dominates M4—ignore Cassie for now), which, in good compositional turn, is determined by the assertional meaning of M11. What's the assertional meaning of M11? As with base nodes, we could say that it is some sort of primitive or else the node itself. We could also say that at this point we must revert to *structural* meaning. That, in turn, suggests that for the *structural* meaning of a *base* node, we could revert to its *assertional* meaning. To make matters more complex, presumably the meaning of, for example, M8 and B2, also play some role in the assertional meaning of M4.

I will leave for another time (and another researcher: Hill, 1994, 1995) the spelling out of the details. But there are two observations to be made: (1) Circular-

ity abounds. (2) Compositionality is *not* necessarily compromised (see Hill, 1994, §§6.5.2, 6.6). I might also note that productivity, systematicity, and isomorphism likewise do not seem to be compromised or rendered inexplicable. (We'll return to circularity, in the next section.)

6.7.2.3. *The analytic-synthetic distinction.* What happened to the analytic-synthetic distinction? The proposal is to save inferential role by limiting it to *analytic* information: Analytic inferential role *is* compositional, so it *can* be identified with meaning. The first thing to notice is that this *removes* “being dangerous” from the meaning of ‘brown cow’ (and *a fortiori* from the meanings of ‘brown’ and ‘cow’). Now, there are advantages and disadvantages to that. One of the disadvantages is that if I do believe that brown cows are dangerous, then that *is* part of the meaning of ‘brown cow’ (and my concept of brown cows is equally part of what ‘dangerous’ means to me). If, for example, the first time I read ‘dangerous’ was in the sentence ‘brown cows are dangerous’, then what ‘dangerous’ *meant*, for me, is: something that brown cows are. Now (as I argued in Rapaport, 1995, §2.6.2), the more occurrences of ‘dangerous’ (or of ‘brown cow’) I encounter, the less likely it will be that ‘brown’, or ‘cow’, or ‘brown cow’ will play a significant role (excuse the expression) in my understanding of ‘dangerous’ (and, *mutatis mutandis*, the less likely it will be that ‘dangerous’ plays a significant role in my understanding of ‘brown cow’). What will be left when such idiosyncratic, contingent aspects of the meaning play smaller and smaller roles (or drop out of my dictionary-like definition of ‘brown cow’ or of ‘dangerous’)? What will be left may well be just the analytic inferential roles: ‘brown cow’ will *mean* “cow that is brown” (although I might still *believe* that brown cows are dangerous, and have a connotation of danger whenever I encounter ‘brown cow’). That’s the *advantage* of analytic inferential role.

Of course, it’s not *enough*. What about the meaning of ‘cow’ *tout court*? We have a few options even within the family of role-type semantics.

Option 1: ‘cow’ means “cow”, where “cow” is a primitive term of Mentalese or of my language of thought (or a SNePS node). Perhaps this is what Fodor has in mind when he makes such claims as that we have innate concepts of, say, carburetors.²⁹ Option 1 is OK as far as it goes, but not very enlightening.

Option 2: ‘cow’ means my entire set of concepts minus “cow”, where “cow” is as in Option 1. That is, the meaning of ‘cow’ is its entire role (or location) in my entire mental network. That’s holism. I think it’s fine, as I argued earlier. But I grant that it seems to be a bit too much. So, when needed, we can choose Option 3:

Option 3: ‘cow’ means that portion of my entire set of concepts (minus “cow”, of course) from which I can infer whatever else I need to know to use and understand ‘cow’—that is, that more or less stable portion of my conceptual net that corresponds to the sort of information given in a dictionary or small encyclopedia. (This would be one implementation of the SCOPE mechanism of Hill, 1994, 1995. Ehrlich, 1995 limits SCOPE by, roughly, the information necessary to categorize the term.)

What about circularity? Accepting—as I do—both compositionality and *conceptual*-role semantics (rather than mere *inferential*-role semantics, analytic or otherwise), we see that compositionality never “bottoms out”. This, I take it, is one of the prices of the holism of conceptual-role semantics. How expensive is it? Well, note first that it rears its head when we inquire into the meanings of base nodes. Perhaps the structural-assertional distinction renders that head less ugly than it might otherwise appear. The other place that circularity appears is when we try to find a natural “stopping place” in the computation of a node’s “full” (that is, both assertional and structural) meaning (cf. Quillian, 1967, 1968). How bad is *that*? Don’t forget: Our network is huge, and includes internal representations of all of the entities that a Harnad-like grounded theory postulates. We *could* say that the meaning of any node can never be given in isolation—to understand one node is to understand the entire network. We *could* say that the meaning of *some* nodes *is* intrinsic or primitive or given in some sense (Perlis, 1991, 1994 seems to say this; cf. my treatment of Lakoff and Johnson in Rapaport, 1996, Ch. 3, §2.2.2.3). Or we *could* say that some smaller portion of the entire network is sufficient (this is the dictionarylike-definition strategy). We could *also* say all of the above, distinguishing different kinds of meaning for different purposes.

Fodor and Lepore aren’t happy with analytic inferential role, however. First, the only way to identify the *analytic* inferences (from all the others) is to see which ones are validated by meanings alone, but the only way to identify meanings is to look at analytic inferences. I have no stake in defending analytic inferential role. I think that the notion of a broader *conceptual* role, limited at times as in Option 3, avoids this problem. As I hinted in §6.7.2., analytic inferences can be identified quite easily: they’re the ones of the form $\forall x[ANx \rightarrow Ax]$ and $\forall x[ANx \rightarrow Nx]$, where *A* is a predicate modifier; e.g., red squares are red, red squares are square. There are, of course, well-known problems with toy guns (which *are* toys, but not (real) guns), alleged murderers (which *are* alleged but not necessarily murderers), and small elephants (which *are* elephants, but only *relatively* small), but even Fodor and Lepore are willing to waive these (p. 334).

Second, they see analytic inferential role as “jeopardizing” “the naturalizability of inferential role semantics” (p. 336), because it can’t be identified with *causal* role, in turn because there is no causal theory of analyticity. I don’t know what a causal theory of analyticity would look like. If it would be a theory explaining why we tend to infer *N* from *AN* (we do, after all, *tend* to think of toy guns as guns, and there is a sense in which small elephants *are* small, at least as far as elephants go), then I see no reason why we would even *want* to identify (analytic inferential) role with *causal* role. The former seems quite abstract and general; the latter seems to be a mere implementation of it, hence less interesting or theoretically important. And why naturalize semantics at all? Put otherwise, isn’t it natural—and ubiquitous—to begin with?

6.7.3. *The Inconsistency*

So the inconsistency that Fodor and Lepore see in the compositionality/role/analytic-synthetic triad is this: if meaning is (inferential) role, then it is not compositional. If meaning is *analytic* inferential role, and if there were a viable analytic-synthetic distinction, then meaning *would be* compositional. Moreover, analytic inferential-role semantics entails the analytic-synthetic distinction. But there is no viable analytic-synthetic distinction.

There appear to be three options: (1) keep compositionality and reject both the analytic-synthetic distinction and both inferential- and analytic-inferential-role semantics, (2) keep non-analytic inferential-role semantics and reject both the analytic-synthetic distinction and compositionality, and (3) reject all of them.³⁰ Of these, Fodor and Lepore ought to opt for (1).

Their first consideration is to resurrect the analytic-synthetic distinction in a limited form, namely, to allow it “only between expressions and their *syntactic constituents*” (p. 338). That’s fine by me (see my discussion of $AN \rightarrow N$ and $AN \rightarrow A$ inferences). The problem with this that Fodor and Lepore see is that it rules out as analytic such statements as that cows are animals (or, presumably, that bachelors are unmarried men). That’s fine by me, too, tradition be damned. Unless ‘bachelor’ is *defined* as ‘unmarried man’, it really *isn’t* analytic that bachelors are unmarried men. A Martian sociologist trying to figure out what’s “natural” about the category of bachelors would not treat the claim that bachelors are unmarried men as analytic (cf. Rapaport, 1981; Lakoff, 1987; see also the discussion of reverse engineering in Weizenbaum, 1976, esp. p. 134). For Fodor and Lepore, that cows are animals must be analytic if what counts is *inferential* role. But, first, that has to be a rather broad definition of inference (for it is a *biological* inference, not a logical one). And, second, it’s just another reason for preferring *conceptual*-role semantics, which doesn’t license any *analytic* or *logical* inferences from cow to animal. As Fodor and Lepore point out, “If Quine’s arguments show anything, they show that there is no way to reconstruct the intuition that ‘brown cow \rightarrow animal’ is definitional and ‘brown cow \rightarrow dangerous’ isn’t” (p. 339). I agree; but there *is* a way to distinguish these from the strictly definitional ‘brown cow \rightarrow brown’, and that’s all we need.

Their second consideration is that the holism of inferential-role semantics entails “that expressions in different languages are semantically incommensurable” (p. 339). Yes; so what? As we saw in §6.7.1(2), that does not prevent us from communicating—successfully—with one another (for other reasons why, see §§6.6 and 7, and Rapaport, 1996, Ch. 5). Ah—but *is* inferential-role semantics thus holistic? Fodor and Lepore think not: They think that the following argument is not a good one (p. 340):

1. The meaning of an expression is determined by *some* of its inferential relations.

2. “There is no principled distinction between those of its inferential relations that constitute the meaning of an expression, and those that don’t” (p. 340).
3. ∴ The meaning of an expression is determined by *all* of its inferential relations.

Premise 1 follows from inferential-role semantics, premise 2 follows from the *lack* of an analytic-synthetic distinction, and the conclusion is holism. They think that this is not a good way to argue for holism, because it is a slippery-slope argument *and* because it depends on denying the analytic-synthetic distinction. The latter is a problem because if you *accept* a principled analytic-synthetic distinction (as I do), you can’t accept (2), and if you *deny* a principled analytic-synthetic distinction, you can’t accept (1), because (1) requires a principled analytic-synthetic distinction. It seems to me that all that this shows is that holism can’t be inferred this way, not that holism is false.

Here’s how I see it: (1) *is* true. In fact, I can give it at least two interpretations on *conceptual-role* semantics, not *inferential-role* semantics:

- (1a) The *structural* meaning of an expression (or node) is determined by the expressions (or nodes) that constitute (or are dominated by) it.
- (1b) The *dictionary-like* meaning of an expression (or node) is determined by *some* of its conceptual relations. (Which ones depend on the contexts in which the cognitive agent has encountered the expression and on which of those are needed to provide a “stable” meaning, as spelled out in Ehrlich, 1995.)

Premise (2) *is* false. There are *lots* of *different* principled distinctions. One is that between *logical* inferences and *non-logical* ones (between ones whose logical form is $AN \rightarrow N$ or $AN \rightarrow A$ and ones whose logical form is $A \rightarrow B$). Another difference is that produced by (1a): the distinction between structural and assertional information. Yet another is that produced by (1b): the distinction between “core” relations and “peripheral” (or “connotational”) ones. (This is also spelled out in Ehrlich, 1995.) Holism, as I see it, is independent of (1) and (2). But it *does* follow from—indeed, it simply *is*—the notion of the *full meaning* of an expression (or node) as given by conceptual-role semantics.

So the “crack in the foundations of” semantics (p. 342) can be patched by using different brands of role semantics, analytic-synthetic distinctions, and maybe compositionality: Buy *conceptual-role* semantics, a logical (or structural) analytic-synthetic distinction, and some version of compositionality—and accept that there are *lots* of aspects to “the” meaning of an expression.

7. How to Compare Roles

One of the leftover problems that Fodor and Lepore saw has to do with the apparent incommensurability of different systems of roles. Perhaps, they suggest pessimistically, one will have to be reconciled to a theory of *similarity* of meaning, rather than of identity of meaning.

There are, I think, cases where roles indeed can’t be cleanly compared. The clearest cases come from language translation. The role of the French preposition

‘à’ is simply not played by any one preposition in English, nor is the role of the English preposition ‘in’ played by any one preposition in French. However, this prevents neither translation nor mutual comprehension. Nor do cases of dissimilar roles among nouns prevent everyday translation or comprehension, though they wreak havoc with literary and poetic translation, not to mention puns and even everyday associations or connotations. So be it. One can always convey the foreign meaning by a suitable, if prosaic and pedantic, gloss (cf. Rapaport, 1981; Jennings, 1985).

There are ways to compare roles “on the fly”, though one has to look at the larger picture—indeed, larger and larger pictures—and one has to settle, sometimes, for only partial agreement. As Nicolas Goodman has put it, “. . . I associate with your words various complexes of memory, behavior, affect, etc., in such a way that I end up with a sentence which can play *more or less* the same role *in my life* as your sentence plays *in your life*” (personal communication; my italics). The important point is that this correspondence (hence, this *semantic* understanding) *can* be set up. As Douglas B. Lenat and Edward A. Feigenbaum (1991) observe about a similar situation, “While this does not guarantee that the genuine meanings of the concepts have been captured, it’s good enough for us” (p. 236). What is “genuine meaning”? Is it an “intended interpretation”? Intended by whom? In the case of Lenat and Feigenbaum’s CYC system—a vast, encyclopedic knowledge base (but one that can be thought of as akin to the mind of a (computational) cognitive agent; cf., however, Smith, 1991)—there is an answer: The genuine meaning of a concept is the one intended by the CYC researchers. But in the case of a human or of a CYC-like system that “changes its mind” and “learns”, *its own* understanding is just syntactic. More importantly for our present concern,

. . . how does one guarantee that one’s neighbor shares the same meanings for terms? The answer is that one doesn’t, at least not formally or exhaustively. Rather, in practice, one defeasibly assumes by default that everyone agrees, but one keeps in reserve the ubiquitous conflict resolution method that says “one may call into question whether they and their neighbor are simply disagreeing over the meaning of some terms”. (Lenat and Feigenbaum, 1991: 236)

That is, communicative negotiation can resolve conflicts, enabling us to understand one another. But that is another story (told in Rapaport, 1996, Ch. 5).

Acknowledgements

I am grateful to my colleagues Stuart C. Shapiro and Jean-Pierre Koenig, as well as to the members of the SNePS Research Group, for comments on earlier drafts of this paper and to National Science Foundation grant #REC-0106338.

Notes

¹Saint Thomas Aquinas, *Summa Theologiae* I-II 32: 8.

²William James, *The Varieties of Religious Experience*, Lecture 1.

³This section is adapted from Rapaport, 1995, §§1.1.3, 1.2.

⁴Kamp, 1984; Kamp and Reyle, 1993.

⁵Brachman and Schmolze, 1985; Woods and Schmolze, 1992.

⁶Schank and Rieger, 1974; Schank and Riesbeck, 1981; Hardt, 1992; Lytinen, 1992.

⁷Sowa, 1984, 1992.

⁸And, on occasion, ‘Oscar’. Cassie is the *Cognitive Agent of the SNePS System—an Intelligent Entity*. Oscar is the *Other SNePS Cognitive Agent Representation*. See Shapiro and Rapaport, 1985; Rapaport et al., 1997.

⁹This question is to be understood as urged in Rapaport, 1988, 1995, §1.1.1 and 2000b, §9.1.

¹⁰Cf.: “In most cases it is not possible to infer the meaning ascribed to a symbol within a given culture from the symbolic form alone. At the very least, we have to see how that form is used, how it is reacted to. We have to see it in the context of other actions and of other speakers” (Renfrew, 1990: 7). Renfrew, however, is talking about *external* links. I would say, instead, that we have to see how ‘that form’ is connected to other symbolic forms.

¹¹But note some potential problems in trying to do this: The network can’t be too *simple*, for then it would be underspecified (cf. Rapaport, 1988: 123–124). It would be a pattern that was too general, that would match too much. But neither can the network be too *complex* (as in the case of CYC): Although a giant pattern-matching procedure as envisaged by Lenat and Feigenbaum (1991) is possible in principle, I don’t see how it could be carried out in practice very easily. Better to let the nodes (some of them, at least) wear their intended interpretations on their sleeves. To switch examples back to SNePS, it is better to let a lex-node labeled ‘rich’ be expressed by the English word ‘rich’ than by something arbitrary. (Even this might not be needed if smaller, more tractable portions of the full knowledge base could be understood in the manner that Lenat and Feigenbaum suggest.) This is what we do when we talk to each other. This is explored in Rapaport, 1996, Ch. 5.

¹²‘The Language of Theories’ (1959/1963), ‘Truth and ‘Correspondence’ (1961/1963), and, especially, ‘Some Reflections on Language Games’ (1955/1963).

¹³On computational theories of intentions to speak and to act, cf. Bruce, 1975; Cohen and Perrault, 1979; Allen and Perrault, 1980; Cohen and Levesque, 1985, 1990; Grosz and Sidner, 1986; and, in the SNePS tradition, Haller, 1993a, b, 1994, 1995; Kumar, 1993a–c, 1994, 1996; Kumar and Shapiro, 1993, 1995.

¹⁴Cf. the description of Figures 1(I) and 1(II) in Rapaport, 1985/1986: 67–71.

¹⁵And perhaps also structured-individual nodes.

¹⁶Although, as my colleague Stuart C. Shapiro pointed out, Morse code is just another way of inscribing language.

¹⁷I owe the style of picture to Perlis, 1994.

¹⁸In Rapaport, 1976, 1978, 1981, 1985/1986, I called α a “Sein-correlate”.

¹⁹*Dictionnaire de Français*, Paris: Larousse, 1989: 187. Translation: A cat is a small domestic animal of which there also exist many wild species. Hardly an adequate definition!

²⁰Shortly after presenting a version of this paper to my research group, my colleague Shapiro looked out the window, pointed, and said (truthfully), ‘It’s snowing!’. Someone talking to us in that context but who didn’t understand English would probably have come to believe that it was snowing and that that’s what Shapiro had said.

²¹Or maybe they do—cf. Harman on wide functionalism and my reply to that, Rapaport, 1996, Ch. 6, §5.2.

²²In addition to the provenance of this paper as given in Fodor and Lepore, 1991: 328fn (i.e., adapted from Fodor and Lepore, 1992 (cf. their Ch. 6) and originally presented at the 1991 Chicago Linguistic Society (and published in its proceedings), it was also read by Fodor at the SUNY Buffalo Center for Cognitive Science Conference on Cognition and Representation (April 1992).

²³And my son's first acquaintance (at the age of 3) with ducks, lambs, cows, pigs, etc., was through pictures, not by seeing the real things. Yet he had (and has) no problem understanding those words or applying them correctly.

²⁴For a suggestion on how to identify this core, see Rapaport, 1988, Appendix 2.

²⁵Note that we have to steer a course between the Russellian Scylla of the non-necessity of communication due to complete understanding and the Charybdis of the impossibility of communication due to complete lack of understanding, as in the *Calvin and Hobbes* cartoon (1 September 1992) in which Calvin observes that since "any word can mean anything", then by "inventing new definitions for common words, . . . we'll be unable to communicate". (Or the *Sibling Revelry* cartoon (Lew Little Enterprises, Universal Press Syndicate, 10 July 1991) in which a girl says, "I never knew what power there is in controlling words. If there's something I don't like, I just change the *word* for it. For example, if something is selfish or stupid, I'll just call it 'neat.' I'm going to do that to *all* the words", to which her brother replies, "What a neat idea.") The Charybdis option takes us back to the problem of translation.

²⁶And if you re-read that sentence, the meanings will be changed by what you subsequently read. As Italo Calvino (1986: 19) has said, "There should therefore be a time in adult life devoted to revisiting the most important books of our youth. Even if the books have remained the same (though they do change, in the light of an altered historical perspective), we have most certainly changed, and our encounter will be an entirely new thing."

²⁷Because I think that 'red squares are red' means that if x is red and square, then x is red *simpliciter*. Clearly, it is not analytic (indeed, it is false) that, say, alleged murderers are murderers. It is interesting to note (a) that small elephants, although quite big, *are* small for elephants and (b) that toy guns are not only toys, but also considered by many to be guns (though not real ones); after all, children often learn the meaning of 'gun' via toy guns. But I digress.

²⁸Hill (1994, 1995) would not consider sensory nodes (at the heads of λ ex arcs) to be base nodes.

²⁹David Cole, personal communication, 30 June 1994.

³⁰Here's why: There are four principles: compositionality, the analytic-synthetic distinction, inferential-role semantics, and analytic-inferential-role semantics. So there are 16 possible combinations. Rejecting the analytic-synthetic distinction eliminates 8 of them (the ones in which the analytic-synthetic distinction is true). The analytic-inferential-role semantics \rightarrow analytic-synthetic distinction relation eliminates another four (the ones in which analytic-inferential-role semantics is true but the analytic-synthetic distinction is false). Of the remaining 4, the inferential-role semantics \rightarrow \neg compositionality relation eliminates the one in which inferential-role semantics and compositionality are true.

References

- Allaire, Edwin B. (1963), 'Bare Particulars', *Philosophical Studies* 14; reprinted in Michael J. Loux (1970), ed., *Universals and Particulars*, Garden City, NY: Anchor Books, pp. 235–244.
- Allaire, Edwin B. (1965), 'Another Look at Bare Particulars', *Philosophical Studies* 16; reprinted in Michael J. Loux (1970), ed., *Universals and Particulars*, Garden City, NY: Anchor Books, pp. 250–257.
- Allen, James F. and Perrault, C. Raymond (1980), 'Analyzing Intentions in Utterance', *Artificial Intelligence* 15, pp. 143–178; reprinted in Barbara J. Grosz, Karen Sparck Jones and Bonnie L. Webber (1986), eds., *Readings in Natural Language Processing*, Los Altos, CA: Morgan Kaufmann, pp. 441–458.
- Apostel, Leo (1961), 'Towards the Formal Study of Models in the Non-Formal Sciences', in Hans Freudenthal, ed., *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*, Dordrecht, Holland: D. Reidel, pp. 1–37.
- Brachman, Ronald J. and Schmolze, James G. (1985), 'An Overview of the KL-ONE Knowledge Representation System', *Cognitive Science* 9, pp. 171–216.

- Baker, Robert (1967), 'Particulars: Bare, Naked, and Nude', *Noûs* 1, pp. 211–212.
- Bruce, Bertram C. (1975), 'Generation as a Social Action', *Theoretical Issues in Natural Language Processing-1*, Morristown, NJ: Association for Computational Linguistics, pp. 64–67; reprinted in Barbara J. Grosz, Karen Sparck Jones and Bonnie L. Webber (1986), eds., *Readings in Natural Language Processing*, Los Altos, CA: Morgan Kaufmann, pp. 419–422.
- Bruner, Jerome (1983), *Child's Talk: Learning to Use Language*, New York: W. W. Norton.
- Carnap, Rudolf (1928), *The Logical Structure of the World*, Rolf A. George (trans.), Berkeley: University of California Press, 1967.
- Carroll, Lewis (1895), 'What the Tortoise Said to Achilles', *Mind* 4, pp. 278–280; reprinted in *Mind* 104, pp. 691–693, and in William Warren Bartley, III (1997), ed., *Lewis Carroll's Symbolic Logic*, New York: Clarkson N. Potter, pp. 431–434. See also commentary, pp. 466–475.
- Castanêda, Hector-Neri (1979), 'Fiction and Reality: Their Fundamental Connections; An Essay on the Ontology of Total Experience', *Poetics* 8, pp. 31–62.
- Castanêda, Hector-Neri (1989), *Thinking, Language, and Experience*, Minneapolis: University of Minnesota Press.
- Cho, Kah-Kyung (1992), 'Re-thinking Intentionality' (in Japanese), in Y. Nitta, ed., *Tasha-no Genshogaku (Phenomenology of the Other)*, Hokuto Publishing Co.
- Cohen, Philip R. and Levesque, Hector J. (1985), 'Speech Acts and Rationality', *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)*, Morristown, NJ: Association for Computational Linguistics, pp. 49–60.
- Cohen, Philip R. and Levesque, Hector J. (1990), 'Rational Interaction as the Basis for Communication', in Philip R. Cohen, Jerry Morgan and Martha E. Pollack, eds., *Intentions in Communication*, Cambridge, MA: MIT Press, pp. 221–256.
- Cohen, Philip R. and Perrault, C. Raymond (1979), 'Elements of a Plan-Based Theory of Speech Acts', *Cognitive Science* 3, pp. 177–212; reprinted in Barbara J. Grosz, Karen Sparck Jones, and Bonnie L. Webber (1986), eds., *Readings in Natural Language Processing*, Los Altos, CA: Morgan Kaufmann, pp. 423–440.
- Dennett, Daniel C. (1982), 'Beyond Belief', in Andrew Woodfield, ed., *Thought and Object: Essays on Intentionality*, Oxford: Clarendon Press, pp. xvi–95.
- Ehrlich, Karen (1995), 'Automatic Vocabulary Expansion through Narrative Context', *Technical Report 95-09*, Buffalo: SUNY Buffalo Department of Computer Science.
- Ehrlich, Karen and Rapaport, William J. (1997), 'A Computational Theory of Vocabulary Expansion', *Proceedings of the 19th Annual Conference of the Cognitive Science Society (Stanford University)*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 205–210; <<http://www.cse.buffalo.edu/~rapaport/Papers/vocab.cogsci.tr.ps>>
- Fodor, Jerry and Lepore, Ernest (1991), 'Why Meaning (Probably) Isn't Conceptual Role', *Mind and Language* 6, pp. 328–343.
- Fodor, Jerry and Lepore, Ernest (1992), *Holism: A Shopper's Guide*, Cambridge, MA: Basil Blackwell.
- Galbraith, Mary (1995), 'Deictic Shift Theory and the Poetics of Involvement in Narrative', in Judith F. Duchan, Gail A. Bruder and Lynne E. Hewitt, eds., *Deixis in Narrative: A Cognitive Science Perspective*, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 19–59.
- Grosz, Barbara J. and Sidner, Candace L. (1986), 'Attention, Intentions, and the Structure of Discourse', *Computational Linguistics* 12, pp. 175–204.
- Haller, Susan M. (1993a), 'Planning for Intentions with Rhetorical Relations', *Proceedings of a Workshop on Intentionality and Structure in Discourse Relations (Ohio State University)*, Morristown, NJ: Association for Computational Linguistics, pp. 23–26.
- Haller, Susan M. (1993b), 'Collaboration in an Interactive Model of Plan Explanation', *Proceedings of a AAAI Fall Symposium on Human-Computer Collaboration: Reconciling Theory, Synthesizing Practice, Technical Report FS93-05*, Menlo Park, CA: AAAI.
- Haller, Susan M. (1994), 'Interactive Generation of Plan Descriptions and Justifications', *Technical Report 94-40*, Buffalo, NY: SUNY Buffalo Department of Computer Science.

- Haller, Susan M. (1995), 'Planning Text for Interactive Plan Explanations', in E. A. Yfantis, ed., *Intelligent Systems: 3rd Golden West International Conference*, Dordrecht, Holland: Kluwer Academic Publishers, pp. 61–67.
- Hardt, Shoshana L. (1992), 'Conceptual Dependency', in Stuart C. Shapiro, ed., *Encyclopedia of Artificial Intelligence, 2nd edition*, New York: John Wiley & Sons, pp. 259–265.
- Harman, Gilbert (1974), 'Meaning and Semantics', in Milton K. Munitz and Peter K. Unger, eds., *Semantics and Philosophy*, New York: New York University Press, pp. 1–16.
- Harman, Gilbert (1975), 'Language, Thought, and Communication', in Keith Gunderson, ed., *Minnesota Studies in the Philosophy of Science; Vol. 7: Language, Mind, and Knowledge*, Minneapolis: University of Minnesota Press, pp. 270–298.
- Harman, Gilbert (1982), 'Conceptual Role Semantics', *Notre Dame Journal of Formal Logic* 23, pp. 242–256.
- Harman, Gilbert (1987), '(Nonsolipsistic) Conceptual Role Semantics', in Ernest Lepore, ed., *New Directions in Semantics*, London: Academic Press, pp. 55–81.
- Harman, Gilbert (1988), 'Wide Functionalism', in Stephen Schiffer and Susan Steele, eds., *Cognition and Representation*, Boulder, CO: Westview Press, pp. 11–18.
- Harnad, Stevan (1990), 'The Symbol Grounding Problem', *Physica D* 42, pp. 335–346.
- Hill, Robin K. (1994), 'Issues of Semantics in a Semantic-Network Representation of Belief', *Technical Report 94-11*, Buffalo: SUNY Buffalo Department of Computer Science.
- Hill, Robin K. (1995), 'Non-Well-Founded Set Theory and the Circular Semantics of Semantic Networks', in E. A. Yfantis, ed., *Intelligent Systems: 3rd Golden West International Conference*, Dordrecht, Holland: Kluwer Academic Publishers, pp. 375–386.
- Jennings, Richard C. (1985), 'Translation, Interpretation and Understanding', paper read at the American Philosophical Association Eastern Division (Washington, DC); abstract, *Proceedings and Addresses of the American Philosophical Association* 59, pp. 345–346.
- Kamp, Hans (1984), 'A Theory of Truth and Semantic Representation', in J. Groenendijk, T. M. V. Janssen and M. Stokhof, eds., *Truth, Interpretation, and Information*, Dordrecht: Foris, pp. 1–41.
- Kamp, Hans and Reyle, Uwe (1993), *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Dordrecht, Holland: Kluwer Academic Publishers.
- Knight, Kevin (1989), 'Unification: A Multidisciplinary Survey', *ACM Computing Surveys* 21, pp. 93–124.
- Kumar, Deepak (1993a), 'A Unified Model of Acting and Inference', in Jay F. Nunamaker, Jr. and Ralph H. Sprague, Jr., eds., *Proceedings of the 26th Hawaii International Conference on System Sciences*, Vol. III: *Decision Support Systems and Knowledge-Based Systems*, Los Alamitos, CA: IEEE Computer Society Press, pp. 483–492.
- Kumar, Deepak (1993b), 'An AI Architecture Based on Message Passing', in James Geller, ed., *Proceedings of the 1993 AAAI Spring Symposium on Innovative Applications of Massively Parallel Architectures*, Menlo Park, CA: AAAI Press, pp. 127–131.
- Kumar, Deepak (1993c), 'Rational Engines for BDI Architectures', in Amy Lansky, ed., *Proceedings of the 1993 AAAI Spring Symposium on Foundations of Automated Planning*, Menlo Park, CA: AAAI Press, pp. 78–82.
- Kumar, Deepak (1994), 'From Beliefs and Goals to Intentions and Actions: An Amalgamated Model of Acting and Inference', *Technical Report 94-04*, Buffalo: SUNY Buffalo Department of Computer Science.
- Kumar, Deepak (1996), 'The SNePS BDI Architecture', *Decision Support Systems* 16, pp. 3–19.
- Kumar, Deepak and Shapiro, Stuart C. (1993), 'Deductive Efficiency, Belief Revision, and Acting', *Journal of Experimental and Theoretical Artificial Intelligence* 5, pp. 167–177.
- Kumar, Deepak and Shapiro, Stuart C. (1995), 'The OK BDI Architecture', in E. A. Yfantis, ed., *Intelligent Systems: 3rd Golden West International Conference*, Dordrecht, Holland: Kluwer Academic Publishers, pp. 307–317.

- Lakoff, George (1987), *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, Chicago: University of Chicago Press.
- Landman, Fred (1986), 'Pegs and Alecs', in Joseph Y. Halpern, ed., *Theoretical Aspects of Reasoning About Knowledge*, Los Altos, CA: Morgan Kaufmann, pp. 45–61.
- Lenat, Douglas B. and Feigenbaum, Edward A. (1991), 'On the Thresholds of Knowledge', *Artificial Intelligence* 47, pp. 185–250.
- Lepore, Ernest and Loewer, Barry (1981), 'Translation Semantics', *Synthese* 48, pp. 121–133.
- Lewis, David (1972), 'General Semantics', in Donald Davidson and Gilbert Harman, eds., *Semantics of Natural Language*, Dordrecht, Holland: D. Reidel, pp. 169–218.
- Lewis, David (1975), 'Languages and Language', in Keith Gunderson, ed., *Minnesota Studies in the Philosophy of Science; Vol. 7: Language, Mind, and Knowledge*, Minneapolis: University of Minnesota Press, pp. 3–35.
- Loar, Brian (1982), 'Conceptual Role and Truth-Conditions', *Notre Dame Journal of Formal Logic* 23, pp. 272–283.
- Loewer, Barry (1982), 'The Role of "Conceptual Role Semantics"', *Notre Dame Journal of Formal Logic* 23, pp. 305–315.
- Lourie, Richard (1992), 'Raskolnikov Says the Darndest Things', *New York Times Book Review* (26 April), p. 24.
- Lycan, William G. (1984) *Logical Form in Natural Language*, Cambridge, MA: MIT Press.
- Lyttinen, Steven L. (1992), 'Conceptual Dependency and Its Descendants', *Computers and Mathematics with Applications* 23, pp. 51–73; reprinted in Fritz Lehmann, ed., *Semantic Networks in Artificial Intelligence* (1992) Oxford: Pergamon Press, pp. 51–73.
- Maida, Anthony S. and Shapiro, Stuart C. (1982), 'Intensional Concepts in Propositional Semantic Networks', *Cognitive Science* 6, pp. 291–330; reprinted in Ronald J. Brachman and Hector J. Levesque (1985), eds., *Readings in Knowledge Representation*, Los Altos, CA: Morgan Kaufmann, pp. 169–189.
- Martins, João and Shapiro, Stuart C. (1988), 'A Model for Belief Revision', *Artificial Intelligence* 35, pp. 25–79.
- Martins, João and Cravo, Maria R. (1991), 'How to Change Your Mind', *Noûs* 25, pp. 537–551.
- Meinong, Alexius (1904), 'Über Gegenstandstheorie', in Rudolf Haller (1971), ed., *Alexius Meinong Gesamtausgabe*, Vol. II, Graz, Austria: Akademische Druck- u. Verlagsanstalt, 1971, pp. 481–535; English translation ('The Theory of Objects') by Isaac Levi et al., in Roderick M. Chisholm, ed., *Realism and the Background of Phenomenology* New York: Free Press, pp. 76–117.
- Moore, G. E. (1939), 'Proof of an External World', reprinted in G. E. Moore (1962), *Philosophical Papers*, New York: Collier Books, pp. 126–148.
- Morris, Charles (1938), *Foundations of the Theory of Signs*, Chicago: University of Chicago Press.
- Parsons, Terence (1975), 'A Meinongian Analysis of Fictional Objects', *Grazer Philosophische Studien* 1, pp. 73–86.
- Pavel, Thomas G. (1986), *Fictional Worlds*, Cambridge, MA: Harvard University Press.
- Pelletier, Jeff (1994a), 'The Principle of Semantic Compositionality', *Topoi* 13, pp. 11–24.
- Pelletier, Jeff (1994b), 'Semantic Compositionality: The Argument from Synonymy', in Roberto Casati, Barry Smith and G. White, eds., *Philosophy and the Cognitive Sciences*, Vienna: Hölder-Pichler-Tempsky, pp. 311–317.
- Pelletier, Jeff (1994c), 'On an Argument against Semantic Compositionality', in Dag Prawitz and D. Westerstahl, eds., *Logic and Philosophy of Science in Uppsala*, Dordrecht, Holland: Kluwer, pp. 599–610.
- Perlis, Donald (1991), 'Putting One's Foot in One's Head—Part I: Why', *Noûs* 25, pp. 435–455.
- Perlis, Donald (1994), 'Putting One's Foot in One's Head—Part II: How', in Eric Dietrich, ed., *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines*, San Diego: Academic Press, pp. 197–224.
- Potts, Timothy C. (1973), 'Model Theory and Linguistics', in Edward L. Keenan (1975), ed., *Formal Semantics of Natural Language*, Cambridge, UK: Cambridge University Press, pp. 241–250.

- Quillian, M. Ross (1967), 'Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities', *Behavioral Science* 12, pp. 410–430; reprinted in Ronald J. Brachman and Hector J. Levesque (1985), eds., *Readings in Knowledge Representation*, Los Altos, CA: Morgan Kaufmann, pp. 97–118.
- Quillian, M. Ross (1968), 'Semantic Memory', in Marvin Minsky, ed., *Semantic Information Processing*, Cambridge, MA: MIT Press, pp. 227–270.
- Quine, Willard Van Orman (1951), 'Two Dogmas of Empiricism', reprinted in *From a Logical Point of View, 2nd Edition, Revised*, Cambridge, MA: Harvard University Press, 1980, pp. 20–46.
- Rapaport, William J. (1976), *Intentionality and the Structure of Existence*, Ph.D. dissertation, Bloomington: Indiana University, Department of Philosophy.
- Rapaport, William J. (1978), 'Meinongian Theories and a Russellian Paradox', *Noûs* 12, pp. 153–180; errata, *Noûs* 13, pp. 125.
- Rapaport, William J. (1981), 'How to Make the World Fit Our Language: An Essay in Meinongian Semantics', *Grazer Philosophische Studien* 14, pp. 1–21.
- Rapaport, William J. (1985/1986), 'Non-Existent Objects and Epistemological Ontology', *Grazer Philosophische Studien* 25/26, pp. 61–95; reprinted in Rudolf Haller (1986), ed., *Non-Existence and Predication*, Amsterdam: Rodopi.
- Rapaport, William J. (1986), 'Searle's Experiments with Thought', *Philosophy of Science* 53, pp. 271–279; preprinted as *Technical Report 216*, Buffalo, NY: SUNY Buffalo, Department of Computer Science, 1984.
- Rapaport, William J. (1988), 'Syntactic Semantics: Foundations of Computational Natural-Language Understanding', in James H. Fetzer, ed., *Aspects of Artificial Intelligence*, Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 81–131; reprinted (with numerous errors) in Eric Dietrich (1994), ed., *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines*, San Diego, CA: Academic Press, pp. 225–273.
- Rapaport, William J. (1991a), 'Predication, Fiction, and Artificial Intelligence', *Topoi* 10, pp. 79–111; pre-printed as *Technical Report 90-11*, Buffalo: SUNY Buffalo, Department of Computer Science, May 1990.
- Rapaport, William J. (1991b), 'Meinong, Alexius I: Meinongian Semantics', in Hans Burkhardt and Barry Smith, eds., *Handbook of Metaphysics and Ontology*, Munich: Philosophia Verlag, pp. 516–519.
- Rapaport, William J. (1995), 'Understanding Understanding: Syntactic Semantics and Computational Cognition', in James E. Tomberlin, ed., *AI, Connectionism, and Philosophical Psychology*, Philosophical Perspectives, Vol. 9, Atascadero, CA: Ridgeview, pp. 49–88; reprinted in Andy Clark and Josefa Toribio (1998), eds., *Language and Meaning in Cognitive Science: Cognitive Issues and Semantic Theory*, Artificial Intelligence and Cognitive Science: Conceptual Issues, Vol. 4, Hamden, CT: Garland.
- Rapaport, William J. (1996), *Understanding Understanding: Semantics, Computation, and Cognition*, *Technical Report 96-26*, Buffalo, NY: SUNY Buffalo, Department of Computer Science; available on line at either:
 <<ftp://ftp.cse.buffalo.edu/pub/tech-reports/96-26.ps.Z>> or
 <<http://www.cse.buffalo.edu/~rapaport/Papers/book.ps>>
- Rapaport, William J. (1999), 'Implementation Is Semantic Interpretation', *The Monist* 82, pp. 109–130; longer version preprinted as *Technical Report 97-15*, Buffalo, NY: SUNY Buffalo, Department of Computer Science, and *Technical Report 97-5*, Buffalo: SUNY Buffalo, Center for Cognitive Science: <<http://www.cse.buffalo.edu/~rapaport/Papers/implementation.pdf>>
- Rapaport, William J. (2000a), 'Cognitive Science', in Anthony Ralston, Edwin D. Reilly and David Hemmendinger, eds., *Encyclopedia of Computer Science, 4th edition*, London: Macmillan Reference; see also William J. Rapaport, 'Cognitive Science', in Anthony Ralston and Edwin D. Reilly, eds., *Encyclopedia of Computer Science, 3rd edition*, New York: Van Nostrand Reinhold, 1993, pp. 185–189.

- Rapaport, William J. (2000b), 'How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition', *Journal of Logic, Language, and Information* 9, pp. 467–490.
- Rapaport, William J. and Ehrlich, Karen (2000), 'A Computational Theory of Vocabulary Acquisition', in Łucja Iwańska and Stuart C. Shapiro, eds., *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language*, Menlo Park, CA/Cambridge, MA: AAAI Press/MIT Press, pp. 347–375.
- Rapaport, William J. and Shapiro, Stuart C. (1995), 'Cognition and Fiction', in Judith F. Duchan, Gail A. Bruder and Lynne E. Hewitt, eds., *Deixis in Narrative: A Cognitive Science Perspective*, Hillsdale, NJ: Lawrence Erlbaum Associates; portions reprinted as 'Cognition and Fiction: An Introduction', in Ashwin Ram and Kenneth Moorman (1999), eds., *Understanding Language Understanding: Computational Models of Reading*, Cambridge, MA: MIT Press, pp. 11–25.
- Rapaport, William J., Shapiro, Stuart C. and Wiebe, Janyce M. (1997), 'Quasi-Indexicals and Knowledge Reports', *Cognitive Science* 21, pp. 63–107; reprinted in Francesco Orilia and William J. Rapaport (1998), eds., *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda*, Dordrecht: Kluwer Academic Publishers, pp. 235–294.
- Récanati, François (1995), 'The Alleged Priority of Literal Interpretation', *Cognitive Science* 19, pp. 207–232.
- Renfrew, Colin (1990), 'Towards a Cognitive Archeology', in "For Participants in the Workshop for an Explicitly Scientific Cognitive Archeology", unpublished ms., Cambridge, UK: University of Cambridge Department of Archeology, 8–13 April 1990.
- Roberts, Lawrence D. and Rapaport, William J. (1988), 'Quantifier Order, Reflexive Pronouns, and Quasi-Indexicals', *Technical Report 88-16*, Buffalo: SUNY Buffalo, Department of Computer Science, August 1988.
- Russell, Bertrand (1918), 'The Philosophy of Logical Atomism', reprinted in Bertrand Russell, *Logic and Knowledge: Essays 1901–1950*, R. C. Marsh (1956), ed., New York: Capricorn, pp. 177–281.
- Ryle, Gilbert (1949), *The Concept of Mind*, New York: Barnes & Noble University Books.
- Schank, Roger C. and Rieger, Charles J. (1974), 'Inference and the Computer Understanding of Natural Language', *Artificial Intelligence* 5, pp. 373–412.
- Schank, Roger C. and Riesbeck, Christopher K., eds. (1981), *Inside Computer Understanding: Five Programs Plus Miniatures*, Hillsdale, NJ: Lawrence Erlbaum.
- Segal, Erwin M. (1995), 'Narrative Comprehension and the Role of Deictic Shift Theory' in Judith F. Duchan, Gail A. Bruder and Lynne E. Hewitt, eds., *Deixis in Narrative: A Cognitive Science Perspective*, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 3–17.
- Sellars, Wilfrid (1963), *Science, Perception and Reality*, London: Routledge & Kegan Paul.
- Searle, John R. (1979), 'The Logical Status of Fictional Discourse', in *Expression and Meaning*, Cambridge, UK: Cambridge University Press, pp. 58–75.
- Searle, John R. (1980), 'Minds, Brains, and Programs,' *Behavioral and Brain Sciences* 3, pp. 417–457.
- Shapiro, Stuart C. (1979), 'The SNePS Semantic Network Processing System', in Nicholas Findler, ed., *Associative Networks: Representation and Use of Knowledge by Computers*, New York: Academic Press, pp. 179–203.
- Shapiro, Stuart C. (1993), 'Belief Spaces as Sets of Propositions', *Journal of Experimental and Theoretical Artificial Intelligence* 5, pp. 225–235.
- Shapiro, Stuart C. and Rapaport, William J. (1987), 'SNePS Considered as a Fully Intensional Propositional Semantic Network', in Nick Cercone & Gordon McCalla, eds., *The Knowledge Frontier: Essays in the Representation of Knowledge*, New York: Springer-Verlag, pp. 262–315; earlier version reprinted as *Technical Report 85-15*, Buffalo: SUNY Buffalo Department of Computer Science, 1985; shorter version appeared in *Proceedings of the 5th National Conference on Artificial Intelligence AAAI-86, Philadelphia*, Los Altos, CA: Morgan Kaufmann, pp. 278–283; revised version of the shorter version appears as "A Fully Intensional Propositional Semantic

- Network”, in Leslie Burkholder, ed., *Philosophy and the Computer*, Boulder, CO: Westview Press, 1992, pp. 75–91.
- Shapiro, Stuart C. and Rapaport, William J. (1991), ‘Models and Minds: Knowledge Representation for Natural-Language Competence’, in Robert Cummins and John Pollock, eds., *Philosophy and AI: Essays at the Interface*, Cambridge, MA: MIT Press, pp. 215–259.
- Shapiro, Stuart C. and Rapaport, William J. (1992), ‘The SNePS Family’, *Computers and Mathematics with Applications* 23, pp. 243–275; reprinted in Fritz Lehmann (1992), ed., *Semantic Networks in Artificial Intelligence*, Oxford: Pergamon Press, pp. 243–275.
- Shapiro, Stuart C. and Rapaport, William J. (1995), ‘An Introduction to a Computational Reader of Narrative’, in Judith F. Duchan, Gail A. Bruder and Lynne E. Hewitt, eds., *Deixis in Narrative: A Cognitive Science Perspective*, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 79–105.
- Shapiro, Stuart C., Rapaport, William J., Cho, Sung-Hye, Choi, Joongmin, Feit, Elissa, Haller, Susan, Kankiewicz, Jason and Kumar, Deepak (1996), ‘A Dictionary of SNePS Case Frames’, <<http://www.cse.buffalo.edu/sneps/Manuals/dictionary.ps>>
- Smith, Brian Cantwell (1985), ‘Limits of Correctness in Computers,’ in Charles Dunlop and Rob Kling (1991), eds., *Computerization and Controversy*, San Diego, CA: Academic Press, pp. 632–646.
- Smith, Brian Cantwell (1987), ‘The Correspondence Continuum’, *Report CSLI-87-71*, Stanford, CA: Center for the Study of Language and Information.
- Smith, Brian Cantwell (1991), ‘The Owl and the Electric Encyclopedia’, *Artificial Intelligence* 47, pp. 251–288.
- Sowa, John F. (1984), *Conceptual Structures: Information Processing in Mind and Machine*, Reading, MA: Addison-Wesley.
- Sowa, John F. (1992), ‘Conceptual Graphs as a Universal Knowledge Representation’, *Computers and Mathematics with Applications* 23, pp. 75–93; reprinted in Fritz Lehmann (1992), ed., *Semantic Networks in Artificial Intelligence*, Oxford: Pergamon Press, pp. 75–93.
- Stich, Stephen (1983), *From Folk Psychology to Cognitive Science: The Case Against Belief*, Cambridge, MA: MIT Press.
- Weizenbaum, Joseph (1976), *Computer Power and Human Reason: From Judgment to Calculation*, New York: W. H. Freeman.
- Wills, Gary (1991), ‘Man of the Year’, *New York Review of Books* (21 November 1991) 12, pp. 14–18.
- Wolterstorff, Nicholas (1970), ‘Bergmann’s Constituent Ontology’, *Noûs* 4, pp. 109–134.
- Woods, William A. (1975), ‘What’s in a Link: Foundations for Semantic Networks’, in Daniel G. Bobrow and Alan M. Collins (1985), eds., *Representation and Understanding: Studies in Cognitive Science*, New York: Academic Press, pp. 35–82; reprinted in Ronald J. Brachman and Hector J. Levesque, eds., *Readings in Knowledge Representation*, Los Altos, CA: Morgan Kaufmann, pp. 217–241.
- Woods, William A. and Schmolze, James G. (1992), ‘The KL-ONE Family’, *Computers and Mathematics with Applications* 23, pp. 133–177; reprinted in Fritz Lehmann (1992), ed., *Semantic Networks in Artificial Intelligence*, Oxford: Pergamon Press, pp. 133–177.
- Wyatt, Richard (1989), ‘The Representation of Opaque Contexts’, *Technical Report 89-13* (1992) Buffalo, NY: SUNY Buffalo, Department of Computer Science.
- Wyatt, Richard (1990), ‘Kinds of Opacity and Their Representations’, in Deepak Kumar, ed., *Current Trends in SNePS—Semantic Network Processing System*, Lecture Notes in Artificial Intelligence, No. 437, Berlin: Springer, pp. 123–144.
- Wyatt, Richard (1993), ‘Reference and Intensions’, *Journal of Experimental and Theoretical Artificial Intelligence* 5, pp. 263–271.
- Zadrozny, Wlodek (1994), ‘From Compositional to Systematic Semantics’, *Linguistics and Philosophy* 17, pp. 329–342.