UNDERSTANDING UNDERSTANDING:
SYNTACTIC SEMANTICS AND COMPUTATIONAL COGNITION

William J. Rapaport

Department of Computer Science and Center for Cognitive Science
State University of New York at Buffalo
Buffalo, NY 14260
rapaport@cs.buffalo.edu

Abstract

John Searle (1993: 68) says: "The Chinese room shows what we knew all along: syntax by itself is not sufficient for semantics. (Does anyone actually deny this point, I mean straight out? Is anyone actually willing to say, straight out, that they think that syntax, in the sense of formal symbols, is really the same as semantic content, in the sense of meanings, thought contents, understanding, etc.?)." *I say:* "Yes". *Stuart C. Shapiro (in conversation, 19 April 1994) says:* "Does that make any sense? Yes: Everything makes sense. The question is: What sense does it make?" This essay explores what sense it makes to say that syntax by itself *is* sufficient for semantics.

(This preprint is SUNY Buffalo Department of Computer Science Technical Report No. 94-28. It is a preliminary version of 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): 49–88. All quotations from, and all citations to, it should be to the published version.)

1 COMPUTATIONAL NATURAL-LANGUAGE UNDERSTANDING AND A COMPUTATIONAL MIND

1.1 Understanding Language.

What does it mean to understand language? "Semantic" understanding is a correspondence between two domains; a cognitive agent understands one of those domains in terms of the other. But if one domain is to be understood in terms of another, how is the other understood? Recursively, in terms of yet another. But, since recursion needs a base case, there must be a domain that is not understood in terms of another. So, it must be understood in terms of itself. How? Syntactically! Put briefly, bluntly, and a bit paradoxically, *semantic understanding is syntactic understanding.* Thus, any cognitive agent—including a computer—capable of syntax (symbol manipulation) is capable of understanding language.

1.1.1 Computers, programs, and processes.

What does it mean for a computer to understand language? Strictly speaking, neither computers nor programs can do so. Certainly, but uninterestingly, no present-day computers or AI programs do so. Some suitably-programmed computers can process a lot of natural language, though none can do it (yet) to the degree needed to pass a Turing Test. Rather, if a suitably programmed computer is ever to pass a Turing Test for natural-language understanding, what will understand will be neither the mere physical computer (the hardware) nor the static, textual program (the software), but the dynamic, behavioral *process*—the program being executed by the computer (cf. Tanenbaum 1976: 12; Smith 1987, §5).

1.1.2 The real thing.

Such a successful natural-language–understanding process will be an example of "strong AI". First, it will probably be "psychologically valid"; i.e., the underlying algorithm will probably be very similar (if not identical) to the one we use. Second, natural-language understanding is at least necessary, and possibly sufficient, for passing the Turing Test. Thus, anything that passes the Turing Test does understand natural language. But such a process will pass the Turing test. So, such a process will do more than merely simulate natural-language understanding; it will *really* understand natural language. Or so I claim. What is needed for any cognitive agent—human or computer—to understand language?

1.1.2.1 Robustness. A cognitive agent that understands language must be "open-ended" or "robust", able to deal with "improvisational audience-participation discourse":

Although some "canned" patterns of conversation will be needed, as in theories of "frames", "scripts" etc. (e.g., Minsky 1975, Schank & Riesbeck 1981), it cannot rely solely on these. For we can use language in arbitrary and unforeseen circumstances. Similarly, the language-understanding process must be able to improvise.

Second, monologues are fine as far as they go; but a language-using entity unable to converse with an interlocutor would not pass the Turing Test. Interaction provides feedback, allowing the two natural-language–understanding systems (the two interlocutors) to reach mutual understanding (to "align" their "knowledge bases"). It also provides causal links with the outside world.

Finally, the process must be able to understand not only isolated sentences, but *sequences* of sentences that form a coherent discourse. What it understands at any point in a discourse will be a function partly of what it understood before. (Cf. Segal et al. 1991: 32.)

1.1.2.2 Natural-Language Competencies. A natural-language–understanding process must understand virtually all input that it "hears" or "reads", whether grammatical or not; after all, *we* do. It must remember what it believed or heard before, as well as what it learns during a conversation. It must be able to perform inference on what it hears and believes; revise its beliefs, as needed; and remember what, that, how, and why it inferred. It must be able to plan and to understand plans: In particular, it must be able to plan speech acts, so that it can *generate* language to answer questions, to ask questions, and to initiate conversation. Thus, it must be both a natural-language–*understanding* process and a natural-language–*generation* process; call this natural-language *competence* (Shapiro & Rapaport 1991). And it must be able to understand the speech-act plans of its interlocutors, in order to understand why speakers say what they do. This, in turn, requires the process to have (or to construct) a "user model"—a theory of the interlocutor's beliefs. Last on this list (though no doubt more is needed), it must be able to learn via language—to learn about non-linguistic things (the external world, others' ideas), and to learn about language, including its own language: It must be able to learn its own language from scratch, as we do from infancy, as well as consciously learn the syntax and semantics of its language, as we do (or should) in school.

1.1.3 Mind.

To do all of this, a cognitive agent who understands natural language must have a "mind"—what AI researchers call a 'knowledge base'. Initially, it will contain what might be called "innate ideas"—anything in the knowledge base before any language use begins. And it will come to contain beliefs resulting from perception, conversation, and inference. Among these will be internal representations of external objects.

For convenience and perspicuousness, let us think of the knowledge base or mind as a propositional semantic network, whose nodes represent individual concepts, properties, relations, and propositions, and whose connecting arcs structure atomic concepts into molecular ones (including structured individuals, propositions, and rules). The specific semantic-network theory we use is the SNePS knowledge representation and reasoning system (see §1.2), but you can think in terms of other such systems, such as (especially) Discourse Representation Theory,[1] the KL-ONE family,[2] Conceptual Dependency,[3] or Conceptual Graphs.[4] (Or, if you prefer, you can think in terms of a connectionist system.)

1.1.4 Syntax suffices.

> Philosophy must be done in the first person, for the first person. (Hector–Neri Castañeda, in conversation, 1984)

Meaning will be, *inter alia*, relations among these internal representations of external objects, on the one hand, and other internal symbols of the language of thought, on the other. A cognitive agent, C, with natural-language competence understands the natural-language output of another such agent, O, "by building and manipulating the symbols of an internal model (an interpretation) of [O's] output considered as a formal system. [C]'s internal model would be a knowledge-representation and reasoning system that manipulates symbols" (Rapaport 1988b: 104). Hence, C's semantic understanding of O is a *syntactic* enterprise.

Two semantic points of view must be distinguished. The *external* point of view is C's understanding of O. The *internal* point of view is C's understanding of itself. There are two ways of viewing the external point of view: the "third-person" way, in which *we*, as external

[1] Kamp 1984, Kamp & Reyle 1993.
[2] Brachman & Schmolze 1985, Woods & Schmolze 1992.
[3] Schank & Rieger 1974, Schank & Riesbeck 1981, Hardt 1992, Lytinen 1992.
[4] Sowa 1984, 1992.

observers, describe C's understanding of O, and the "first-person" way, in which C understands its own understanding of O. Traditional referential semantics is largely irrelevant to the latter, primarily because external objects *can* only be dealt with via internal representations of them. It is first-person and internal understanding that I seek to understand and that, I believe, can only be understood syntactically. I have argued for these claims in Rapaport 1988b. The rest of this essay is an investigation into what kind of sense this makes.

1.2 A Computational Mind.

The specific knowledge-representation and reasoning (KRR) system I will use to help fix our ideas is the SNePS *S*emantic *Ne*twork *P*rocessing *S*ystem (Shapiro 1979; Shapiro & Rapaport 1987, 1992, 1995). *As a knowledge-representation system,* SNePS is symbolic (or "classical"; as opposed to connectionist), propositional (as opposed to being a taxonomic or "inheritance" hierarchy), and 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, 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.

There is no loss of generality in focussing on such a *symbolic* system. A connectionist system that passed the Turing Test would make my points about the syntactic nature of understanding equally well. For a connectionist system is just as computational—as syntactic—as a classical symbolic system (Rapaport 1993).

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 taxomonic have automatic inheritance features; in SNePS, this 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 cognitive agents named 'Cassie' and 'Oscar'.[5] If Cassie passes the Turing Test, then she *is* intelligent and *has* (or perhaps *is*) a mind. (Or so I claim.) Her 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),

[5] Cassie is the *C*ognitive *A*gent of the *S*NePS *S*ystem—an *I*ntelligent *E*ntity. Oscar is the *O*ther *S*NePS *C*ognitive *A*gent *R*epresentation. Shapiro & Rapaport 1985; Rapaport, Shapiro, & Wiebe 1986.

properties, relations, propositions, events, situations, etc. 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, inten*t*ional—hence inten*s*ional—objects. They are not extensional objects in the external world, though, of course, they may bear some 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 & Shapiro 1982; Shapiro & Rapaport 1987, 1991, 1992, 1995; Rapaport 1988b, 1991; Rapaport & Shapiro 1995.)

Does Cassie understand English? (This question is to be understood as urged in §1.1.1.) 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.

2 SEMANTICS AS CORRESPONDENCE.

2.1 The Fundamental Principle of Understanding.

It has been said that you never really understand a complex theory such as quantum mechanics—you just get used to it. This suggests the following *Fundamental Principle of Understanding*:

> To understand something is either
>
> 1. to understand it *in terms of something else*, or else
> 2. to "get used to it".

In type-1 understanding, one understands something *relative* to one's understanding of another thing. This is a *correspondence theory* of understanding (or meaning, or semantics—terms that, for now, I will take as rough synonyms). The correspondence theory of truth is a special case.

Type-2 understanding is *non-relative*. Or, perhaps, it *is* relative—but to itself: To understand something by getting used to it is to understand it in terms of *itself*, perhaps to understand *parts* of it in terms of the *rest* of it. The coherence theory of truth is a special case.

Type-1 understanding is *externally* relative; type-2 understanding is *internally* relative. Type-1 understanding concerns correspondences between two domains; type-2 understanding concerns syntax.

Since type-1 understanding is relative to the understanding of something *else*, one can only understand something in this first sense if one has *antecedent* understanding of the other thing. How does one understand the other thing? Recursively speaking, either by understanding it relative to some third thing, or by understanding it *in itself*—by being used to it. Either this "bottoms out" in some domain that is understood non-relativistically, or there is a large circle of domains each of which is understood relative to the next. In either case, our understanding bottoms out in "syntactic" understanding of that bottom-level domain or of that large domain consisting of the circle of mutually or sequentially understood domains.

'Correspondence' and 'syntactic understanding' are convenient shorthand expressions that need explication. Before doing so, let me make it clear that I use the terms 'syntax' and 'semantics' in Morris's classic sense (Morris 1938: 6–7): *Syntax* concerns the relations that symbols have among themselves and the ways in which they can be manipulated. *Semantics* concerns the relations between symbols, on the one hand, and the things the symbols "mean", on the other. Classically, then, semantics always concerns two domains: a domain of things taken as symbols and governed by rules of syntax, and a domain of other things. Call these two domains, respectively, the 'syntactic domain' and the 'semantic domain'. There must also be a relation between these two domains—the "semantic relation".

Understanding, in the usual and familiar sense of type-1 understanding, is a semantic enterprise in Morris's sense of semantics. But this has some surprising ramifications. Once these are seen, we can turn to the less familiar, type-2 sense of understanding as a syntactic enterprise (§3).

When faced with some new phenomenon or experience, we seek to understand it. Perhaps this need to understand has some evolutionary survival value; perhaps it is uniquely human. Our first strategy in such a case is to find something, no matter how incomplete or inadequate, with which to *compare* the new phenomenon or experience. By thus *interpreting* the "unknown" or "new" in terms of the "known" or "given", we seek analogies that will begin to satisfy, at least for the moment, our craving for understanding. For instance, I found the film, *My Twentieth Century*, to be very confusing (albeit quite entertaining—part of the fun was trying to figure it out, trying to understand it). I found that I could understand it—at least as a working hypothesis—by mapping the carefree character Lili to the pleasure-seeking, hedonistic aspects of 20th-century life; another character—her serious twin sister, Dora—to the revolutionary political activist, social-caring aspects of 20th-century life; and the third main character—a professor—to the rational, scientific aspects of 20th-century life. The film, however, is quite complex, and these mappings—these correspondences or analogies—provided for me at best a weak, inadequate understanding. The point, however, is that I *had to*—I was *driven to*—find *something* in terms of which I could make sense of what I was experiencing.

This need for connections as a basis for understanding, as an anchor in unchartered waters, can also be seen in the epiphenal well-house episode in the life of Helen Keller. With water from the well running over one hand while Annie Sullivan finger-spelled 'w-a-t-e-r' in the other, Helen suddenly understood that 'w-a-t-e-r' meant water (Keller 1905). This image of one hand literally in the semantic domain and the other literally in the syntactic domain is striking. By "co-activating" her knowledge (her understanding) of the semantic domain (viz., her experiences of the world around her) and her knowledge of the syntactic domain (viz., her experiences of finger-spellings), she was able to "integrate" (or "bind") these two experiences and thus understand (cf. Mayes 1991: 111).

Is there an alternative to the classical view of semantics as correspondence? Many philosophers and linguists look with scorn upon formal or model-theoretic semantics. However, as long as one is willing to talk about "pairings" of sentences (or their structural descriptions) with meaning (cf. Higginbotham 1985: 3), there is no alternative. That is, if we are to talk *at all* about "the meaning *of* a sentence", we must talk about *two* things: sentences and meanings. Thus, there must be two domains: the domain of sentences (described syntactically) and the domain of the semantic interpretation.

There is, however, another kind of semantics, which linguists not of the formal persuasion study. Here, one is concerned not with what the meanings of linguistic items are, but with semantic relationships among linguistic items: synonymy, implication, etc.[6] These relationships are usually distinct from, though sometimes dependent upon, syntactic relationships. But note that they are, nonetheless, relationships *among linguistic, i.e., syntactic, items*. Hence, on our terms, they, too, are "syntactic", not "semantic".[7] So, semantics is either correspondence or else syntactic.

2.2 Tarskian Semantics.

2.2.1 Syntactic systems.

On the standard view, the syntactic domain is usually some (formalized) language L, described syntactically. That is, one specifies a stock of symbols and rules for forming well-formed formulas (WFFs) from them. (What I intend by 'symbols' are just marks, (perhaps) physical inscriptions or sounds, that have only some very minimal features such as having distinguished, relatively unchanging shapes capable of being recognized when encountered again.) A *language* is sometimes augmented with a *logic*: Certain WFFs of L are distinguished as axioms (or "primitive theorems"), and rules of inference are provided that specify how to produce "new" theorems from "old" ones. The general pattern should be familiar (see, e.g., Rapaport 1992ab). The point is that all we have so far are symbols and (syntactic) rules for manipulating them either linguistically (to form WFFs) or logically (to form theorems)—syntax in Morris's sense.

2.2.2 Semantic interpretations.

Given a syntactic domain such as L, one can ask purely "internal", syntactic, questions: What are L's WFFs and theorems? One can also ask: What's the meaning of all this? What do L's symbols mean (if anything)? What, e.g., is so special about the WFFs or the theorems? To answer this sort of question, we must go outside the syntactic domain, providing "external" entities that the symbols mean, and showing the mappings—the associations, the correspondences—between the two domains.

Now a curious thing happens: I need to show you the semantic domain. If I'm very lucky, I can just point it out to you—we can look at it together, and I can describe the correspondences ("The symbol A_{37} means that red thing over there."). But, more often, I have to describe the semantic domain to you in ... symbols, and hope that the meaning of *those* symbols will be obvious to you.

As an example, let's see how to provide a semantic interpretation of L. Assuming L has individual terms, function symbols, and predicate symbols—combinable in various (but not arbitrary) ways—I need to provide meanings for each such symbol as well as for their legal combinations. So, we'll need a non-empty domain $\mathbf{D}$ of things that the terms will mean and sets $\mathbf{F}$ and $\mathbf{R}$ of things that L's function and relation symbols will mean, respectively. These three sets can be collectively called $\mathbf{M}$ (for **M**odel). $\mathbf{D}$ contains anything you want to talk or think about. $\mathbf{F}$ and $\mathbf{R}$ contain functions and relations on $\mathbf{D}$ of various arities—i.e., anything you want to be able to say about the things in $\mathbf{D}$. That's our *ontology*, what there is.

Now for the correspondences. To say what a symbol of L means in $\mathbf{M}$, we can define an

[6]Kean Kaufmann and Matthew Dryer helped me see this.

[7]Kaufmann says that *cognitive* linguistics is not to be included here, presumably because it pairs sentences with meanings "in the head" ("cognitive" meanings), in which case, of course, it is a correspondence theory of semantics.

interpretation mapping I that will assign to each symbol of L something in $\mathbf{M}$. Again, the general way of doing this should be familiar (cf. Rapaport 1992ab). Typically, I is a homomorphism; i.e., it satisfies a principle of compositionality: The interpretation of a molecular symbol is determined by the interpretations of its atomic constituents in the usual recursive manner. Ideally, I is an *isomorphism*—a 1–1 and onto homomorphism; i.e., *every* item in $\mathbf{M}$ is the meaning of *just one* symbol of L.[8] (Being onto is tantamount to L's being "complete".) In this ideal situation, $\mathbf{M}$ is a virtual duplicate of L. (Indeed, $\mathbf{M}$ could *be* L itself (cf. Chang & Keisler 1973: 4ff), but that's not very interesting or useful for *understanding* L.) Less ideally, there might be symbols of L that are *not* interpretable in $\mathbf{M}$: I would be a *partial* function. Such is the case when L is English and $\mathbf{M}$ is the world ('unicorn' is English, but unicorns don't exist), though if we "enlarge" or "extend" $\mathbf{M}$ in some way, then we can make I total (e.g., we could take $\mathbf{M}$ to be Meinong's *Aussersein* instead of the actual world; cf. Rapaport 1981). In another less ideal circumstance, "Horatio's Law" might hold: There are more things in $\mathbf{M}$ than in L; i.e., there are elements of $\mathbf{M}$ not expressible in L: I is not onto. Or I might be a relation, not a function, so L would be ambiguous. There is another, more global, sense in which L could be ambiguous: By choosing a different $\mathbf{M}$ (and a different I), we could give the symbols of L entirely distinct meanings. Worse, the two $\mathbf{M}$s need not be isomorphic.

Suppose that L is a language for ordinary propositional logic and that $\mathbf{M}$ is a model for it consisting of states of affairs and Boolean operations on them. As an experiment, one could devise an exotic formal symbol system L' using, say, boxes and other squiggles as symbols, and give it a syntax like—but not *obviously* like—that of L (say, with only postfix notation, to make it more disorienting). Not realizing that it was syntactically isomorphic to L, one could only understand L' by getting used to manipulating its symbols, laboriously creating WFFs and proving theorems: doing grammatical and logical syntax. But one could provide relief by giving a semantic interpretation of L' in terms of a model whose domain is L's atomic *symbols*. Of course, I could also have told you what L''s symbols mean in terms of L's model, $\mathbf{M}$. In that case, L' just *is* ordinary propositional logic, exotically notated. In the first way, the model for L' is itself a syntactic formal symbol system (viz., L) whose meaning can be given in terms of $\mathbf{M}$, but L''s meaning can be given either in terms of L *or* in terms of $\mathbf{M}$.

Obviously, the exotic L' is not a very "natural" symbol system. Usually, when one presents the syntax of a formal symbol system, one already has a semantic interpretation in mind, and one *designs* the syntax to "capture" that semantics: In a sense that will become clearer, the syntax is a model—an implementation—of the semantics.

We also see that it is possible and occasionally even useful to allow *one syntactic* formal symbol system to be the semantic interpretation of *another*. Of course, this is only useful if the interpreting syntactic system is antecedently understood. How? In terms of *another* domain with which we are antecedently familiar! So, in our example, the unfamiliar L' was interpreted in terms of the more familiar L, which, in turn, is interpreted in terms of $\mathbf{M}$. And how is it that we understand what states of affairs in the world are? Well ... we've just gotten used to them.

In out example, L is a sort of "swing" domain, serving as L''s *semantic* domain and as $\mathbf{M}$'s

[8] Perhaps isomorphism is less than ideal, at least for the case of natural languages. When one studies, not isolated or made-up sentences, but "real, contextualised utterances ... it is often the case that all the elements that one would want to propose as belonging to semantic structure have no overt manifestations in syntactic structure. ... [T]he degree of isomorphism between semantic and syntactic structure is mediated by pragmatic and functional concerns ..." (Wilkins 1992: 154).

syntactic domain. We can have a "chain" of domains, each of which except the first is a semantic domain for its predecessor, and each of which except the last is a syntactic domain for its successor. To understand any domain in the chain, we must understand its successor. How do we understand the last one? Syntactically. But I'm getting ahead of myself. Let's first look at some "chains".

2.3 The Correspondence Continuum: Data.

Let's begin with examples of *pairs* of things: One member of each pair plays the role of the syntactic domain; the other plays the role of the semantic domain.

1. The first example is the obvious one: L and $\mathbf{M}$ (or L' and L).

2. The next examples come from what I'll call (after Wartofsky 1979) *The Muddle of the Model in the Middle*. There are two notions of "model" in science and mathematics: We speak of a "mathematical model" of some physical phenomenon, by which we mean a mathematical, usually formal, theory of the phenomenon. In this sense, a model is a *syntactic* domain whose intended semantic interpretation is the physical phenomenon being "modeled". But we also speak of a semantic interpretation of a syntactic domain as a "model", as in the phrase 'model-theoretic semantics'. In this sense, a model is a *semantic* domain. We have the following syntax/semantics pairs:

   data/formal theory (i.e., theory as interpretation of the data),
   formal theory/set-theoretic (or mathematical) model (i.e., a model of the theory),
   set-theoretic (or mathematical) model/real-world phenomenon.

   The latter is closely related to—if not identical with—the data that we began with, giving us a cycle of domains! (Cf. Rosenblueth & Wiener 1945: 316.)

3. A *newspaper photograph* can be thought of as a semantic interpretation of its *caption*. But a cognitive agent reading the caption and looking at the photo makes further correspondences: (a) There will be a mental model of the caption—the reader's semantic interpretation of the caption-as-syntax; (b) there will be a mental model of the photo—the reader's semantic interpretation of the photo-as-syntax; and, (c) there may be a single mental model that collates the information from each of these and which, in turn, is a semantic interpretation of the picture+caption unit. (Srihari & Rapaport 1989, 1990; Srihari 1991ab.)

4. A *musical score*, say, Bach's *Goldberg Variations*, is a piece of syntax; a *performance of* it is a semantic interpretation. And, of course, there could be a performance of the *Goldberg Variations* on piano or on a harpsichord. E.g., a piano transcription of a symphony is a semantic interpretation of the symphony (cf. Pincus 1990; conversely, Smith (1985: 636) considers "musical scores as models of a symphony").

5. Similarly, the *script* of a play is syntax; a *performance* of the play is a semantic interpretation. For a performance to be a semantic interpretation of the script, an actual *person* would play the role (i.e., be the semantic interpretation) of a *character* in the play. (Scripts are like computer programs; performances are like computer processes; see example 17; cf. Rapaport 1988a.)

6. A *movie* or *play* based on a *novel* can be considered a semantic interpretation of the text. In this case, there must be correspondences between the characters, events, etc., in the book and

the play or movie, with some details of the book omitted (for lack of time) and some things in the play or movie added (decisions must be made about the colors of costumes, which might not have been specified in the book, just as one can *write* about a particular elephant without specifying whether it's facing left or right, but one can't *show*, *draw*, or *imagine* the elephant without so specifying).

7. Consider a narrative text as a piece of syntax: a certain sequence of sentences and other expressions in some natural language. The *narrative* tells a *story*—the story is a semantic interpretation of the text. On this way of viewing things, the narrative has a "plot"—descriptions of certain events in the story, but not necessarily ordered in the chronological sequence that the events "actually" occurred in. Thus, one story can be told in many ways, some more interesting or suspenseful than others. The story takes place in a "story world". Characters, places, times, etc., in the story world correspond to linguistic descriptions or expressions of them in the narrative. (Cf. Segal 1995.)

8. The reader of the narrative constructs a mental model of the narrative as he or she reads it. This mental story is a semantic interpretation of the syntactic narrative. Or one could view it as a *theory* constructed from the narrative-as-data (cf. Bruder et al. 1986; Duchan et al. 1995).

9. Examples 4–8 suggest a tree of examples: Some *narrative text* might be interpreted as a *play*, on which an *opera* is based. There could be a *film* of a *ballet* based on the *opera*, and these days one could expect a "*novelization*" of the film. Of course, a (different) ballet could be based directly on the play, or a film could have been based directly on the play, then novelized, then re-filmed. Or a symphony might have been inspired by the play, and then have several performances.

10. The *linguistic and perceptual "input"* to a cognitive agent can be considered as a syntactic domain whose semantic interpretation is provided by the agent's *mental model* of his or her (or its) sensory input. (The mental model is the agent's "theory" of the sensory "data"; cf. examples 2, 8.)

11. The *mental model*, in turn, can be considered as a syntactic language of thought whose semantic interpretation is provided by the *actual world*. In this sense, a person's beliefs are true to the extent that they correspond to the world.

12. In Kamp's Discourse Representation Theory, there is a discourse (i.e., a linguistic text—a piece of syntax), a (sequence of) discourse representation structures, and the actual world (or a representation thereof), with mappings from the discourse to the discourse representation structures, from the discourse to the world, and from the discourse representation structures to the world. Each such mapping is a semantic interpretation. One can also consider the correspondences, if any, between the story world and the actual world; these, too, are semantic. (Cf. examples 7, 8, 10, and 11.)

13. The *Earth* is the semantic domain for a global *map*.

14. A *house* is a semantic interpretation of a *blueprint* (cf. Potts 1973, Rapaport 1978, Smith 1985).

15. A *scale model* (say, of an airplane) corresponds to the *thing modeled* (say, the airplane itself) as syntax to semantic interpretation. And, of course, the thing modeled could itself be a scale

model, say, a statue; so I could have a model of a statue, which is, in turn, a model of a person. (Cf. Smith 1985, Shapiro & Rapaport 1991).

16. A *French translation* of an *English text* can be seen from the French speaker's point of view as a *semantic* interpretation of the English syntax, and from the English speaker's point of view as a *syntactic* expression of the English (cf. Gracia 1990: 533).

17. A computer *program* is a static piece of syntax; a computer *process* can be thought of as its semantic interpretation. And, according to Smith, one of the concerns of knowledge representation is to interpret *processes* in terms of the actual world: "It follows that, in the traditional terminology, the *semantic* domain of traditional programming language analyses [which "take ... semantics as the job of mapping programs onto processes"] should be the knowledge representer's so-called *syntactic* domain" (Smith 1987: 15, 17–18).

18. A *data structure* (such as a stack or a record) provides a semantic interpretation of (or, a way of categorizing) the otherwise inchoate and purely syntactic *bits* in a computer (Tenenbaum & Augenstein 1981, Schneiderman 1993). Suppose we have a computer program intended to model the behavior of customers lining up at a bank. Some of its data structures will represent customers. This gives rise to the following transitive syntax–semantics chain: syntactic bits are semantically interpreted by data structures, which, in turn, are semantically interpreted as customers. (Cf. Smith 1982: 11.)

No doubt you can supply more examples. My conclusion is this:

> Semantics and correspondence are co-extensive. *Whenever* two domains can be put into a correspondence (preferably, but not necessarily, a homomorphism), one of the domains (which can be considered to be the *syntactic domain*) can be understood in terms of the other (which will be the *semantic domain*).

2.4 The Correspondence Continuum: Implications.

The syntactic domain need not be a "language", either natural or formal. It need only be analyzable into parts (or symbols) that can be combined and related—i.e., manipulated—according to rules. (Cf. Wartofsky 1979: 6.)

Moreover, *the so-called "syntactic" and "semantic" domains must be treated on a par*; i.e., one cannot say of a domain that it is syntactic except relative to another domain which is taken to be the semantic one, and vice versa: "[T]he question of whether an element is syntactic or semantic is a function of the point of view; the syntactic domain for one interpretation function can readily be the semantic domain of another (and a semantic domain may of course include its own syntactic domain)" (Smith 1982: 10).

Finally, *what makes something an appropriate* ***semantic*** *domain is that it be antecedently understood*. This is crucial for promoting semantics as "mere" correspondence to the more familiar notion of semantics as meaning or understanding. And ultimately such antecedent understanding is syntactic manipulation of the items in the semantic domain.

Suppose that something identified as the semantic domain is *not* antecedently understood, but that the putative syntactic domain *is*. Then, by switching their roles, one can learn about the former semantic domain by means of its syntactic "interpretation" (cf. Rosenblueth and Wiener 1945: 318, Corless 1992: 203).[9]

2.5 The Correspondence Continuum of Brian Cantwell Smith.

What I have referred to as the "correspondence continuum" has received its most explicit statement and detailed investigation in the writings of Brian Cantwell Smith (from whom I have borrowed the term).

2.5.1 Worlds, models, and representations.

In an important essay on computer ethics, Smith (1985) sets up the Wartofskian "model muddle" as follows:

> When you design ... a computer system, you first formulate a model of the problem you want it to solve, and then construct the computer program in its terms. ...
>
> To build a model is to conceive of the world in a certain delimited way. ... [C]omputers have a special dependence on these models: *you write an explicit description of the model down inside the computer*, in the form of a set of rules or ... *representations*—... linguistic formulae encoding, in the terms of the model, the facts and data thought to be relevant to the system's behaviour. ... [T]hat's really what computers are (and how they differ from other machines): they run by manipulating representations, and representations are always formulated in terms of models. (Smith 1985: 636.)

The model is an abstraction of the real-world situation. For instance, "a hospital blueprint would pay attention to the structure and connection of its beams, but not to the arrangements of proteins in the wood the beams are made of ..." (Smith 1985: 637). The model is itself "modeled" or *described* in the computer program; the model, thus, is a "swing domain", playing the role of syntactic domain to the real world's semantic domain, and the role of semantic domain to the computer program's syntactic—indeed, linguistic—description of it.

Smith calls the process of abstraction (which for him includes "every act of conceptualization, analysis, categorization", in addition to the mere omission of certain details) a necessary "act of violence—[if you] don't ignore some of what's going on—you would become so hypersensitive and so overcome with complexity that you would be unable to act" (Smith 1985: 637). Of course, one ought to do the least amount of violence consistent with not being overwhelmed. This might require successive approximations to a good model that balances abstraction against adequacy. Lakoff's complaints about "objectivism" (1987) can be seen as a claim that "classical" categories defined by necessary and sufficient conditions do too much violence, so that the resulting models are inadequate to the real-world situations.

But I fail to see why complexity makes acting difficult. The real-world situation has precisely the maximal degree of complexity, yet a human *is* capable of acting. Moreover, a complete and

[9] If one understands *neither* domain antecedently, then one might be able to learn both together, either by seeing the same structural patterns in both, or by "getting used to" them both. (Although, possibly, this contradicts the second observation, above.) In this case, neither is the syntactic domain—or else both are!

complex model of some real-world situation might be so complex that a mere human trying to understand *it* might "drown" in its "infinite richness" (Smith 1985: 637), much as a human can't typically hand-trace a very long and complex computer program. Yet a computer can execute that program without "drowning" in its complexity.

Nevertheless, for Smith, "models are inherently *partial*. All thinking, and all computation, ... similarly *have* to be partial: that's how they are able to work" (Smith 1985: 637). Note that some of the partiality of thinking and computation is inherited from the partiality of the model and then compounded: To the extent that thinking and computation use partial descriptions of partial models of the world, they are doubly partial. Much inevitably gets lost in translation, so to speak. Models certainly need to be partial, at least to the extent that the omitted "implementation" details are irrelevant, and certainly to the extent that they (or their descriptions) are discrete whereas the world is continuous. But does thinking "have to be partial" in order to be "able to work"? A *real* thinking thing isn't partial—it is, after all, part of the real world—though its descriptions of models of the world might be partial. And that's really Smith's point—thinking things (and computing things) work with partial models. They "represent the world *as being a certain way*", "*as being one way as opposed to another*" (Smith 1987: 4, 51n.1): They present a fragmentary point of view, a facet of a complete, complex real-world situation—they are objects under a (partial) description (cf. Castañeda 1972).

So we have the following situation. On one side is the real world in all its fullness and complexity. On the other side are partial models of the world and—embedded in computer programs—partial descriptions of the models. But there is a gap between full reality, on the one hand, and partial models and descriptions, on the other, insofar as the latter fail to capture the richness of the former, which they are intended to interact with: Action "is not partial .... When you reach out your hand and grasp a plow, it is the real field you are digging up, not your model of it ... [C]omputers, like us, participate in the real world: they take real actions" (Smith 1985: 637–638). This holds for programs with natural-language competence, too. Their actions are speech acts, and they affect the real world to the extent that communication between them and other natural-language–using agents is successful.

To see how the "reaching out" can fail to cross the gap, consider a blocks-world robot I once saw, programmed with a version of an AI program (Winston 1977) for picking up and putting down small objects at various locations in a confined area. This robot really dealt with the actual world—it was not a simulation. But it did so successfully only by accident. If the blocks were *perfectly* arranged in the blocks-world area, all went well. But if they were slightly out of place—as they were on the day I saw the demo—the robot blindly and blithely executed its program and behaved as if it were picking up, moving, and putting down the blocks. More often, it failed to pick them up, knocked others down as it rotated, and dropped those it hadn't grasped correctly. It was humorous, even pathetic. The robot was doing what it was "supposed" to do, what it was programmed to do, but its partial model was inadequate. Its *successful* runs were, thus, accidental—they worked only if the real world was properly aligned to allow the robot to affect it in the "intended" manner. Clearly, a robot with a more complete model would do better. The Rochester checkers-playing robot, for example, has a binocular vision system that enables it to "see" what it's doing and to bring its motions into alignment with a changing world (Marsh et al. 1992).

So, computers participate in the real world *without interpretations of their behavior by humans* and without the willing participation of humans.[10] Does a computer with natural-language competence really "use language" or "communicate" without a human interpreter? There are two answers: 'yes' and 'no, but so what?':

Yes: As long as the computer is using the vocabulary of some natural language according to its rules of grammar, it is thereby using that language, even if there is no other language-using entity around, including a human. This is true for humans, too: Even if I talk to myself without uttering a sound, I mean things by my silent use of language; sound or other external signs of language-use are not essential to language (Cho 1992). And, therefore, neither is a hearer or other interlocutor.[11]

No; but so what?: A human might interpret the computer's natural-language output differently from how the computer "intended" it. Or one might prefer to say that the computer's output is meaningless until a human interprets it. The output would be mere syntax; its semantics would have to be provided by the human, *although it could be provided by another natural-language–using computer.* However, interpretation problems can arise in human–to–human communication, too. Nicolaas de Bruijn once told me roughly the following anecdote: Some chemists were talking about a certain molecular structure, expressing difficulty in understanding it. De Bruijn, overhearing them, thought they were talking about mathematical lattice theory, since everything they said could be—and was—interpreted by him as being about the mathematical, rather than the chemical, domain. He told them the solution of their problem in terms of lattice theory. They, of course, understood it in terms of chemistry. Were de Bruijn and the chemists talking about the same thing? No; but so what? They *were* communicating!

It is also important to note that when a natural-language–competent computer interacts with a human or another natural-language–competent computer, both need to be able to reach a more-or-less stable state of mutual comprehension. If the computer uses an expression in an odd way (perhaps merely because it was poorly programmed or did not adequately learn how to use that expression), the human (or other interlocutor) must be able to correct the computer—*not* by reprogramming it—but by *telling* it, in natural language, what it should have said. Similarly, if the human uses an expression in a way that the computer does not recognize, the computer must be able to figure out what the human meant. (Cf. Rapaport 1988b and §§2.6.2, 3, below.)

2.5.2 The model–world gap and the third-person point of view.

The gap between model and world is difficult, perhaps impossible, to bridge:

> ... we in general have no guarantee that the models are right—indeed we have no *guarantee* about much of anything about the relationship between model and world. ...
>
> ... [T]here is a very precise mathematical theory called "model theory." You might think that it would be a theory about what models are, what they are good for, how they correspond to the worlds they are models of .... Unfortunately, ... model theory doesn't address the model–world relationship at all. Rather, what model theory

[10] This has *moral* implications, too, as Smith emphasizes in his essay.

[11] Though without an interlocutor, it could not pass the Turing Test; cf. §1.1.2.1.

does is to tell you how your descriptions, representations, and programs *correspond to your model*. (Smith 1985: 638.)

To "address the model–world relationship" requires a language capable of dealing with *both* the model *and* the world. This would, at best, be a "Russellian" language that allowed sentences or propositions to be constructed out of real-world objects (Russell 1903, Moore 1988). It would have to have sentences that explicitly and directly linked parts of the model with parts of the world (recall Helen Keller at the well house). How can such model–world links be made? The only way, short of a Russellian language, is by having *another* language that describes the world, and then provide links between *that* language and the model. (That would have to be done in a meta-language. I am also assuming, here, that the model is a language—a description of the world. If it is a non-linguistic model, we would need yet another language to describe *it*.) But this leads to a Bradleyan regress, for how will we be able to address the relationship between the world and the language that describes it? This parallels the case of the mind, which, insofar as it has no direct access to the external world, has no access to the reference relation.

According to Smith, model theory discusses only the relation (call it R_1) between a model and its description. It does not deal with the relation (call it R_2) between the model and the real-world situation. But if semantics is correspondence, the two cases should be parallel; one ought to be able to deal with both R_1 and R_2, or with neither. Yet we have just seen that R_2 cannot be dealt with except indirectly. Consider R_1. Is it the case that the relation between the computer and the model is dealt with by model theory? No; as Smith says, it deals with the relation between a *description* of the model and the model. After all, the computer is part of the real world (cf. Rapaport 1985/1986: 68). So the argument about the model–world relationship also holds here, for, in the actual computer, there is a physical (real-world) implementation of the model.

Thus, a relation between a syntactic domain and a semantic domain can be understood only by taking an independent, external, third-person point of view. There must be a standpoint—a language, if you will—capable of having equal access to *both* domains. A semantic relation can obtain between two domains, but neither domain can describe that relation by itself. From the point of view of the model, nothing can be said about the world. Only from the point of view of some agent or system capable of taking *both* points of view simultaneously can comparisons be made and correspondences established.

2.5.3 Assymetry.

Smith begins "The Correspondence Continuum" by considering such core semantic or intentional relations as representation and knowledge, "asymmetric" relations that "characterise phenomena that are *about* something, that refer to the world, that have meaning or content" (Smith 1987: 2). What kind of asymmetry is this? Wartofsky has argued that any domain can be used to represent another (cf. our data in §2.3) and that the modeling relation (cognitive agent S takes domain x as a model of domain y) is asymmetric. Yet Wartofsky says that it is not merely that x and y cannot be switched, but rather that in order for S to take x as a model of y, x must (be believed by S to) have fewer relevant properties than y, because if it were "equally rich in the same properties … it would be identical with its object", and if it were "*richer* in properties, … these would then not be ones relevant to its object; it [the object] wouldn't possess them, and so the model couldn't be taken to represent them in any way" (Wartofsky 1979: 5–7).

But it is more appropriate to locate the asymmetry in the fact that the model must be antecedently understood: Suppose an antecedently understood model M of some state of affairs or object O has fewer properties than O, the case that Wartofsky takes to be the norm. Here, the asymmetry between M and O could be ascribed either to M's having fewer properties (as Wartofsky would have it) or to M's being antecedently understood (as I would have it). Suppose, next, that M and O have the same properties. On Wartofsky's view, the asymmetry is lost, but if I antecedently understood M, I could still use M as a model of O: This is the situation of Dennett's Ballad of Shakey's Pizza Parlor (1982: 53–60): Since all Shakey Pizza Parlors are indistinguishable, I can use my knowledge of one to understand the others (e.g., to locate the rest rooms). Finally, suppose that M has *more* (or perhaps merely *different*) properties than O. For example, one could use (the liquidity of) milk as a model of (the liquidity of) mercury (at least, for certain purposes),[12] even though milk has more (certainly, different) properties. These extra (or different) properties are "implementation details"; but they are *merely* that—hence, to be ignored. As long as I antecedently understand M, I can use it as a model of O, no matter how many properties it has. But if I *don't* antecedently understand M, then I *can't* use it as a model (cf. n. 9). And how do I antecedently understand M? By getting used to it.

2.5.4 The continuum.

Smith sees the classical semantic enterprise as a special case of a general theory of correspondence. I see *all* cases of correspondence as being semantic.

Smith distinguishes various types of correspondences. Some semantic relations, e.g., are transitive; others aren't (Smith 1987, e.g., p. 27). Consider, as he suggests, a photo (P_2) of a photo (P_1) of a ship (S). Smith observes that P_2 is not, on pain of use–mention confusion, a photo of S, but that this is "pedantic". Clearly, there are differences between P_1 and P_2: Properties of P_1 *per se* (say, a scratch on the negative) might appear in P_2 and be mistakenly attributed to S. But consider a photo of a map of the world; I *could* use the *photo* as a map to locate, say, Vichy, France. As Smith points out, the photo of the map isn't a map (just as P_2 isn't a photo of S). Yet *information* is preserved, so the photo can be *used as* a map (or: to the extent that information is preserved, it can be so used).

Another of Smith's examples is a document-image–understanding system, which has a knowledge representation of a digital image of a photo (cf. example 3, above). Does the knowledge representation represent the digitized image, or does it represent the photo? The practical value of such a system lies in the knowledge representation representing the photo, not the (intermediate) digitized image. But perhaps, to be pedantic about it, we should say that the knowledge representation does represent the digitized image even though *we* take it *as* representing the photo. After all, the digitized image is internal to the document-image–understanding system, which has no direct access to the photo. Of course, neither do we. Smith seems to agree:

> The true situation ... is this: a given intentional structure—language, process, impression, model—is set in correspondence with one or more other structures, each of which is in turn set in correspondence with still others, at some point reaching (we hope) the states of affairs in the world that the original structures were genuinely about.

[12]Not for understanding its meniscus; the example is due to Kripa Sundar.

It is this structure that I call the 'correspondence continuum'—a semantic soup in which to locate transitive and non-transitive linguistic relations, relations of modelling and encoding, implementation and realisation .... (Smith 1987: 29.)

But can one distinguish among this variety of relations? What makes modeling different from implementation, say? Perhaps one can distinguish between transitive and non-transitive semantic relations,[13] but within those two categories, useful distinctions cannot really be drawn, say, between modeling, encoding, implementation, etc. Perhaps one can say that there are "intended" distinctions, but these cannot be pinned down. Perhaps one can say that it is the person doing the relating who decides, but is that any more than giving different names or offering external purposes? Indeed, Smith suggests (p. 29) that the only differences are individual ones.

He thinks, though, that not all "of these correspondence relations should be counted as genuinely semantic, intentional, representational" (p. 30), citing as an example the correspondences between (1a) an optic-nerve signal and (1b) a retinal intensity pattern, between (2a) the retinal intensity pattern and (2b) light-wave structures, between (3a) light-wave structures and (3c) "surface shape on which sunlight falls", and between (4a) that sunlit surface shape and (4b) a cat. He observes that "it is the cat that I see, not any of these intermediary structures" (p. 30). But so what? Some correspondence relations are not present to consciousness. Nonetheless, they can be treated as semantic.

Not so, says Smith: "correspondence is a far more general phenomenon than representation or interpretation" (p. 30). Perhaps to be "genuinely semantic" (p. 30) is to be *about* something. But why *can't* we say that the retinal intensity pattern is "about" the light-wave structures? Or that the light-wave structures are "about" the sunlit surface shape? The relation between two of these purely physical processes is one of information transfer, so it is surely semantic. Note that it seems to be precisely when phenomena are information-theoretic that models of them *are* the phenomena themselves: Photos of maps *are* maps; models of minds *are* minds.

In any case, what is important for my purposes is Smith's claim that

> the correspondence continuum challenges the clear difference between "syntactic" and "semantic" analyses of representational formalisms ... . [N]o simple "syntactic/semantic" distinction gets at a natural joint in the underlying subject matter. (Smith 1987: 38.)

Although he might be making the point that there can be no "pure" syntactic (or semantic) analyses—that each involves the other—his discussion suggests that the "challenge" is the existence of swing domains.

[13] "In a case where the elements of syntactic domain S correspond to elements of semantic domain D_1, and the elements of D_1 are themselves linguistic, bearing their own interpretation relation to another semantic domain D_2, then the elements of the original domain S are called *metalinguistic*. Furthermore, the semantic relation is taken to be *non-transitive*, thereby embodying the idea of a strict use–mention distinction, and engendering the familiar hierarchy of metalanguages" (Smith 1987: 9). However, it's not clear that S really *is* linguistic (although D_1 *is*), for S will typically consist of *names* of items in D_1, but names are not linguistic in Smith's sense. Second, suppose that S = French, D_1 = English, and D_2 = the actual world. Then the semantic relation *is* transitive, and there is *no* use–mention issue. Here, I am thinking of a machine-translation system, *not* of the case of a French-language textbook written in English (i.e., a textbook whose object language is French and whose metalanguage is English). Clearly, though, there *are* systems of the sort described in this assumption.

2.5.5 Smith's Gap, revisited.

So we have a continuum, or at least a chain, of domains that correspond to one another, each (except the last) understandable in terms of the next. Yet where the last domain is the actual world, Smith's Gap separates it from any model of it. Nonetheless, if that model of the world is in the mind of a cognitive agent—if it is *Cassie's* mental model of the world—then it was constructed (or it developed) by means of perception, communication, and other direct experience or direct contact with the actual world. In terms of Smith's three-link chain (§2.5.1) consisting of a part of the actual world (W), a set-theoretic model of it (M_W), and a linguistic description (in some program) of the model (D_{M_W}), what we have in Cassie's case is that her mental model of the world is simultaneously M_W and D_{M_W}. It is produced by causal links with the external world. Thus, the gap is, in fact, bridged (in this case, at least). Bridged, but not comprehended. In formalizing something that is essentially *in*formal, one can't *prove* (formally, of course) that the formalization is correct; one can only discuss it with other formalizers and come to some (perhaps tentative, perhaps conventional) agreement about it. Thus, Cassie can never check to see if her formal M_W really does match the informal, messy W. Thus, the gap remains. But, once bridged, M_W is independent of W, except when Cassie interacts with W by conversing, asking a question, or acting. That is the lesson of methodological solipsism. Let us turn to Cassie's construction of M_W.

2.6 Cassie's Mental Model.

How does Cassie (or any (computational) cognitive agent) construct her mental model of the world, and what does that model look like? I will focus on her language-understanding abilities—her mental model of a conversation or narrative. (For visual perception, cf. example 3, above.) Details of Cassie's language-understanding abilities have been discussed in a series of earlier papers.[14] Here, I will concentrate on a broad picture of how she processes linguistic input, and a consideration of the kind of world model she constructs as a result.

2.6.1 Fregean semantics.

Frege wanted to divorce logic and semantics from psychology. He told us that terms and expressions (signs, or symbols) of a language "express" a "sense" and that to some—but not all—*senses* there "corresponds" a "referent". So expressions indirectly "designate" or "refer" (or fail to designate or refer) to a referent. Further, the sense is the "way" in which the referent is presented by the expression. Except for the mentalistic notion of an "associated idea", which he does not take very seriously, all of this is very objective or non-cognitive (Frege 1892).

Something exactly like this goes on in cognition, when Cassie—or any natural-language–understanding cognitive agent—understands language:

1. Cassie perceives (hears or reads) a sentence.

2. By various computational processes (namely, an augmented-transition-network parser with lexical and morphological modules, plus various modules for dealing with anaphora resolution,

[14] See, e.g., Shapiro 1982, 1989; Rapaport 1986, 1988b, 1991; Rapaport, Shapiro, & Wiebe 1986; Wiebe & Rapaport 1986, 1988; Almeida 1987; Neal & Shapiro 1987; Shapiro & Rapaport 1987, 1991, 1995; Peters, Shapiro, & Rapaport 1988; Neal et al. 1989; Peters & Rapaport 1990; Wyatt 1990, 1993; Wiebe 1991, 1994; Rapaport & Shapiro 1995.

computing belief spaces and subjective contexts, etc.), she constructs (or finds) a molecular node in the semantic network that is her mental model.

3. That node constitutes her understanding of the perceived sentence.

Now, the procedures that input pieces of language and output nodes are algorithms—*ways* in which the nodes are associated with the linguistic symbols. They are, thus, akin to senses, and the nodes are akin to referents (cf. Wilks 1972). Here, though, all symbols denote, even 'unicorn' and 'round square': If Cassie hears or reads about a unicorn, she constructs a node representing her concept (her understanding) of that unicorn. Her nodes represent the things she has thought about, whether or not they exist—they are part of her "epistemological ontology" (Rapaport 1985/1986).

A very different correspondence can also be set up between natural-language understanding and Frege's theory. According to this correspondence, it is the node in Cassie's mental model that is akin to a sense, and it is an object (if one exists) in the actual world to which that node corresponds that is akin to the referent. On this view, Cassie's unicorn-node represents (or perhaps is) the sense of what she read about, although (unfortunately) there is no corresponding referent in the external world. Modulo the subjectivity or psychologism of this correspondence (Frege would not have identified a sense with an expression of a language of thought), this is surely closer in spirit to Frege's enterprise.

Nonetheless, the first correspondence shows how senses can be interpreted as algorithms that yield referents (a kind of "procedural semantics"; see, e.g., Winograd 1975, Smith 1982). It also avoids the problem of non-denoting expressions: If no "referent" is found, one is just constructed, in a Meinongian spirit (cf. Rapaport 1981).

The various links between thought and language are direct and causal. Consider natural-language generation, the inverse of natural-language understanding. Cassie has certain thoughts; these are private to her.[15] By means of various natural-language–generation algorithms, she produces—directly and causally, from her private mental model—public language: utterances or inscriptions. I hear or read these; this begins the process of natural-language *understanding*. By means of *my* natural-language–understanding algorithms, I interpret her utterances, producing—directly and causally—my private thoughts. Thus, I interpret another's private thoughts indirectly, by directly interpreting her public expressions of those thoughts, which public expressions are, in turn, her direct expressions of her private thoughts. (Cf. Gracia 1990: 495.)

The two direct links are both semantic interpretations. The public expression of Cassie's thoughts is a semantic interpretation (an "implementation" or physical "realization") of her thoughts. And my understanding of what she says is a semantic interpretation of her public utterances. Thus, the public communication language (Shapiro 1993) is a "swing domain".

2.6.2 The nature of a mental model.

Cassie's mental model of the world (including utterances expressed in the public communication language) is expressed in her language of thought. That is, the world is modeled, or represented,

[15] Except, of course, that I, as her programmer and a "computational neuroscientist" (so to speak), have direct access to her thoughts and can manipulate them "directly" in the sense of not having to manipulate them via language. That is, as her programmer, I can literally "read her mind" and "put thoughts into her head". But I ought, on methodological (if not moral!) grounds, to refrain from doing so (as much as possible). I should only "change her mind" via conversation.

by expressions—sentences—of her language of thought (which, for the sake of concreteness, I am taking to be SNePS). There may, of course, be more: e.g., mental imagery. But since Cassie can think and talk about images, they must be linked to the part of her mental model constructed via natural-language understanding (Srihari 1991ab). Hence, we may consider them part of an extended language of thought that allows such imagery among its terms. This extended language of thought, then, is propositional with direct connections to imagistic representations.

In Section 1, I asked how we understand language. This is the challenge of Searle's (1980) Chinese Room Argument: How could Searle-in-the-room come to know the semantics of the Chinese squiggles? One question left open in that debate is whether Searle-in-the-room even knows what their *syntax* is. He could not come to know the syntax (the grammar) just by having, as Searle suggests, a SAM-like program (i.e., a program for global understanding of a narrative using scripts; cf., e.g., Schank & Riesbeck 1981); a syntax-learning program is also needed (see, e.g., Berwick 1979; cf. §1.1.2.2, above). So let us assume that Searle-in-the-room's instruction book includes this.

Given an understanding of the syntax, how can semantics be learned? The meaning of some terms is best learned ostensively, or perceptually: We must see (or hear, or otherwise experience) the term's referent. This ranges from terms for such archetypally medium-sized physical objects as 'cat' and 'cow', through 'red' (cf. Jackson 1986) and 'internal combustion engine', to such abstractions as 'love' and 'think' (cf. Keller 1905: 40f, 300).

But the meaning of many, perhaps most, terms is learned "lexically", or linguistically. Such is dictionary learning. But equally there is the learning, on the fly, of the meaning of new words from the linguistic contexts in which they appear. If 'vase' is unknown, but one learns that Tommy broke a vase, then one can compute that a vase is that which Tommy broke. Initially, this may appear less than informative, though further inferences can be drawn: Vases, whatever they are, are breakable by humans, and all that that entails.[16] As more occurrences of the word are encountered, the "simultaneous equations" (Higginbotham 1989: 469) of the differing contexts, together with background knowledge and some guesswork, help constrain the meaning further, allowing us to revise our theory of the word's meaning. Sooner or later, a provisionally steady state is achieved (pending future occurrences). (See §3.2.2.)

Both methods are contextual. For ostension, the context is physical and external—the real world (or, at least, our perception of it); for the lexical, the context is linguistic (Rapaport 1981). Ultimately, however, the context is mental and internal: The meaning of a term represented by a node in a semantic network is dependent on its location in—i.e., the surrounding context of—the rest of the network (cf. Quine 1951, Quillian 1967, Hill 1994). Such holism has a long and distinguished history, and its share of skeptics (most recently, Fodor and Lepore (1992)). It certainly appears susceptible to charges of circularity (cf., e.g., Harnad 1990), but a chronological theory of how the network is constructed can help to obviate that: Granted that the meaning of 'vase' (for me) may depend on the meaning of 'breakable' *and vice versa*, nonetheless if I learned the meaning of the latter first, it can be used to ground the meaning of the former (for me). Holism, though, has benefits: The meanings of terms get enriched, over time, the more they—or their closest-linked terms in the network—are encountered. For instance, in the research for this essay, certain themes reappeared in various contexts, each appearance enriching the others. In writing, however, one must begin somewhere—writing is a more or less sequential, not a parallel or even

[16]Example due to Karen Ehrlich.

holistic, task. Though this is the first mention of holism in the essay, it was not the first in my research, nor will it be the only one.

Understanding, we see again, is recursive. Each time we understand something, we understand it in terms of all that has come before. Each of those things, earlier understood, were understood in terms of what preceded them. The base case is, retroactively, understandable in terms of all that has come later:

> 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. (Calvino 1986: 19.)

But initially, the base case was understandable solely in terms of itself (or in terms of "innate ideas" or some other mechanism; see Hill 1994 on the semantics of base nodes in SNePS).

But *is* "knowledge of the semantics" (Barwise & Etchemendy 1989: 209) achieved by speakers? If this means knowledge of the relations between word and thing in such a way that it requires knowledge of *both* the words (syntactic knowledge) *and* the things, then: No. For we can't have (direct) knowledge of the things. This is Smith's Gap. It also means, by the way, that ostensive learning is really mental and internal, too: I learn what 'cat' means by seeing one, but really what's happening is that I have a mental representation of that which is before my eyes, and what constitutes the ostensive meaning is a (semantic) link that is established between my internal node associated with 'cat' and the *internal* node that represents what is before my eyes.

Thus, "knowledge of the semantics" means (1) knowledge of the relations *between* our linguistic concepts and our "purely conceptual" concepts (i.e., that correspond to, or are caused by, external input) and (2) knowledge of the relations *among* our purely linguistic concepts. The former (1) is "semantic", the latter (2) "syntactic", as classically construed. Yet, since the former concerns relations among our internal concepts (cf. Srihari 1991ab), it, too, is syntactic.[17]

Barwise and Etchemendy (1989) conflate such an internal semantic theory with a kind of external one, identifying "*content of a speaker's knowledge* of the truth conditions of the sentences of his or her language" with "*the relationship between sentences and non-linguistic facts* about the world that would support the truth of a claim made with the sentence" (p. 220, my italics). I take "the content of a speaker's knowledge of ... truth conditions" to involve knowing the relations between linguistic and non-linguistic *internal* concepts. This is the internal, Cassie-approach to semantics. In contrast, giving an "account of the relationship between sentences and non-linguistic facts" (p. 220) is an *external* endeavor, one that *I* can give concerning Cassie, but not one that *she* can give about herself. This is because *I* can take a "God's-eye", "third-person" point of view and see both Cassie's mind and the world external to it, thus being able to relate them, whereas she can only take the "first-person" point of view.

However, a "third person" cannot, in fact, have direct access to either the external world or Cassie's concepts (except as in n. 15). So what the third person is *really* comparing (or finding correspondences between) is the third person's *representations* of Cassie's concepts and the third person's *own concepts* representing the external world. That is, the third person *can* establish a

[17] The first time you read this, you either found it incomprehensible or insane. By now, it should be less of the former, if not the latter, since its role in the web of my theory should be becoming clearer.

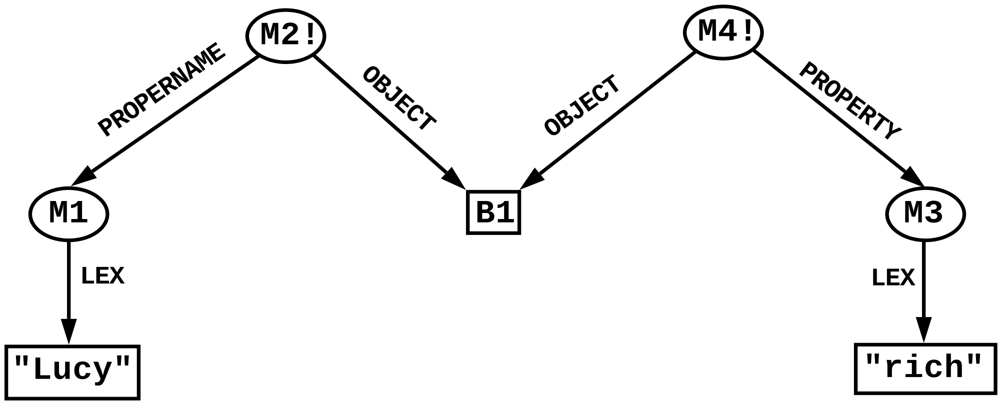

Figure 1: Cassie's belief that Lucy is rich. (Linearly abbreviated: `M4` = `B1` is rich; `M2` = `B1` is named 'Lucy'. Node labels with '`!`' appended are "asserted", i.e., believed by Cassie.)

---

semantic correspondence (in the classic sense) between two domains. From the third person's point of view, the two domains are the syntactic domain of Cassie's concepts and the semantic domain of the external world. But, in fact, the two domains are *the third person's representations* of Cassie's concepts and *the third person's representations* of the external world. These are both *internal* to the third person's mind! And internal relations, even though structurally *semantic*—i.e., even though they are correspondences between two domains—are fundamentally *syntactic* in the classic sense: They are relations *among* (two classes of) symbols in the third person's language of thought.

What holds for the third person holds also for Cassie. Since she doesn't have direct access to the external world either, she can't have knowledge of "real" semantic correspondences. The best she can do is to have a correspondence between certain of her concepts and her representations of the external world. What might her "knowledge of truth conditions" look like? Here is one possibility: When she learns that Lucy is rich, she builds the network shown in Figure 1. Thus, Cassie might think to herself something like: "My thought that [Figure 1] is true iff ($\exists\, x \in$ external world)[x = Lucy & x is rich]". This would require, for its full development, (1) an internal truth predicate (cf. Maida & Shapiro 1982, Neal & Shapiro 1987), (2) an existence predicate (cf. Hirst 1991), (3) a duplication of the network, and (4) a biconditional rule asserting the equivalence (see Figure 2 for a possible version).

The picture we have of Cassie's mental model of the world (including utterances) is, in part, this: If Cassie hears or reads a sentence, she constructs a mental representation of that sentence qua linguistic entity, *and* she constructs a mental representation of the state of affairs expressed by that sentence. These will be linked by a Tarski-like truth-biconditional (`M8`) asserting that the representation of the sentence (`M7`) is true (`M6`) iff the representation of the state of affairs (`M7`) is believed (`M7!`). If Cassie sees something, she constructs a mental representation of it, *and* she constructs a mental propositional representation of the state of affairs she sees. These will be linked in ways extrapolatable from §2.3, example 3, above. These networks, of course, are not isolated, but embedded in the entire network that has been constructed so far. What is newly perceived is understood in terms of all that has gone before. This is purely syntactic, since both sides of the biconditional are expressed in Cassie's language of thought. Thus, the best Cassie can do is to have a theory of truth as coherence among her own concepts.

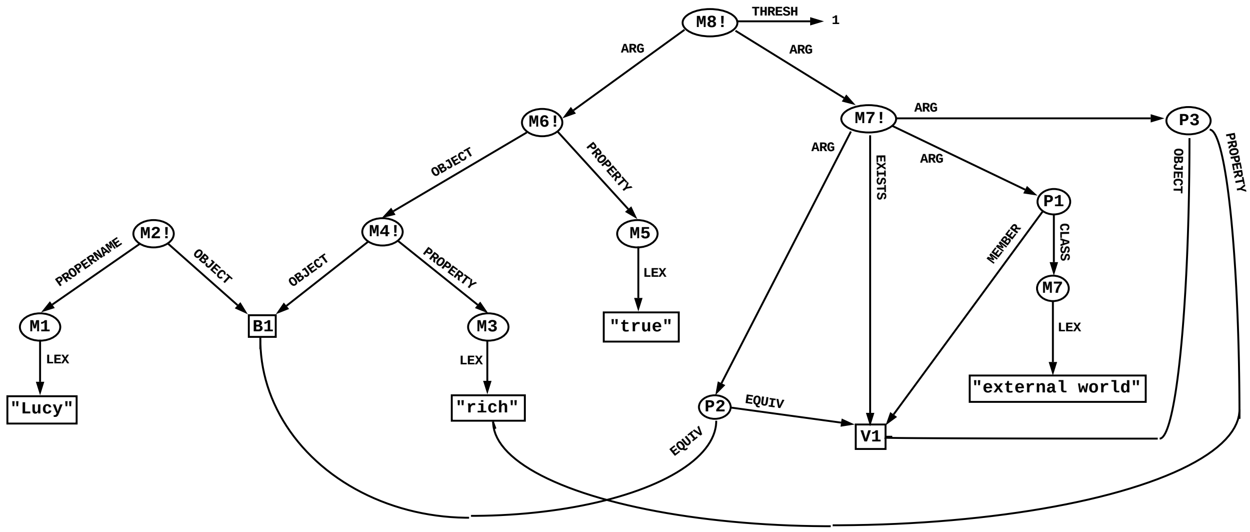

Figure 2: A biconditional rule (`M8`) asserting the equivalence of `M6` = that Lucy is rich is true, and `M7` = something in the external world is Lucy and is rich. (See Shapiro 1979, Shapiro & Rapaport 1987 for the semantics of `thresh`. The truth condition for `M2` is not shown.)

3 SEMANTICS AS SYNTAX.

3.1 The Story So Far.

To understand language is to construct a semantic interpretation—a model—of the language. In fact, we *normally* understand something by modeling it and then determining correspondences between the two domains. In some cases, we are lucky: We can, as it were, keep an eye on each domain, merging the images in our mind's eye. In other cases, notably when one of the domains is the external world, we are not so lucky—Smith's Gap cannot be crossed—and so we can understand that domain *only* in terms of the model. Lucky or not, we understand one thing in terms of another by modeling that which is to be understood (the syntactic domain) in that which we antecedently understand (the semantic domain). But how is the antecedently understood domain antecedently understood? In the base case of our recursive understanding of understanding, a domain must be understood in terms of *itself*, i.e., syntactically.

3.2 Syntactic Understanding.

3.2.1 Familiarity breeds comprehension.

What is type-2, syntactic understanding? What does it mean to "get used to" something? In some sense, it should be obvious:

> In today's chess, only the familiarly shaped Staunton pieces are used. ... [One] reason is the unfamiliarity, to chess players, of other than Staunton pieces. ... [In Reykjavik, in 1973, two grandmasters] started to play [with a non-Staunton set], and the conversation ran something like:

> "What are you doing? That's a pawn."
> "Oh. I thought it was a bishop."
> "Wait! Maybe it is a bishop."
> "No, maybe it really is a pawn."
>
> Whereupon the two grandmasters decided to play without the board. They looked at each other and this time the conversation ran:
>
> "D5"
> "C4"
> "E6"
> "Oh, you're trying *that* on me, are you? Knight C3."
>
> And they went along that way until they finished their game. (Schonberg 1990: 38–39.)

In a game played with Staunton pieces, the players are "used to" the pieces. Even in a game played with no physical pieces at all, the players are "used to" the symbolic notation for the pieces. But in a game played with non-Staunton pieces, clearly they are not.

Suppose that the semantic relation is (merely) a correspondence relation. Suppose, further, that it is a homomorphism mapping the syntactic domain into the semantic domain. To understand something in terms of itself would then be to take the syntactic domain as its own semantic domain, treating the homomorphism as an *automorphism*. Such an automorphism would be a relation among the symbols of the syntactic domain, hence a classically syntactic relation. Yet it would also be a *semantic* relation, because it is a correspondence between "two" domains (better: between two roles played by the same domain). Indeed, the very first example of a semantic model in Chang & Keisler 1973 is such a mapping. The syntactic domain, now considered as its own *semantic* domain, is syntactic twice over: once by way of its own, purely syntactic, features, and once by way of the semantic automorphism. (Recall the way some linguists do semantics; cf. §2.1.)

Now, the automorphism is either the identity mapping, or it isn't. If it is, then the symbol manipulations (the syntax) that constitute the semantics are just those of the syntactic domain itself. This is the core meaning of understanding by "getting used to" the system (as in the syntactic way of understanding §2.2.2's L') We do the same thing when we learn how to solve algebraic equations by manipulating symbols (Rapaport 1986).

If the automorphism is *not* the identity mapping, then it must map some elements onto others (or sets of others). So some parts of the syntactic domain will be understood in terms of others. (There may be "fixed points"—symbols that *are* mapped into themselves; see §3.2.3.) Let's see what this means for our central case—natural-language understanding.

3.2.2 Dictionary definitions and algebra.

Dictionary-like definitions are an obvious example of this sort of automorphism. Indeed, this is probably what *most* people mean by "meaning", as opposed to philosophers, logicians, and cognitive scientists—though *some* cognitive scientists are sympathetic: "meanings are, if anything, only other symbols" (Wilks & Fass 1992: 205; cf. Wilks 1971, 1972: 86).

And, as noted earlier (§2.6.2), we learn the meaning of many (if not most) new words in linguistic contexts—either in explicit definitions or "on the fly" in ordinary conversational or literary discourse. The unknown word, like the algebraic unknown, simply means whatever is necessary to give meaning to the entire context in which it appears. The meaning of the unknown word is (the meaning of) the surrounding context—the context "minus" the word. Finding the meaning is, thus, "solving" the context for the unknown: "The appearance of a word in a restricted number of settings suffices to determine its position in the language as a whole" (Higginbotham 1985: 2; cf. Wilks 1971: 519–520). As Wilks 1971 notes, the context must be suitably large to get the "correct" or at least "intended" meaning. But, as the 'vase' example shows (§2.6.2), any context will do for starters. One's understanding of the meaning of the word will *change* as one comes across more contexts in which it is used (or: as the total context becomes larger); ultimately, one's understanding of the meaning will reach a stable state (at least temporarily—everything is subject to revision). Thus, learning a word is theory construction: One's understanding of the word's meaning is a *theory*, subject to revision.

The first time I read the word 'brachet', I did not know what it meant (do you?). Here is the context of that first occurrence:

> [T]here came a white hart running into the hall with a white brachet next to him, and thirty couples of black hounds came running after them with a great cry. (Malory 1470: 66.)

My first hypothesis (believe it or not) was that a brachet was a buckle on a harness worn by the hart. Although this hypothesis goes beyond the algebraic picture I've been painting, the algebraic metaphor is still reasonable if we extend the notion of context to include the background knowledge ("world knowledge" or "commonsense knowledge") that I bring to bear on my understanding of the narrative (cf. Rapaport 1991, Rapaport & Shapiro 1995). Nor does it matter whether this hypothesis is good, bad, indifferent, or just plain silly; if I never see the word again, it won't matter, but, if I do, I will have ample opportunity to revise my beliefs about its meaning. Indeed, after 18 more occurrences of the term,[18] I stabilized on the following theory of its meaning: A brachet is a hound or hunting dog, perhaps a lead hound. Not bad, considering that the *Oxford English Dictionary* defines it as "A kind of hound which hunts by scent". (For details and further references, see Rapaport 1981; Ehrlich & Rapaport 1992; Ehrlich, forthcoming.)

This is purely syntactic: First, there is no external semantic domain: I did not see a brachet (or a picture of one). Second, when I read the word (or when Cassie does), I build a mental representation of that word embedded in a mental representation of its context. These mental representations are part of the entire network of mental representations in my mind. Thus, the background knowledge I contribute is part and parcel of the mental representation of the new word in context. It is that system of mental representations that constitutes the syntactic domain in which is located "the meaning" of—i.e., my understanding of—the word.

Representing meaning in such a dictionary-like network goes back at least to Quillian 1967, though he was more concerned with merely representing the information in a dictionary, whereas I am concerned with representing meaning as part of a cognitive agent's entire complex network of beliefs. This is a brand of holistic, conceptual-role semantics, since I take the meaning of a word to be, algebraically, the role it plays in its context.

[18] The protocols appear in Ehrlich, forthcoming.

3.2.3 Understanding the parts.

Another thing that using parts of the syntactic domain to understand the rest of it might mean is that those parts are primitives. How are *they* understood? What do *they* mean?

They might be "markers" with no intrinsic meaning. But such markers *get* meaning the more they are used—the more roles they play in providing meaning to *other* nodes. A helpful analogy comes from Wartofsky:

> '[M]ental' objects, or 'internal representations' are derivative, and have their genesis in our primary activity of representing, in which we take external things,—most typically, what we also designate as physical objects—as representations. Moreover, I take our *making* of representations to be, in the first place, the actual praxis of creating concrete objects-in-the-world, *as* representations; or of taking the made objects as representational. (1979: xxi–xxii.)

In this *primary* activity, what do we take the external physical object to be a representation *of*? Its use and history? What does that have to do with the common properties in terms of which one thing can represent another? More likely, it is that, once made, it can *remind us* of its use or of its manufacture and therefore represent those things for us. The *first* time we see an unfamiliar object, it is a meaningless thing (except insofar as it shares any properties with anything familiar, allowing us to form hypotheses about it and to place it in our semantic network). The *second* time we see it, it can remind us of something, if only of itself on its first appearance. Subsequent encounters produce familiarity, which entrench it in our network, and allow newer objects to be understood in terms of *it*. This is how holism works.

Alternatively, the fixed points or the markers (or—for that matter—any of the nodes) are somehow "grounded" in another domain. This, of course, is just to say that they have meaning in the correspondence sense of semantics, and ultimately we will be led to question the way in which we understand that other domain.

3.2.4 The symbol-grounding problem.

The symbol-grounding problem, according to Harnad (1990), is that without grounding, a hermetically sealed circle of nodes can only have circular meaning. And, presumably, circles are vicious and to be avoided.

It is well known that a dictionary is a closed circle of meanings: Each word is defined in terms of other words. Assuming that all words used in the definitions are themselves defined, we have a circle (in fact, several of them). Now, before agreeing with Harnad that such circles cannot yield meaning or understanding, consider that we do use dictionaries fairly successfully to learn meanings. How can this be? A very small circle may indeed not be informative, especially if we don't antecedently know the meaning of the definiens (e.g., 'being' is defined as "existence", and 'exist' as "have being" in *Webster's Vest Pocket Dictionary*). However, the larger the circumference of the circle, so to speak, the more likely it is that it will be informative, on the assumption that the definition of the word whose meaning we seek, and the definitions of the words in that definition, and so on, will contain *lots* of words that we antecedently understand. So, we can easily "solve" the "equation" for the unknown word—i.e., dictionary definitions are most useful to the extent that they are like '$x = (4 - 3)/2$', rather than like '$x = (4y - 3)/2$' (where there is a further unknown in the definiens).

Nonetheless, some words will still only be poorly defined in terms of other words, notably (but not exclusively) nouns like 'cat' or 'cow'. For these, *seeing* a cat or cow (or for 'love', *experiencing* love) is worth a thousand-word definition. Illustrated dictionaries handle this with a type distinction in the syntax of the definiens: Terms can be of the type *word* or the type *picture*. But although we now have grounding in an extra-linguistic system, *it is still part of the dictionary*. And, of course, the pictures could, with a suitable indexing scheme, themselves be definiendum entries: A picture of a cat could have as its "definiens" the word 'cat', as in a visual dictionary or a field guide to cats. This, of course, only widens the circle. We could widen it further, albeit at some expense and inconvenience: Let every dictionary come with a real cat; ditto for all other better-ostensively-defined terms. We still have a circle, but now, I think, Harnad would have to agree that we've also got grounding—we've merely incorporated the groundings into the dictionary.

The same holds for the mind. Harnad says a link is needed between (some) mental nodes (say, our "cat" node) and items in the external world (say, a cat). Although we can't import such items directly into our minds, we *can* have mental representations of them. And it is the relation between our "cat" node and a node representing a perceived cat that grounds the former. We saw how in §2.6.2: What we (or Harnad) *think* is the relation between word and world is really a connection between an internal representation of a perceived *word* and an internal representation of the perceived *world*.

Do not misunderstand me: Experience certainly enriches our understanding. Consider "immersion" learning of a foreign language. "Thinking in French" is understanding French holistically, without any correspondences to one's native language (say, English). It is helped immeasurably by living in a francophone community. When we ask "What does the French word '*chat*' mean?", and we give the answer ("cat") in English words, we are doing pure *syntax* (here, relating symbols from one system to those of another) *that is also semantic* (understanding one system in terms of another). This is no different than answering the question in French ("*un chat est un petit animal domestique, dont il existe aussi plusieurs espèces sauvages*"[19])—except for choice of language for the definiens. Giving the definition in English is just as much symbol grounding as pointing to a cat would be. Symbol grounding, thus, does *not* necessarily get us out of the circle of words—at best, it widens the circle. That is my point: Syntactic understanding—the base case of understanding—is just a *very* wide circle.

Suprisingly, Harnad's own examples of grounding are internal in just the ways we have been considering. He distinguishes between "symbolic" and "non-symbolic" representations (Harnad 1990: 335, Abstract). But *both* are *internal* representations. Harnad says that non-symbolic "*iconic representations* ... are **internal** analog transforms of the projections of distal objects on our sensory surfaces" (p. 342; Harnad's italics, my boldface). Such a projection could be a retinal image, say. So an iconic representation is some "analog transform" of *that*, stored (or created) somewhere further along the optic pathway. Thus, it could be part of our semantic network (cf. Srihari 1991ab). Furthermore, the symbolic and non-symbolic representational systems must be linked; hence, because of Smith's Gap, they must all be internal.

One would *expect* internal items to be "grounded" in *external* ones. But in Harnad's hier-

[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!

archy (p. 335, Abstract), symbolic representations are "grounded" in "elementary symbols", which are "names" of "categories", which categories are "assigned on the basis of" categorial representations, which representations are "derived" from sensory projections; and iconic representations are "analogs" of those sensory projections. In fact, the only actual use of the term 'grounding' is between symbolic representations and elementary symbols, *both of which are internal*. Indeed, *all* the items on this hierarchy are internal!

Curiously, Harnad only mentions "grounding in the world" in a footnote:

> If a candidate model [for a cognitive system] were to exhibit all ... behavioral capacities, both *linguistic* ["produce" and "respond to descriptions of ... objects, events, and states of affairs"] ... and *robotic* ["discriminate, ... manipulate, ... [and] identify ... the objects, events and states of affairs **in the world they live in**] ..., it would pass the "total Turing test" .... A model that could pass the total Turing test, however, would be **grounded in the world**. (Harnad 1990: 341fn13; Harnad's italics, my boldface.)

Recall the blocks-world and checkers-playing robots (§2.5.1). The former is blind and methodologically solipsistic. The latter can see. But is it grounded? Could it be fooled as the blind robot was? Possibly: by Cartesianly deceiving its eyes. It would then "live in" a world in which to be was to be perceived. Of course, such a Berkeleyan robot *would* be grounded *in the world that it lives in*, which happens not to be the actual world, but a purely intensional one. (In this case, note that the grounding system *and* the grounded one are *both* internal, hence part of a single network.) What would such a robot's symbols mean *to it*? Here, internal semantic interpretation would be done by internal links only.

Even more curious is the fact that grounding for Harnad—even grounding in the external world—does not seem to serve a semantic function:

> Iconic representations no more "mean" the objects of which they are the projections than the image in a camera does. Both icons and camera images can of course be *interpreted* as meaning or standing for something, but the interpretation would clearly be derivative rather than intrinsic. (p. 343.)

Harnad seems to be saying here that the causal connection of the iconic representations with its real-world counterpart is irrelevant to its intrinsic meaning. In that case, Harnad owes us answers to two questions: (1) what does such a causal grounding *do* in his theory, and (2) what *is* the intrinsic meaning of an iconic representation? Harnad may have identified an interesting problem, but he doesn't seem to have solved it.

My position is this: The mind–world gap cannot be bridged by the mind. There are causal links between them, but the only role these links play in semantics is this: The mind's internal representations of external objects (which internal representations are caused by the external objects) *can* serve as "referents" of other internal symbols, but, since they are *all* internal, meaning is in the head and is syntactic.

A purely syntactic system is ungrounded, up in the air, self-contained. But there are arbitrarily many ways to ground it; that is, there are infinitely many possible interpretations for any syntactic system. By "communicational negotiation", we (agree to) ground our language of thought in equivalent ways for all practical purposes (cf. Bruner 1983). Harnad seeks a *natural* grounding (cf. "intended interpretation"). Some candidates, such as the (human) body, are convenient.[20] But such natural groundings are merely one or two among many. The only one that is non-arbitrary is the "null" grounding, the "self" grounding: the purely syntactic, "internal" mode of understanding. " 'Explanations come to a stop' as ... Wittgenstein would put it; 'there is a last house in the lane' " (Leiber 1991: 54; cf. Wittgenstein 1958: §1, p. 3e, and §29, p. 14e). The last semantic domain in a correspondence continuum is the "last house in the lane". It can only be understood syntactically. Hence, all understanding rests on syntactic understanding.

This is one of the flaws in Searle's Chinese-Room Argument. Part of his argument is that computers can never understand natural language because (1) understanding natural language requires (knowledge of) semantics, (2) computers can only do syntax, and (3) syntax is insufficient for semantics. I take *my* argument to have shown that (3) is false (and that, therefore, (2) is misleading, since the kind of syntax that computers do *ipso facto* allows them to do semantics).

4 SUMMARY.

We understand one domain recursively in terms of an antecedently understood one. The "base" case is the case in which a domain is understood in terms of itself. When a syntactic domain is its own semantic domain, the semantic interpretation function either maps the symbols to themselves or else to other symbols. In the former case, we understand the domain by "getting used to" its syntax. In the latter case, if there are no fixed points—if each symbol is mapped to a different one—then we have the situation we face when using a dictionary. The difference is that since all external items are also mapped into internal ones, the symbol-grounding problem can be avoided. If there *are* fixed points, then they come to be understood either retroactively in terms of the role they play in the understanding of other terms, or else by "grounding" them to "non-linguistic"—albeit *internal*—symbols.

In any case, we have a closed network of meaning—a holistic, "conceptual-role semantics". And that is how semantics can arise from syntax.

5 REFERENCES.

1. Almeida, Michael J. (1987), "Reasoning about the Temporal Structure of Narratives," *Technical Report 87–10* (Buffalo: SUNY Buffalo Department of Computer Science).

2. Barwise, Jon, & Etchemendy, John (1989), "Model-Theoretic Semantics," in M. I. Posner (ed.), *Foundations of Cognitive Science* (Cambridge, MA: MIT Press): 207–243.

3. Berwick, Robert C. (1979), "Learning Structural Descriptions of Grammar Rules from Examples," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence (IJCAI–79, Tokyo)* (Los Altos, CA: William Kaufmann): 56–58.

4. Brachman, Ronald J., & Schmolze, James G. (1985), "An Overview of the KL-ONE Knowledge Representation System," *Cognitive Science* 9: 171–216.

[20] For discussion of various aspects of this, see Leiber 1980, Johnson 1987, Lakoff 1987, Turner 1987, Kirsh 1991, Perlis 1991, Dreyfus 1992.

5. Bruder, Gail A.; Duchan, Judith F.; Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; & Zubin, David A. (1986), "Deictic Centers in Narrative: An Interdisciplinary Cognitive-Science Project," *Technical Report 86–20* (Buffalo: SUNY Buffalo Department of Computer Science).

6. Bruner, Jerome (1983), *Child's Talk: Learning to Use Language* (New York: W. W. Norton).

7. Calvino, Italo (1986), "Why Read the Classics?" *The New York Review of Books* (9 October 1986): 19–20.

8. Castañeda, Hector–Neri (1972), "Thinking and the Structure of the World," *Philosophia* 4 (1974) 3–40; reprinted in 1975 in *Critica* 6 (1972) 43–86.

9. Chang, C. C., & Keisler, H. J. (1973), *Model Theory* (Amsterdam: North–Holland).

10. Cho, Kah–Kyung (1992), "Re-thinking Intentionality" (in Japanese), in Y. Nitta (ed.), *Tasha-no Genshogaku (Phenomenology of the Other)* (Hokuto Publishing Co.).

11. Corless, R. M. (1992), "Continued Fractions and Chaos," *American Mathematical Monthly* 99: 203–215.

12. Dennett, Daniel C. (1982), "Beyond Belief," in A. Woodfield (ed.), *Thought and Object* (Oxford: Clarendon Press): xvi–95.

13. Dreyfus, Hubert L. (1992), *What Computers* Still *Can't Do: A Critique of Artificial Reason* (Cambridge, MA: MIT Press).

14. Duchan, Judith F.; Bruder, Gail A.; & Hewitt, Lynne (eds.) (1995), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).

15. Ehrlich, Karen (forthcoming), "Automatic Vocabulary Expansion through Natural-Language Contexts," *Technical Report* (Buffalo: SUNY Buffalo Department of Computer Science).

16. Ehrlich, Karen, & Rapaport, William J. (1992), "Automatic Acquisition of Word Meanings from Natural-Language Contexts," *Technical Report 92–03* (Buffalo: SUNY Buffalo Center for Cognitive Science, July 1992). Fodor, Jerry A. (1975), *The Language of Thought* (New York: Thomas Y. Crowell Co.).

17. Fodor, Jerry, & Lepore, Ernest (1992), *Holism: A Shopper's Guide* (Cambridge, MA: Basil Blackwell).

18. Frege, Gottlob (1892), "On Sense and Reference," M. Black (trans.), in P. Geach & M. Black (eds.), *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Basil Blackwell, 1970): 56–78.

19. Gracia, Jorge J. E. (1990), "Texts and Their Interpretation," *Review of Metaphysics* 43: 495–542.

20. Hardt, Shoshana L. (1992), "Conceptual Dependency," in S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence*, 2nd edition (New York: John Wiley & Sons): 259–265.

21. Harnad, Stevan (1990), "The Symbol Grounding Problem," *Physica D* 42: 335–346.

22. Higginbotham, James (1985), "On Semantics," reprinted in E. LePore (ed.), *New Directions in Semantics* (London: Academic Press, 1987): 1–54.

23. Higginbotham, James (1989), "Elucidations of Meaning," *Linguistics and Philosophy* 12: 465–517.

24. Hill, Robin (1994), "Issues of Semantics in a Semantic-Network Representation of Belief," *Technical Report No. 94–11* (Buffalo: SUNY Buffalo Department of Computer Science).

25. Hirst, Graeme (1991), "Existence Assumptions in Knowledge Representation," *Artificial Intelligence* 49: 199–242.

26. Jackson, Frank (1986), "What Mary Didn't Know," *Journal of Philosophy* 83: 291–295.

27. Johnson, Mark (1987), *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* (Chicago: University of Chicago Press).

28. Kamp, Hans (1984), "A Theory of Truth and Semantic Representation," in J. Groenendijk, T. M. V. Janssen, & M. Stokhof (eds.), *Truth, Interpretation, and Information* (Dordrecht: Foris): 1–41.

29. Kamp, Hans, & 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).

30. Keller, Helen (1905), *The Story of My Life* (Garden City, NY: Doubleday, 1954).

31. Kirsh, David (1991), "Foundations of AI: The Big Issues," *Artificial Intelligence* 47: 3–30.

32. Lakoff, George (1987), *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press).

33. Leiber, Justin (1980), *Beyond Rejection* (New York: Ballantine Books).

34. Leiber, Justin (1991), *An Invitation to Cognitive Science* (Cambridge, MA: Basil Blackwell).

35. Lytinen, Steven L. (1992), "Conceptual Dependency and Its Descendants," *Computers and Mathematics with Applications* 23: 51–73.

36. Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks," *Cognitive Science* 6: 291–330.

37. Malory, Sir Thomas (1470), *Le Morte Darthur*, ed. and trans. by R. M. Lumiansky (New York: Collier Books, 1982).

38. Marsh, Brian; Brown, Chris; LeBlanc, Thomas; Scott, Michael; Becker, Tim; Quiroz, Cesar; Das, Prakash; & Karlsson, Jonas (1992), "The Rochester Checkers Player: Multimodal Parallel Programming for Animate Vision," *Computer*, Vol. 25 No. 2 (February 1992) 12–19.

39. Mayes, A. R. (1991), Review of [*inter alia*] H. Damasio & A. R. Damasio, *Lesion Analysis in Neuropsychology*, in *British Journal of Psychology* 82: 109–112.

40. Minsky, Marvin (1975), "A Framework for Representing Knowledge," in P. H. Winston (ed.), *The Psychology of Computer Vision* (New York: McGraw-Hill).

41. Moore, Robert C. (1988), "Propositional Attitudes and Russellian Propositions," *Report No. CSLI–88–119* (Stanford, CA: Center for the Study of Language and Information).

42. Morris, Charles (1938), *Foundations of the Theory of Signs* (Chicago: University of Chicago Press).

43. Neal, Jeannette G., & Shapiro, Stuart C. (1987), "Knowledge Based Parsing," in L. Bolc (ed.), *Natural Language Parsing Systems* (Berlin: Springer–Verlag).

44. Neal, Jeannette G.; Thielman, C. Y.; Dobes, Zuzanna; Haller, Susan M.; & Shapiro, Stuart C. (1989), "Natural Language with Integrated Deictic and Graphic Gestures," *Proceedings of the DARPA Speech and Natural Language Workshop* (Morgan Kaufmann): 14.

45. Perlis, Donald (1991), "Putting One's Foot in One's Head—Part I: Why," *Noûs* 25: 435–455.

46. Peters, Sandra L., & Rapaport, William J. (1990), "Superordinate and Basic Level Categories in Discourse: Memory and Context," *Proceedings of the 12th Annual Conference of the Cognitive Science Society (Cambridge, MA)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 157–165.

47. Peters, Sandra L.; Shapiro, Stuart C.; & Rapaport, William J. (1988), "Flexible Natural Language Processing and Roschian Category Theory," *Proceedings of the 10th Annual Conference of the Cognitive Science Society (Montreal)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 125–131.

48. Pincus, Andrew L. (1990), "The Art of Trnascription Sheds New Light on Old Work," *The New York Times*, Arts and Leisure (Sect. 2) (23 September 1990).

49. Potts, Timothy C. (1973), "Model Theory and Linguistics," in E. L. Keenan (ed.), *Formal Semantics of Natural Language* (Cambridge, Eng.: Cambridge University Press, 1975): 241–250.

50. Quillian, M. Ross (1967), "Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities," *Behavioral Science* 12: 410–430.

51. Quine, Willard Van Orman (1951), "Two Dogmas of Empiricism," reprinted in W. V. Quine, *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 2nd ed., revised, 1980): 20–46.

52. Rapaport, William J. (1978), "Meinongian Theories and a Russellian Paradox," *Noûs* 12: 153–180; errata, *Noûs* 13 (1979) 125.

53. Rapaport, William J. (1981), "How to Make the World Fit Our Language: An Essay in Meinongian Semantics," *Grazer Philosophische Studien* 14: 1–21.

54. Rapaport, William J. (1985/1986), "Non-Existent Objects and Epistemological Ontology," *Grazer Philosophische Studien* 25/26: 61–95.

55. Rapaport, William J. (1986), "Logical Foundations for Belief Representation," *Cognitive Science* 10: 371–422.

56. Rapaport, William J. (1988a), "To Think or Not to Think," *Noûs* 22: 585–609.

57. Rapaport, William J. (1988b), "Syntactic Semantics: Foundations of Computational Natural-Language Understanding," in J. H. Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht, Holland: Kluwer Academic Publishers): 81–131.

58. Rapaport, William J. (1991), "Predication, Fiction, and Artificial Intelligence," *Topoi* 10: 79–111.

59. Rapaport, William J. (1992a), "Logic, Predicate," in S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence*, 2nd edition (New York: John Wiley): 866–873.

60. Rapaport, William J. (1992b), "Logic, Propositional," in S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence*, 2nd edition (New York: John Wiley): 891–897.

61. Rapaport, William J. (1993), "Cognitive Science," in A. Ralston & E. D. Reilly (eds.), *Encyclopedia of Computer Science*, 3rd edition (New York: Van Nostrand Reinhold): 185–189.

62. Rapaport, William J., & Shapiro, Stuart C. (1995), "Cognition and Fiction," in J. F. Duchan, G. A. Bruder, & L. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).

63. Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1986), "Quasi–Indicators, Knowledge Reports, and Discourse," *Technical Report 86–15* (Buffalo: SUNY Buffalo Department of Computer Science).

64. Rosenblueth, Arturo, & Wiener, Norbert (1945), "The Role of Models in Science," *Philosophy of Science* 12: 316–321.

65. Russell, Bertrand (1903), *The Principles of Mathematics* (New York: W. W. Norton, 1937).

66. Schank, Roger C., & Rieger, Charles J. (1974), "Inference and the Computer Understanding of Natural Language," *Artificial Intelligence* 5: 373–412.

67. Schank, Roger C., & Riesbeck, Christopher K. (eds.) (1981), *Inside Computer Understanding: Five Programs Plus Miniatures* (Hillsdale, NJ. Lawrence Erlbaum).

68. Schneiderman, Ben (1993), "Data Type," in A. Ralston & E. D. Reilly (eds.), *Encyclopedia of Computer Science*, 3rd edition (New York: Van Nostrand Reinhold): 411–412.

69. Schonberg, Harold C. (1990), "Some Chessmen Don't Make a Move," *New York Times* (15 April 1990), Sect. 2, pp. 38–39.

70. Searle, John R. (1980), "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3: 417–457.

71. Searle, John R. (1993), "The Failures of Computationalism," *Think* (Tilburg, The Netherlands: Tilburg University Institute for Language Technology and Artificial Intelligence) 2 (June 1993) 68–71.

72. Segal, Erwin M. (1995), "Stories, Story Worlds, and Narrative Discourse," in J. F. Duchan, G. A. Bruder, & L. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).

73. Segal, Erwin M.; Duchan, Judith F.; & Scott, Paula J. (1991), "The Role of Interclausal Connectives in Narrative Structuring: Evidence from Adults' Interpretations of Simple Stories," *Discourse Processes* 14: 27–54.

74. Shapiro, Stuart C. (1979), "The SNePS Semantic Network Processing System," in N. Findler (ed.), *Associative Networks* (New York: Academic Press): 179–203.

75. Shapiro, Stuart C. (1982), "Generalized Augmented Transition Network Grammars for Generation from Semantic Networks," *American Journal of Computational Linguistics* 8: 12–25.

76. Shapiro, Stuart C. (1989), "The CASSIE Projects: An Approach to Natural Language Competence," *Proceedings of the 4th Portugese Conference on Artificial Intelligence (Lisbon)* (Springer–Verlag): 362–380.

77. Shapiro, Stuart C. (1993), "Belief Spaces as Sets of Propositions," *Journal of Experimental and Theoretical Artificial Intelligence* 5: 225–235.

78. Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in N. Cercone & G. McCalla (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer–Verlag): 262–315.

79. Shapiro, Stuart C., & Rapaport, William J. (1991), "Models and Minds: Knowledge Representation for Natural-Language Competence," in R. Cummins & J. Pollock (eds.), *Philosophy and AI: Essays at the Interface* (Cambridge, MA: MIT Press): 215–259.

80. Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family," *Computers and Mathematics with Applications* 23: 243–275.

81. Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narrative," in J. F. Duchan, G. A. Bruder, & L. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).

82. Smith, Brian Cantwell (1982), "Linguistic and Computational Semantics," *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics (University of Toronto)* (Morristown, NJ: Association for Computational Linguistics): 9–15.

83. Smith, Brian Cantwell (1985), "Limits of Correctness in Computers," in C. Dunlop & R. Kling (eds.), *Computerization and Controversy* (San Diego: Academic Press, 1991): 632–646.

84. Smith, Brian Cantwell (1987), "The Correspondence Continuum," *Report No. CSLI–87–71* (Stanford, CA: Center for the Study of Language and Information).

85. Sowa, John F. (1984), *Conceptual Structures: Information Processing in Mind and Machine* (Reading, MA: Addison-Wesley).

86. Sowa, John F. (1992), "Conceptual Graphs as a Universal Knowledge Representation," *Computers and Mathematics with Applications* 23: 75–93.

87. Srihari, Rohini K. (1991a), "PICTION: A System that Uses Captions to Label Human Faces in Newspaper Photographs," *Proceedings of the 9th National Conference on Artificial Intelligence (AAAI-91, Anaheim)* (Cambridge, MA: AAAI/MIT Press): 80–85.

88. Rohini K. Srihari (1991b), "Extracting Visual Information from Text: Using Captions to Label Faces in Newspaper Photographs," *Technical Report 91–17* (Buffalo: SUNY Buffalo Department of Computer Science.

89. Srihari, Rohini K., & Rapaport, William J. (1989), "Extracting Visual Information From Text: Using Captions to Label Human Faces in Newspaper Photographs," *Proceedings of the 11th Annual Conference of the Cognitive Science Society (Ann Arbor, MI)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 364–371.

90. Srihari, Rohini K., & Rapaport, William J. (1990), "Combining Linguistic and Pictorial Information: Using Captions to Interpret Newspaper Photographs," in D. Kumar (ed.), *Current Trends in SNePS—Semantic Network Processing System*, *Lecture Notes in Artificial Intelligence, No. 437* (Berlin: Springer–Verlag): 85–96.

91. Tanenbaum, Andrew S. (1976), *Structured Computer Organization* (Englewood Cliffs, NJ: Prentice-Hall).

92. Tenenbaum, Aaron M., & Augenstein, Moshe J. (1981), *Data Structures using Pascal* (Englewood Cliffs, NJ: Prentice-Hall).

93. Turner, Michael (1987), *Death Is the Mother of Beauty* (Chicago: University of Chicago Press).

94. Wartofsky, Marx W. (1979), *Models: Representation and the Scientific Understanding* (Dordrecht: D. Reidel).

95. Wiebe, Janyce M. (1991), "References in Narrative Text", *Noûs*, 25: 457–486.

96. Wiebe, Janyce M. (1994), "Tracking Point of View in Narrative," *Computational Linguistics* 20.

97. Wiebe, Janyce M., & Rapaport, William J. (1986), "Representing *De Re* and *De Dicto* Belief Reports in Discourse and Narrative," *Proceedings of the IEEE* 74: 1405–1413.

98. Wiebe, Janyce M., & Rapaport, William J. (1988), "A Computational Theory of Perspective and Reference in Narrative" *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics (SUNY Buffalo)* (Morristown, NJ. Association for Computational Linguistics): 131–138.

99. Wilkins, David P. (1992), "Interjections as Deictics," *Journal of Pragmatics* 18: 119–158.

100. Wilks, Yorick (1971), "Decidability and Natural Language," *Mind* 80: 497–520.

101. Wilks, Yorick (1972), *Grammar, Meaning and the Machine Analysis of Language* (London: Routledge & Kegan Paul).

102. Wilks, Yorick, & Fass, Dan (1992), "The Preference Semantics Family," *Computers and Mathematics with Applications* 23: 205–221.

103. Winograd, Terry (1975), "Frame Representations and the Declarative/Procedural Controversy," in D. G. Bobrow & A. M. Collins (eds.), *Representation and Understanding* (New York: Academic Press): 185–210.

104. Winston, Patrick Henry (1977), *Artificial Intelligence* (Reading, MA: Addison–Wesley).

105. Wittgenstein, Ludwig (1958), *Philosophical Investigations: The English Text of the Third Edition*, trans. by G. E. M. Anscombe (New York: Macmillan).

106. Woods, William A., & Schmolze, James G. (1992), "The KL-ONE Family," *Computers and Mathematics with Applications* 23: 133–177.

107. Wyatt, Richard (1990), "Kinds of Opacity and Their Representations," in D. Kumar (ed.), *Current Trends in SNePS—Semantic Network Processing System*, Lecture Notes in Artificial Intelligence, No. 437 (Berlin: Springer-Verlag): 123–144.

108. Wyatt, Richard (1993), "Reference and Intensions," *Journal of Experimental and Theoretical Artificial Intelligence* 5: 263–271.