

How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room

William J. Rapaport

**Department of Computer Science and Engineering,
Department of Philosophy, Department of Linguistics,
and Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260-2000**

`rapaport@cse.buffalo.edu`

`http://www.cse.buffalo.edu/~rapaport/`

August 2, 2006

Abstract

The shortest answer to the question, “How can a computer come to understand natural language?” is: “The same way that Helen Keller did”, viz., by using “syntactic semantics”—a theory of how syntax can suffice for semantics, i.e., how semantics for natural language can be provided by means of computational symbol manipulation. This essay considers real-life approximations of Chinese Rooms, focusing on Helen Keller’s experiences growing up deaf and blind, locked in a sort of Chinese Room yet learning how to communicate with the outside world. Using a computational knowledge-representation system (SNePS), the essay analyzes her belief that learning that “everything has a name” was the key to her success, enabling her to partition her mental concepts into mental representations of words, mental representations of objects, and mental representations of the naming relations between them. It next looks at Herbert Terrace’s theory of naming, which is akin to Keller’s, and which only humans are supposed to be capable of. The essay suggests that computers at least, if not also non-human primates, are also capable of this kind of naming.

1 Chinese Rooms, Syntactic Semantics, and SNePS

Philosophy is the history of a deaf-blind person writ large.

—Helen Keller (1903: 25), cited in Leiber 1996: 426.

The shortest answer to the question, “How can a computer come to understand natural language?” is: “The same way that Helen Keller did.”

—Albert Goldfain, personal communication, 2006.

This essay investigates the nature of Chinese Rooms and how to escape from them. We first consider a few real-life approximations of them before focusing on one in particular: Helen Keller’s experiences growing up deaf and blind. We will examine what her experiences were like locked in a sort of Chinese Room yet learning how to communicate successfully—indeed, brilliantly—with the outside world.

In Keller’s own view, learning that “everything has a name” was the key to her success. We’ll use the SNePS knowledge-representation and reasoning system to help explain the sense in which this was true and its relationship to “syntactic semantics”—a theory of how the kind of syntactic symbol manipulation that computers do so well suffices for providing the kind of semantics needed for natural-language understanding (*pace* Searle 1980, 1993: 68).

I also critique Herbert Terrace’s theory of naming, which seems akin to Keller’s. Terrace believes that only humans are capable of naming in his sense, hence that non-human primates cannot learn language. I shall argue that at least computers, if not also non-human primates, are also capable of naming in Terrace’s sense, hence that, if it *is* necessary for learning language, then computers (and, perhaps, non-human primates) can understand language.

1.1 A Quiz

Let’s begin with a brief reminder of what a Chinese Room is. Here’s a little quiz to help you remember: Who said the following, and when?

Consider a box B inside of which we have a man L with a desk, pencils and paper. On one side B has two slots, marked *input* and *output*. If we write a number on paper and pass it through the input slot, L takes it and begins performing certain computations. If and when he finishes, he writes down a number obtained from the computation and passes it back to us through the output slot. Assume further that L has with him explicit deterministic instructions of finite length as to how the computation is to be done. We refer to these instructions as P. Finally,

assume that the supply of paper is inexhaustible, and that B can be enlarged in size so that an arbitrarily large amount of paper work can be stored in it in the course of any single computation. . . . I think we had better assume, too, that L himself is inexhaustible, since we do not care how long it takes for an output to appear, provided that it does eventually appear after a finite amount of computation. We refer to the system B-L-P as M. . . . In the approach of Turing, the symbolism and specifications are such that the entire B-L-P system can be viewed as a digital computer Roughly, to use modern computing terms, L becomes the logical component of the computer, and P becomes its program. In Turing's approach, the entire system M is hence called a *Turing machine*. (Rogers 1959: 115, 117.)

If you answered "John Searle, in 1980", you've got the wrong person and you're off by some 20 years: It was written by Hartley Rogers in 1959 (pp. 115, 117)! Searle (1980), of course, later made such a "box" famous as the "Chinese Room", used in a thought experiment to create a counterexample to the Turing test (Turing 1950, Shieber 2004, Rapaport 2005cd). What, then, is a Chinese Room?

1.2 Chinese Rooms

Let us define a "Chinese Room" as a particular kind of Turing machine conceived as in Rogers's specification, i.e., a system consisting of two subsystems: a program and a program executor.

1.2.1 The Program

The first subsystem is a (computer) program, i.e., an implementation, in some language, of an algorithm. In a Chinese Room, the program (1) accepts as input what are in fact (written or spoken) expressions in a natural language (archetypally Chinese), and (2) manipulates the input (3) to yield as output "appropriate" expressions in that language.¹

The "appropriateness" of the output depends on the kind of input. Thus, if the input is a story together with "reading comprehension" questions about it, then "appropriate" output would be whatever would count as "appropriate" answers to

¹Often, programs are characterized merely in input-output terms. I use the phrase "manipulate the input to yield the output" in order to emphasize the *algorithmic* nature of the program. (The "manipulation", of course, could be a null operation, in which case the algorithm (or program) would indeed be a mere input-output device.) However, to be even more accurate, we should say that the program *describes how to* accept and manipulate the input, since a program is a static (usually textual) object, as opposed to a dynamic "process". See §1.2.2, below.

the questions. They need not be correct or non-trivial² answers, but at least they should be relevant ones. Or, if the input is a fragment of a conversation in Chinese, then “appropriate” output would be whatever would count as an “appropriate” continuation of the conversation in Chinese. This could include changing the subject in the way that sometimes happens in ordinary conversations, but should probably not include the kinds of abrupt changes that one sees in Eliza-like programs that try to avoid topics that they are not designed to handle (Weizenbaum 1966).

These are *de re* descriptions of the input and output, i.e., descriptions from a third-person point of view. Viewed *de dicto*—i.e., from a first-person point of view, from the *program’s* point of view, so to speak—the program accepts as input uninterpreted marks (‘meaningless squiggles’; Searle 1980: 417),³ and manipulates them to produce other such marks as output.

This is a natural-language-processing Chinese Room. There could also be Chinese Rooms for visual processing or other kinds of cognition. Given the ‘AI-completeness’ of these tasks,⁴ probably any Chinese Room needs to handle all such tasks. Thus, a more general description of a Chinese Room would not restrict the input-output to natural-language expressions but—as with Rogers’s box—would allow any expressions from some syntactic system.

1.2.2 The Program Executor

Since a program is a static object, a second subsystem is required that causes the program to be executed, i.e., that creates a dynamic ‘process’ consisting of actually *taking* the input, *manipulating* it, and *producing* the output.⁵

In a computer, this second subsystem would be the central processing unit (CPU), which fetches and executes the instructions (by implementing the switch-settings specified by the program). In Searle’s Chinese Room Argument, the second subsystem is a human, named ‘Searle’, who (consciously) follows the instructions of the program. In Rogers’s box, it is L (‘the logical component of the computer’).

²Albert Goldfain (personal communication, 2006) pointed out to me that some of Eliza’s responses are trivially appropriate (see Weizenbaum 1966).

³For discussion on the relation of ‘symbols’ to what I call ‘marks’ (roughly, uninterpreted symbols, for readers who will excuse the apparent oxymoronic nature of that phrase), see Rapaport 1995 and, especially, Rapaport 2000.

⁴A problem is AI-complete if a (computational) solution for it requires or produces (computational) solutions to *all* problems in AI (Shapiro 1992).

⁵Cf. Richmond Thomason’s (2003) characterization of a computer as a device that ‘change[s] variable assignments’, i.e., that accepts certain assignments of values to variables as input, changes (i.e., manipulates) them, and then outputs the changed values.

In the Chinese Room Argument, Searle-in-the-room (i.e., the executor) by hypothesis does not understand Chinese. Thus, the “squiggles” that are input and output are meaningless *to Searle-in-the-room*; *he* does not, or cannot, interpret them. All that he has access to is the *syntax* of the input. On one interpretation of the Chinese Room Argument (apparently Searle-the-philosopher’s interpretation), even though Searle-in-the-room does not understand Chinese, he is having a fluent conversation in Chinese (i.e., is passing a Turing test for Chinese natural-language understanding). Therefore (either by analogy or by universal generalization), no computer—more precisely, no Chinese Room executor—equipped with a Chinese natural-language-understanding program could understand Chinese. Thus, natural-language understanding, or cognition more generally, is not computable.

1.2.3 The System

Searle-the-philosopher assumes that it is Searle-in-the-room who does not understand Chinese, i.e., that the Chinese Room executor does not understand Chinese. But Searle-the-philosopher concludes that no *computer* can understand Chinese solely computationally. Yet the computer is something over and above the two subsystems; it is a third thing: the system itself, consisting of the program interacting with the executor. Some readers might complain that the combined system is not a single (perhaps “emergent”) thing but is completely analyzable (or reducible) to its two subsystems. If this is the case, then so much the worse for Searle’s analogy. For humans are surely single entities (at least from the perspective of cognitive psychology), and it is a human who understands natural language, not a part of a human (i.e., a brain). Since, by analogy, no one thinks it necessary for a CPU to “know” or “understand” what it is doing, there is no reason for Searle-in-the-room’s lack of knowledge or understanding of Chinese to prevent the *system* from understanding Chinese. It is the entire computer that processes data, not any single part of it (e.g., only its CPU or its memory; cf. Rogers, above, and §1.3, below).

In any case, the system is *not* a mere set of its two subsystems. The program by itself can do nothing, and the executor without a program to execute has nothing to do. On this view, it is not Searle-in-the-room *simpliciter* (the executor) who is conversing in Chinese. Rather, the entire Chinese Room *system* consisting of the executor interacting with the program is conversing in Chinese. This “systems reply” to the Chinese Room Argument seems clearly related to notions of “wide” or “extended” cognition (Hutchins 1995ab, Clark & Chalmers 1998; cf. Giere 2002.) But to say that the *system* understands Chinese might be to describe the situation from the system’s first-person point of view. Alternatively, one could say that the system can be *said* to understand Chinese. But this must be to describe the

situation from a third-person point of view, say, the point of view of the native Chinese speaker outside the room. A third alternative might be this: Chinese natural-language understanding is being produced. This more neutral description seems to be a “view from nowhere” (Nagel 1986). But can there be understanding without an understander? (Cf. my discussions of Taylor 2002 and of Proudfoot 2002 in Rapaport 2006.)

On any of these alternatives—whether it is Searle-in-the-room or the entire Chinese Room system that is alleged to understand Chinese, or whether “it is understanding” in the same way that we say “it is raining” (i.e., understanding is occurring without anything *doing* the understanding, just as rain occurs without a “rainer”)—Searle-the-philosopher argues that the mere syntactic symbol-manipulation being undertaken in the Room does not suffice to yield semantic understanding (Searle 1980, 1993). What appears to be missing are meanings to be “attached” to the squiggles (Searle 2002: 53; cf. Rapaport 2006). In the “robot reply” to the Chinese Room Argument, these are entered into the Room via sensors.⁶ But, according to Searle-the-philosopher, these are just more symbols, not meanings, and so nothing is gained. I will suggest below that more symbols is all that a cognitive agent gets, hence that they *must* suffice for understanding, and so nothing is lost.

1.3 Syntactic Semantics

How could Searle-in-the-room come to have knowledge of the semantics of the (Chinese) squiggles? How can *we* have knowledge of the semantics of *our* language? This is the challenge posed by Searle’s Chinese Room Argument. Allegedly, such knowledge is required in order for Searle-in-the-room to “understand” Chinese (more generally, to “understand” the natural-language input; more generally still, to “understand” *simpliciter*). As we have seen, however, the program executor in a Chinese Room does *not* need to have knowledge of the semantics of the input, for it is not the executor who needs such knowledge or who understands. Rather, it is the entire system that needs such knowledge or that understands. In Rogers’s system, it is not L who computes, but “the entire system M” consisting of L, the program, and the box itself—i.e., the Turing machine.

For Searle-the-philosopher, this is the task of explaining how Searle-in-the-room (or the system) could know what the symbols *are about*, not (merely)

⁶Effectors are also provided, to enable Searle-in-the-room or the system to manipulate the environment, though this may be less essential, since almost no one would want to claim that a quadriplegic or a brain-in-a-vat with few or no effectors was not capable of cognition. Cf. Maloney 1987, 1989, and discussion in Rapaport 1993, 1998, 2000. See also Anderson 2003, §5; Chrisley 2003, fn. 25.

what their grammatical syntax is, i.e., how they are related *to other things* (their meanings), not merely how they are related *to each other* (their grammatical syntax). Of course, Searle-in-the-room (or the system) also needs to know the grammatical syntax. How? Not, presumably, *just* by having the sort of program that Searle 1980 envisioned, namely, a Schankian Script Applier Mechanism (Cullingford 1981). A syntax-learning program is necessary. There has certainly been a lot of work on the problem of learning grammatical syntax, so this is not science fiction. So let's assume that Searle-in-the-room (or the system) has such a program (cf. Rapaport 1995, §2.6.2). Given an understanding of the grammatical syntax, how much semantics can be learned? Quite a bit, or so I have argued in a series of earlier essays on the theory of "syntactic semantics".⁷

The theory of syntactic semantics has three basic theses. *Thesis 1*: The first of these is that a computer (or a cognitive agent) can take two sets of symbols with relations between them and treat their union as a single syntactic system in which the previously "external" relations are now "internalized". Note that initially there are *three* things: two sets (of things)—which may have a non-empty intersection—and a third set (of relations between them) that is not part of either set of things, i.e., that is external to both sets. One set of things can be thought of as a cognitive agent's mental entities (thoughts, concepts, etc.). The other can be thought of as "the world" (in general, the meanings of the thoughts, concepts, etc., that are in the first set). The relations are intended to be the semantic relations of the mental entities to the objects in "the world". Note that these semantic relations are neither among the agent's mental entities nor in "the world". (Of course, all three sets are "in the world" in the sense that *everything* is in the world.) At best, they are accessible only from a third-person point of view (though I will argue later that even that is not the case). In the case of a Chinese Room, one set might be the squiggles, the other set might be their meanings, and the "external" relations might be semantic interpretations of the former in terms of the latter.

When the two sets are unioned, the meanings become "internalized". That is, in order for the agent to understand the world, he or she (or it) must "push the world into the mind" (Jackendoff 2002, §10.4; cf. Rapaport 2003a), so that (analogues of) the world-set and the relations-set are among the agent's mental entities. Now the agent's mental entities include *both* the original thoughts, concepts, etc., *and* representatives of the (formerly) external objects, *and* representatives of the (formerly) external relations. Before the union, the semantic relations obtained between *two* sets; now, their mental analogues obtain within a *single* set. Hence, they are syntactic: Classically, semantics is the study of relations between two sets,

⁷Rapaport 1985, 1986b, 1988ab, 1990, 1993, 1995, 1996, 1998, 1999, 2000, 2002, 2003a, 2005b, 2006.

and syntax is the study of relations among the members of a single set (Morris 1938).

As an example, Stanislas Dehaene (1992: 30–32) argues that “numbers may be represented mentally in three different codes”: an “auditory verbal code” (i.e., a mental “word sequence”), a “visual arabic [sic] code” (i.e., mental Arabic numerals), and an “analogue magnitude code” (i.e., a mental number line). The first two are what I would call “purely” syntactic representational systems. The third, I would call a “semantic” syntactic system: It is syntactic, because it is mental—implemented in the very same kind of neurons as the other two codes. Yet it provides an internal semantic interpretation for the other two codes.

In this way, the syntax of the new, unioned set *can* suffice for semantics; i.e., it can syntactically handle the semantic system. In this way, computers and cognitive agents *can* understand natural language via a syntactic, holistic, conceptual-role semantics, with the help of negotiation (Rapaport 2002, 2003a).⁸ Note that, for a network, any semantics that does not go outside the network must be a conceptual-role semantics (see Rapaport 2002). This theory has received a partial computational implementation in a research project on “contextual vocabulary acquisition”.⁹

Thesis 2: Some might prefer to say that the purely syntactic system (merely “models” the semantic relationship between the symbols (e.g., the Chinese squiggles) and their (real-world/external-world) meanings. But I take the internalization of these real-world/external-world meanings seriously; this is, in fact the second thesis of the theory: We only have direct access to internal representatives of objects in the external world. I discuss this further in §§2.1 and 2.2.

Thesis 3: The third thesis is that understanding is recursive. Briefly, we understand one system (a syntactic domain) in terms of another (a semantic domain) that is antecedently understood. The base case must be a system that is understood in terms of itself, i.e., syntactically. We’ll look at this at the end of the essay (§7).

I have argued elsewhere (e.g., Rapaport 2000) that syntactic semantics shows how syntax does suffice for semantics (contra Searle 1980) and allows Searle-in-the-room to “escape” from the Chinese Room. The present essay offers further evidence in support of this.

One standard objection to such a theory is that the system is “closed” or

⁸And/or some kind of “dynamic” or “incremental” semantics along the lines of, e.g., Discourse Representation Theory (Kamp & Reyle 1993) or Dresner’s (2002) algebraic-logic approach.

⁹Ehrlich 1995, 2004; Ehrlich & Rapaport 1997, 2004; Rapaport & Ehrlich 2000; Rapaport & Kibby 2002; Rapaport 2003b, 2005a; Kibby et al., forthcoming.

“circular” (e.g., Harnad 1990; cf. Rapaport 1995).¹⁰ Just as all words in a dictionary are ultimately defined in terms of other words in the dictionary (see, e.g., Spärck Jones 1967), in our unioned set all elements are highly interconnected and, almost by definition, not connected to anything “outside” the set. To fully understand any term defined in a dictionary, one apparently has to have some independent knowledge that links the dictionary information to the external world. First, this may not strictly be true even for dictionaries, especially if they contain pictures. Such pictures correspond to internalizations of the external objects. More significantly, our brain is just such a closed system: All information that we get from the external world, along with all of our thoughts, concepts, etc., is represented in a single system of neurons (or neuron firings). Any description of that real neural network, by itself, is a purely syntactic one. Yet the syntax of that real neural network is what provides our semantic understanding of language and the world. If this implies some kind of Quinean holism or indeterminacy, so be it (see note 24).

1.4 The SNePS Knowledge-Representation, Reasoning, and Acting System

In the next section, we will look at some real-life Chinese Rooms. The first example (in §2.1) may need a bit of background that will be useful later on as well.

SNePS is a computational system for knowledge-representation, reasoning, and acting. SNePS can be viewed as a propositional semantic network whose nodes represent propositions and (their constituent) concepts (including other propositions—typically, embedded ones) and whose labeled, directed arcs encode the relationships of constituent concepts to the concepts and propositions of which they are constituents.¹¹

A simple SNePS network is shown in Figure 1. Here, node m3 represents the proposition that humans are mammals.¹² The exclamation point on the node label

¹⁰Objections to holistic theories in general are replied to in Rapaport 2002, 2003a.

¹¹For other ways of viewing SNePS, see Shapiro et al. 2006. For details, see, e.g., Shapiro 1979; Shapiro & Rapaport 1987, 1992, 1995; Shapiro 2000. Further information is available online at: [<http://www.cse.buffalo.edu/sneps>] and at: [<http://www.cse.buffalo.edu/~rapaport/snepskrra.html>].

¹²Here, it is represented via a subclass-superclass relationship. There are other ways to represent this in SNePS. E.g., it could also be represented as one of the universally quantified propositions:

For all x , if x is human (i.e., if object x has the property of being human),
then x is a (member of the class) mammals,

or else:

For all x , if x is a human (i.e., if x is a member of the class humans),

indicates that the proposition is ‘believed’ (technically, ‘asserted’, i.e., accepted as true) by the system. The proposition is analyzed as claiming that a class represented by node *m1*, and expressed in English by the lexeme ‘humans’, is a subclass of a superclass represented by node *m2*, lexicalized as ‘mammals’.¹³ It is important to note that although the network “encodes” the notions of ‘subclass’, ‘class’, and ‘lexical entry’, it does not explicitly *represent* these. The only things that are explicitly represented are (a) two lexical entries (represented by: *humans*, *mammals*), (b) two entities (represented by: *m1*, *m2*) that are characterized only structurally as concepts associated with those lexemes (the association is shown by the arcs; see Shapiro & Rapaport 1987 for further discussion of the semantics of the *lex* case frame), and (c) another entity (represented by: *m3*) that is an asserted proposition structurally characterized only in terms of *m1* and *m2*.¹⁴ SNePS can only ‘talk about’ or have ‘beliefs about’ *nodes*, not arcs.¹⁵ ‘Cassie’ is the name we have given to a computational cognitive agent implemented in SNePS; this network shows that Cassie believes that humans are mammals.¹⁶

Information can be entered into SNePS in one of two ways: The SNePS User Language (SNePSUL) uses a Lisp-like syntax to describe networks directly in terms of their labeled, directed graph structure. Thus, for example, the network of Figure 1 can be built by first defining the arc labels *lex*, *subclass*, and *superclass*, and then issuing the command

```
(assert subclass (build lex humans)
              superclass (build lex mammals))
```

This can be read ‘inside out’: First, nodes *m1* and *m2* are built (or retrieved, if they had previously been built) with *lex* arcs emanating from them, one to a

then *x* is a (member of the class) mammals,

or in various other ways depending on the choice of ontology (which, in turn, determines the choice of arc labels (or ‘case frames’)). This, however, is a choice that SNePS leaves up to each user. E.g., an alternative to the *lex* arc is discussed in §3.4.

¹³The labels on the nodes at the heads of the *lex* arcs are somewhat arbitrary. For expository convenience, I am using English plural nouns. But they could just as well have been singular nouns or even arbitrary symbols (e.g., ‘b1’, ‘b2’). The important points are that (1) the nodes *represent* lexical expressions in some language and (2) the nodes are ‘aligned’ (in a technical sense; see Shapiro & Ismail 2003) with entries in a lexicon. E.g., had we used ‘b1’ instead of ‘humans’, the lexical entry for *b1* could indicate that its morphological ‘root’ is ‘human’, and an English morphological synthesizer could contain information about how to modify that root in various contexts.

¹⁴On the notion of ‘structural’, as opposed to ‘assertional’, characterizations, see: Woods 1975; Shapiro & Rapaport 1987, 1991.

¹⁵For further discussion, see Shapiro & Rapaport 1987 and § 3.4, below.

¹⁶For more information on Cassie, see: Shapiro & Rapaport 1987, 1991, 1992, 1995; Shapiro 1989, 1998; Rapaport, Shapiro, & Wiebe 1997; Rapaport 1991a, 1998, 2000, 2002, 2003a; Ismail & Shapiro 2000; Shapiro, Ismail, & Santore 2000; Shapiro & Ismail 2003; Santore & Shapiro 2004.

node labeled `humans` and one to a node labeled `mammals`. Second, node `m3` is built (or retrieved) with a `subclass` arc to `m1` and a `superclass` arc to `m2`. Finally, `m3` is “asserted”, which means that it is treated as a proposition that Cassie “believes” or takes as true, denoted by the exclamation mark. (Unasserted propositional nodes represent propositions that Cassie is merely thinking about without being committed to.)

The second language for entering information into a network, which we will discuss further beginning in §3.4.2, is SNePSLOG, a language with a syntax like that of a language for predicate logic. In such a language, a predicate or function-symbol is characterized in terms of a *sequence* of arguments. In SNePSLOG, a function symbol¹⁷ is defined in terms of a *set* of arguments (a “case frame”) distinguished by (arc) labels. To create the network of Figure 1 using SNePSLOG, we would first define two case frames (i.e., sets of arc labels). In SNePS (hence in both SNePSUL and SNePSLOG), the user (or “knowledge engineer”) has quite a bit of freedom to choose arc labels and case frames (the only restrictions are due to some pre-defined case frames used by the reasoning system; see Shapiro & Rapaport 1987). In SNePSLOG, the user also has freedom to choose function symbols that, essentially, name these case frames. A function symbol (i.e., a case-frame name) can be identical to one of its arc labels; this often has mnemonic value for the human user. For clarity, however, I will use different names. For example, here are two function-symbol definitions:

1. `define-frame thing-called (nil lex)`
2. `define-frame AKO (nil subclass superclass)`

Case-frame 1 defines a function symbol that can be used to build nodes `m1` and `m2` of Figure 1, and case-frame 2 defines a predicate (actually, another function symbol) that can be used to build node `m3`, as follows:

`AKO(thing-called(humans), thing-called(mammals)).`

(The period causes `m3` to be asserted.) The semantics of these case frames is as follows:¹⁸

1' $[[\text{thing-called}(w)]]$ = the concept associated with (or expressed by) the lexical entry (i.e., the word) w .

¹⁷Actually, SNePS *only* has function symbols, but no predicates; all well-formed formulas are terms, none are sentences in the usual sense, although some terms can be “asserted”, which means that Cassie (or the system) treats them as true sentences.

¹⁸Here and elsewhere, I use surrounding double-brackets (“[[]]”) to denote a function that takes as input the symbol inside the brackets and returns a meaning for it.

2' $[[AKO(x,y)]] = x$ is a kind of y (or: x is a subclass of the superclass y).

(The “hil” in the `define-frame` syntax serves to ensure that the function symbol is not also an arc label; hence, it plays no role in the semantics.) Thus, the SNePSLOG proposition that built `m3` can be understood as saying that a thing called ‘humans’ is a kind of (or: is a subclass of a) thing called ‘mammals’.

In what follows, I will use SNePSLOG syntax in the text for readability (because of its similarity to a language for predicate logic)¹⁹ as well as network diagrams in the figures (for their graphical value).

¹⁹The *logic* underlying SNePS is *not* that of ordinary predicate logic! It is a paraconsistent relevance logic with an automatic belief-revision system that can allow for any believed (i.e., asserted) proposition to be withdrawn (i.e., unasserted), along with any of its inferred consequents, in the light of later beliefs. See Shapiro & Rapaport 1987, Martins & Shapiro 1988, Shapiro 2000.

2 What Is It Like to Be in a Chinese Room?

Before seeing how to “escape” from a Chinese Room, let us consider a few *actual* situations that give the reader some semblance of what it might be like to be in one.

2.1 The Japanese Room

Readers who are conversant in reading SNePS networks can get a feel for what it might be like to be Cassie by considering the network shown in Figure 2 for a Japanese-speaking computational cognitive agent implemented in SNePS (from Arahi & Momouchi 1990: 2). My first reaction on seeing this SNePS network was that indeed I couldn’t understand it. But why should I? It only matters for *Cassie* (or her Japanese counterpart) to understand it. But even that isn’t quite accurate. Cassie does not need to *understand it*. Rather, she only has to be able to understand *with* it; i.e., she only has to *use* that network in order to understand *other* things (e.g., language). In the same way, we humans do not need to understand our neuron firings in order for us to be able to understand with, or by means of, our neuron firings.

I, of course, can only understand the semantic network in Figure 2 by mapping it to *my* concepts, and there’s insufficient information in that figure alone for me to do that in any but a non-arbitrary way. Each SNePS network in Arahi & Momouchi 1990 (as in Figure 2) has *English* arc labels but *Japanese* node labels for the nodes at the heads of `lex` arcs (call these `lex` nodes, for convenience). The Japanese networks’ use of English arc labels makes it *appear* that the arc labels convey some information to Cassie. They don’t. They only convey information to *us*; but that’s irrelevant. (For an elaboration of this, see McDermott 1981.) They serve as “punctuation” or structuring devices only. True, the “reader” of a network uses the arc labels to help understand the network, but that’s akin to the use of self-documenting variable names in a high-level programming language: useful to an “external” reader but not essential to the compiler or interpreter. True, too, SNePS’s natural-language module uses the arc labels to generate appropriate English expressions (see §3.4.2, below), but—again—the arc labels could equally well have been more-or-less arbitrary strings of characters. What counts is *that* the arcs are labeled in certain ways, not *what* the labels are. The “*content*” of the networks is in the `lex` nodes (cf., e.g., Rapaport 1988b).

Nonetheless, the networks seem incomprehensible! When I first saw them, I felt like Searle-in-the-room. I can manipulate the networks, but I don’t know what they mean. What’s missing is a link—a correspondence—between the Japanese `lex` nodes and *my* neural analogues of `lex`-nodes *or else* between the Japanese `lex` nodes and things in the world. The latter, though, is impossible; at best that

would be *representations* in my mind of things in the world. So this reduces to the first alternative (cf. Rapaport 2000, 2002, 2003a).

At this point, a digression concerning the second thesis of syntactic semantics is necessary. Simply put, it is the Kantian thesis that we cannot have direct access to “things-in-themselves”, i.e., to real-world objects external to our minds. All we can have access to are internal entities that are representatives of them, filtered through our cognitive mechanisms. (Following Shapiro & Ismail 2003, I will call the external items “objects” and the internal ones “entities”.) (The legion of readers who may disagree with the second thesis of syntactic semantics can consult Rapaport 2000, 2005b for my reasoning, which is a version of the Argument from Illusion. See also §2.2, below.) Thus, on this view, semantic relations obtain between two systems of mental entities. The first might be, for example, internal representatives of Chinese squiggles, and the second might be, for example, internal representatives of external objects. The latter would have to be augmented with other mental entities that might not have direct external counterparts (e.g., such “abstract” concepts as love, justice, $\sqrt{-1}$, etc., as well as “non-existents”—or Meinongian objects—such as unicorns, golden mountains, or round squares; see, e.g., Rapaport 1981).²⁰ Each of these two systems will have its own grammatical syntax. (The syntax of a set of squiggles will be a grammatical syntax in the linguistic sense. The syntax of a set of internal representatives of external objects will be an ontology, in the AI sense.) The fully internal relations between the two systems (which can be considered as internal representatives of the classical word-world semantic relations) will be purely syntactic, too. More precisely, their union will have two kinds of (internal) relations: grammatical ones and semantic ones. For instance, in natural language, there are grammatical properties and relations among the words in the sentence “Ricky loves Lucy” (subject, verb, object), and there are semantic relations between the words ‘lawyer’ and ‘attorney’ (synonymy), or ‘tall’ and ‘short’ (antonymy). But, in the classical sense of syntax (Morris 1938), both of these sorts of relations are syntactic.

The important point is that if all items are internal (e.g., all are represented by neuron firings), then they are purely syntactic. The alternative is that we are restricted to relations between internal entities and external objects. In that case, we are in no different a situation than a computer: Any relations that we have with the external world could be had by a computer. But, if we can understand the semantics of natural language in this way, so can a computer. Note, though, that such relations are neither in our minds nor among the objects in the world that we are trying to understand (to repeat: except insofar as everything is in the world),²¹

²⁰Cf. Meinong 1904; Rapaport 1976, 1978, 1981, 1985/1986, 1991b.

²¹Such internal-external relations are still being *used* syntactically, after the fashion of the

so it is difficult to account for how we would use or understand them.

Here is another example: I recently attended a lecture on phonology that was almost incomprehensible to most of the non-phonologists in the audience because of the speaker's extensive use of unexplained technical jargon. Most of the unfamiliar terms were names or descriptions of certain (kinds of) sounds and of various parts of our anatomy. I felt as if I were in a Chinese Room listening to aural squiggles. My first reaction was Searlean: I wanted to know the meanings of these terms—I needed *semantics*. Upon calmer reflection, I realized that what I really needed was to *hear the sounds* or to *see pictures* of the anatomical structures. A computer would need, and could have, the same *additional input*—i.e., annotations (as in the Semantic Web). But such input, together with the original terms and links among them, would just be more internal syntax. No external semantic links would be needed.

In the Semantic Web, Web pages containing text and images are associated (or “annotated”, typically on their source pages) with other symbolic (i.e., *syntactic*) information that provides “semantic” information about the Web-page content. (For this reason, Ceusters 2005 cynically refers to the Semantic Web as the “Syntactic” Web!) Tim Berners-Lee appears to be of two minds concerning this. In 1999, he said:

a piece of information is really defined only by what it's related to, and how it's related. There really is little else to meaning. The structure is everything The brain has no knowledge until connections are made between neurons. All that we know . . . comes from the way our neurons are connected. (Berners-Lee & Fischetti 1999: 12; cf. pp. 184ff.)

This sounds very much like a purely syntactic, conceptual-role semantics. Yet, a few years later, he wrote this (perhaps under the influence of his co-authors):

The computer *doesn't truly “understand”* any of this information [provided by the Semantic Web], but it can now *manipulate the terms much more effectively*²² in ways that are useful and meaningful to the human user. (Berners-Lee et al. 2001; my emphasis.)

Of course, we are not told what “true understanding” is. Perhaps it is first-person understanding—understanding *for the computer* (or for the Web itself?)—as opposed to third-person understanding *for a “human user”*. The theory of syntactic

Laputians in Swift's *Gulliver's Travels*, who speak by using objects, not words (Rapaport 2000).

²²As Albert Goldfain (personal communication, 2006) put it: It can *understand with* the Semantic Web information.

semantics claims that first-person understanding is just this sort of syntactic manipulation, and the more ways there are to ‘manipulate the terms’, the more understanding there will be. Elsewhere, we have argued that a cognitive agent (including a cognitive computer) *can have* such first-person—or ‘intrinsic’—understanding: For example, a computer that can calculate greatest common divisors (GCDs) can be said to understand them if it has a sufficiently large network of information about GCDs and mathematics in general.²³

Searle’s frustration stems from his desire to understand such networks semantically. But this is a mistake; the only option open to him is to understand them syntactically. Granted, there aren’t enough samples of the Japanese networks for *me* to be able to make much sense out of them. But given enough networks and time, I should be able to.²⁴

2.2 The Library Room.

But where is the *experience*—that qualium, if you will—of understanding (cf. Galbraith & Rapaport 1995)? Where is that feeling of ‘Aha! *Now* I understand’? The short answer is that the understanding is ‘in’ the complex network. Now, on the face of it, that’s not very satisfactory. If I can detect no understanding from a few networks, why should it arise from a few more? The long answer is that the more ‘interlocking’ networks there are, and the more experience the person in the room has in manipulating them, the more understanding there will be. How might such understanding and experience of understanding arise?

To see how, consider a second Chinese Room: my understanding of Library of Congress catalog numbers: I have never studied, nor do I know, the rules of syntax or semantics for them, but neither do I need to in order to use them ‘fluently’ to find books. However, through frequent exposure to, and use of, them, I have come to learn, inductively, certain rules, both of syntax—for example:

LOC-catalog-number ::=

²³See Rapaport 1988b for discussion, and Goldfain 2004, 2006 for a preliminary computational implementation and theory.

²⁴Would I be able to do so uniquely or ‘correctly’? Or would the theoretical existence of an infinite number of models for any formal theory mean that the best I might be able to hope for is an understanding that would be unique (or ‘correct’) ‘up to isomorphism’? Or (following Quine 1960, 1969) might the best I could hope for be something like equivalent but inconsistent translation manuals? John McCarthy (personal communication, April 2006) thinks that I would eventually come to a unique or correct understanding. These are interesting and important issues, but beyond the scope of this essay. For my present purposes, it suffices that the understander be able to make *some* sense out of the networks, even if it is not the intended one, as long as the understanding is consistent with all the given data and modifiable (or ‘correctable’) in the light of further evidence. (Cf. Rapaport & Ehrlich 2000, Rapaport 2005a.)

letter₁ + numeral + ‘.’ + letter₂ + numeral + year-numeral

and of semantics—for example:

- [[letter₁]] = a category of certain books;
 - usually, [[letter₂]] = initial letter of author’s last name;
 - usually, [[year-numeral]] = year of publication;
- etc.

The more links I make with my knowledge of books and libraries, the more I know of the syntax and semantics (though it doesn’t necessarily help me to communicate any better); and the more I know of the syntax and semantics, the more I understand of what I’m doing. Searle-in-the-Library-of-Congress-room would also come to have such understanding. Why shouldn’t Searle-in-the-Chinese-room?

I hasten to admit that this requires some experience with the external world: I need to match LOC codes with actual books. So, it might be objected that I am not dealing only with a syntactic symbol system, but—inevitably—with such a system *and* with the real world. Objections to theories are best viewed as problems that the theory must deal with (Castañeda 1980, 1984; Rapaport 1982). My preferred way of dealing with this particular problem is, once more, the second thesis of syntactic semantics, namely, internalizing my experiences with the real world. Here is how it happens in the present case: I go to the library with an LOC number to find the book whose number it is; I find the book, see that the letter categorizes it along with other books on roughly the same topic, see that the letter is the initial letter of the author’s last name, and so on. But my visual perceptions of these things are internal—they are images in my mind, caused (to be sure) by external objects, but as much in my mind as are my mental images of the LOC numbers. My claim, then, is that the semantic relations between LOC numbers and their meanings are all (internal) syntactic relations among two sets of (mental) symbols.

But suppose you don’t agree with this argument. Then, for me to explain what the semantics of LOC numbers is, I need to describe the numbers and their meanings to you. But those descriptions of their meanings—of the real-world objects—are syntactic items that are themselves representatives of the real-world objects. So, once again, we are stuck with syntactic relations between two symbol systems (Smith 1982, Rapaport 1995).

2.3 The Helen Keller Room.

Has anyone ever *really* been in a Chinese Room for a sufficiently long time? The answer, I suggest, is ‘Yes’: Helen Keller has.²⁵

²⁵A similar observation has been made by Justin Leiber (1996, esp. p. 435):

The morning after my teacher [Anne Sullivan] came she . . . gave me a doll. . . . When I had played with it a little while, Miss Sullivan slowly spelled into my hand the word “d-o-l-l.” I was at once interested in this finger play and tried to imitate it. When I finally succeeded in making the letters correctly I was flushed with childish pleasure and pride. Running downstairs to my mother I held up my hand and made the letters for doll. *I did not know that I was spelling a word or even that words existed; I was simply making my fingers go in monkey-like imitation.* (Keller 1905: 35; my emphasis.)

At the beginning of this passage, one expects that the antecedently-played-with doll would be associated with the finger-spelled word ‘d-o-l-l’. But, as can be seen from Keller’s later claim of ignorance (“I did not know . . .”), her statement that she “made the letters *for doll*” (my italics) must be taken *de re*,²⁶ since, clearly, Keller did not know that she was “making . . . *letters*” (my italics) or that they were “for doll”.

The italicized last sentence of Keller’s quote is, I think, a significant passage. It is a wonderful description of pure syntax. Searle would be pleased. Sullivan, on the other hand, no doubt would have had reason to believe that Keller *did* know what she was doing. Sullivan plays native-Chinese speaker to Keller’s Searle-in-the-room.

The passage continues:

In the days that followed I learned to spell in this uncomprehending way a great many words, among them *pin, hat, cup* and a few verbs like *sit, stand* and *walk*. But my teacher had been with me several weeks before I understood that everything has a name. (Keller 1905: 35.)

[T]he suspicion that Keller did not live in the real world, could not mean what she said, and was a sort of symbol-crunching language machine . . . suggests a prejudice against the Turing test so extreme that it carries the day even when the Turing test passer has a human brain and body, and the passer does not pass as simply human but as a bright, witty, multilingual product of a most prestigious university, and professional writer about a variety of topics.

For readers unfamiliar with Helen Keller (1880–1968), she was blind and deaf because of a childhood illness (see below), yet graduated from Radcliffe College of Harvard University, wrote three autobiographies, and delivered many public lectures on women’s rights, pacifism, and helping the blind and deaf.

²⁶When I was very young, I heard what I took to be a single word that my parents always used when paraphrasing something; here, I’ll spell it ‘inotherwords’, but at the time I had no idea how to spell it. It took me several hearings before I understood that this was really the three-word phrase ‘in+other+words’. Similarly, from Keller’s point of view, her finger spellings were not letters+for+dolls; they were an unanalyzed ‘lettersfordolls’.

Here, ‘name’ should be understood in a broad sense, synonymous with ‘word’; for Keller and Sullivan, even the adjective ‘very’ is the “name” of a “thing” (Keller 1905: 261). Although sounding a bit odd to modern philosophical ears, this is not an entirely idiosyncratic usage. Tarski (1969) wrote: “When saying something of an object, one always uses a name of this object and not the object itself, even when dealing with linguistic objects.”

Again, these descriptions of Keller’s own experiences, given long after the fact, are best understood *de re*. She experienced external things, and she experienced meaningless finger manipulations, but she did not link them in an appropriate way. Such linking between experiences of a word (a finger spelling) and experiences of an external object would have yielded semantic understanding. According to the theory of syntactic semantics, semantic understanding would actually have come via Keller’s linking of *internal* representations of *both* of those external experiences. They would have played different roles: one the role of syntax, one that of semantics. One (the finger spellings) would not have been comprehended; the other (the physical objects) would have been familiar (they were, after all, part of the world she lived in every day). One would have been a name for a thing; the other, a thing named.

As we know, Keller succeeded remarkably well in her understanding of natural language. Here is how she described her experience immediately after the well-house²⁷ event during which one of her hands was immersed in water while Sullivan finger-spelled ‘w-a-t-e-r’ in her other hand, and Keller allegedly first realized that the finger-spellings had meaning (we’ll examine this event in detail later):

As we returned to the house every object which I touched seemed to quiver with life. That was because I saw everything with the strange, new sight that had come to me. (Keller 1905: 36.)

And this, I claim, would eventually be the experience of Searle-in-the-room, who would then have *semantic* methods for doing things in addition to purely *syntactic* ones (just as logicians have both syntactic and semantic methods of proof). The semantic methods, however, are strictly internal: They are not correspondences among words and things, but correspondences among *internal nodes for* words and things. As such, they are syntactic.

Jim Swan (1994) has pointed out how important Keller’s *hand* was to her ability to communicate.²⁸ At the well house, both the name and the object were communicated via the *same* sense modality: touching her hand. He also points

²⁷Keller referred to the location as the “well”-house, whereas Annie Sullivan, her teacher, referred to it as the “pump”-house. Keller’s house is now a museum; their website (helenkellerbirthplace.org) also calls it a “pump”. However, I will use Keller’s term in this essay.

²⁸On hands and cognition, cf. Papineau’s (1998) review of Wilson 1998.

out how she had to learn about the visual aspects of the world *as expressed via language* (rather than perceived). All of this lends credence to the view that Keller's understanding of language and the world was an internal understanding, with all input encoded into a single syntactic system. Mental concepts ("nodes", in SNePS terminology) for both words and objects were built on the basis of tactile (and olfactory) sensation. One of the reasons the well-house episode was significant was that it was the event that enabled Keller to distinguish some of her internal nodes from others, categorizing some as representing the world and others as names of the former. For Keller, initially, language and the non-linguistic part of the world were indistinguishable.

Swan discusses, from a psychoanalytical point of view, the difficulty for Keller of distinguishing between self and other, between her words and those of others.²⁹ Before the well-house episode, she could use signs, but had difficulties with certain ones, in particular, with those for container vs. contained ('mug' vs. either 'milk' or 'water'; see §§3.2.6, 3.2.8–3.2.10, and 3.3.1, below). Perhaps, *before* the well-house episode, she could not distinguish words from (other) objects: Words *were* objects, part of the holistic fabric of her world. *After* the well-house episode, she could distinguish between two kinds of objects in the world: (non-linguistic) objects and words for them. That placed a syntactic-semantic structure on her mental network. And it resulted, as we know, in the blossoming of her understanding. Searle-in-the-room could do no worse.

Gilbert Harman (1987: 57) has said that

a language, properly so called, is a symbol system that is used both for communication and thought. If one cannot think in a language, one has not yet mastered it. A symbol system used only for communication, like Morse code, is not a language.

Before the well house, Keller used symbols to communicate, but not to think. Her pre-well-house signs were like non-understood Morse code. (Do you understand what "... . -.. -.. - - -" means if you know no Morse code? But if I finger-spelled that into your hand each time I saw you, and forced you to do the same to me, would you get the hang of doing it? Would you *then* understand it?)

Unless the symbols are part of a larger network, they have no (or very little) meaning. To that extent, perhaps Searle has a point. But the more they *are* used for thinking, the more language-like they are. And they *have* to be part of a larger network—partitioned into syntactic and semantic regions—else they could not be

²⁹Keller had been accused of plagiarism, when, in fact, it is possible that she had merely made a grievous use-mention confusion, viz., not having learned how to use quotation marks; cf. Keller 1905.

used to communicate. They have meaning if and only if (and to the extent that) they are part of such a larger, partitioned network. Searle denies the “if” part of this—i.e., that being part of such a network confers meaning—but Helen Keller, I suggest, was a living counterexample. (Cf. my remarks in §2.1, above, about understanding GCD and about the mathematical codes we seem to have in our minds.)

3 Names for Things: From “Monkey-Like Imitation” to Natural-Language Understanding.

3.1 A Puzzle.

I have suggested that the case of Helen Keller offers a real-life Chinese Room situation, and I have given some reasons why the epiphenal well-house episode—paradigmatic of the relationship between a syntactic domain and a semantic domain, with Keller simultaneously having one hand immersed in syntax and the other in semantics—was so significant for her.

But, really, why should it have been? By Keller’s and Sullivan’s own testimony, Keller seemed able to use (finger-spelled) words for things, as well as (self-invented) signs and gestures for things, *before* the well house. So what made the well house so significant?

3.2 What Did Helen Keller Understand, and When Did She Understand It?

It is not easy to determine the chronology of Keller’s language learning. There are at least two distinct, if not independent, first-person accounts: (1) the student’s: Keller’s autobiography (Keller 1905), written, of course, long after the events, and (2) the teacher’s: Sullivan’s contemporaneous letters. The latter are probably to be trusted more.³⁰ As Keller herself says, “When I try to classify my earliest impressions, I find that fact and fancy look alike across the years that link the past with the present. The woman paints the child’s experiences in her own fantasy”³¹ (Keller 1905: 23; hereafter, all page citations are to Keller 1905 unless otherwise indicated). Even though Sullivan’s letters are “incomplete” as scientific “records” (Macy, in Keller 1905: 224, 239; cf. p. 241), they are the closest thing available. Together, the two sources provide a reasonable—if tantalizing—picture. For Keller’s earliest years, however, we must rely on her own report, since Sullivan was not present.

(There are also Keller’s somewhat more contemporaneous letters, but these—while intrinsically interesting, and exhibiting (especially in the early ones) her gradual mastery of language—do not contain much information on *how* she learned language. There are also Sullivan’s speeches and reports. Although they contain

³⁰Trusting a third-person viewpoint over a first-person viewpoint is consistent with trusting the native Chinese speaker’s viewpoint over Searle-in-the-room’s viewpoint. Cf. Rapaport 2000.

³¹This has an overtone of holistic reinterpretation; cf. Rapaport 1995, §2.6.2: We understand the present in terms of all that has gone before, and the past in terms of all that has come after.

some useful information and some valuable insights—especially into the nature of teaching—they, like Keller’s autobiography, were written *ex post facto*; cf. Macy, in Keller 1905: 278.)

3.2.1 Before the Illness

Keller was born, healthy, on 27 June 1880. At 6 months, she could speak a few words or phrases, e.g., ‘How d’ye’, ‘tea’, and—significantly—‘wah-wah’ (‘water’) (p. 25). But did she understand them?

By way of comparison with Keller’s linguistic abilities at 6 months, my son at the same age could only say ‘Da-da-da’, and said it for the first time when *not* in my presence. My granddaughter at the same age did not even say that much. The difficulty of interpreting what infants ‘say’, hence of deciding what they understand, is illustrated nicely in a ‘For Better or for Worse’ cartoon in which a girl (Elizabeth) holds her baby sister (April), who is looking down (presumably at their dog) and saying ‘Goggg-gog Gogg! Go-Go Gogg Gogog’. Elizabeth says, ‘Mom! Listen! April’s trying to say *doggie!*’ Their mother explains, ‘She’s just making sounds, honey. She doesn’t know what she’s saying.’ Then April, in the presence of her bottle, says, ‘Babab. Ba Ba. Ba-ba. Bababa. Ba-ba. Ba!’, and Elizabeth exclaims, ‘She’s saying Ba-Ba for bottle!’ Again, their mother rationally explains, ‘She doesn’t know the word ‘bottle’, Elizabeth. It’s all just noise to her.’ April then looks at her mother and utters, ‘Mum-mum mamamama ma-ma’; the astonished mother calls her friend Connie on the phone to announce, ‘It’s true, Connie! I heard her! She said Ma-Ma!’ Similarly, my granddaughter at age 1;2,21³²—still without language but babbling with English intonation—pointed to a swing and said something that sounded to us like ‘ist?’; *we* interpreted her as saying ‘this’. Linguistic communication is at least 50% the *hearer’s* interpretation.

3.2.2 The Illness

In February 1882, at age 1;7, Keller contracted the illness that left her deaf and blind, and, like many deaf children, she did not learn to speak. Nevertheless, she could make sounds—again, significantly, the sound ‘wah-wah’ for ‘water’, which ‘I ceased making . . . only when I learned to spell the word’ (p. 25).

For another anecdotal comparison, at about this age (1;6,23–1;7,5), my son could say (impolitely!) to our neighbor’s 2-year-old daughter Marielle, ‘Shut up, Ma’iel’; he used and seemed to understand ‘OK’ and ‘good’; he could say ‘See ya’ to his older brother leaving the house; he used ‘zipper’ and ‘wait’ correctly; etc. I noted in my diary at the time that each day he had many new

³²I.e., at age 1 year, 2 months, 21 days.

words and new abilities to comprehend. It occurred to me that he was not merely learning to speak; rather, he spoke a “foreign” language that his mother and I could interpret to some extent. Words meant different things for him than for us. For example, ‘Bambi’ sometimes meant ‘videotape’; ‘water’ sometimes meant ‘liquid’. But these seemed neither to be overgeneralizations nor metaphors; such characterizations, in any case, are *de re*, third-person descriptions. From my son’s (*de dicto*) point of view, those words *didn’t* mean what *we* think they do.

3.2.3 After the Illness

After her recovery, Keller could communicate via touch, via “crude signs”—

A shake of the head meant “no” and a nod, “yes”, a pull meant “come” and a push meant “go” (p. 27)

—and via (other) imitative motions, including some rather complex ones, e.g.:

when I wanted to go egg-hunting, . . . I would double my hands and put them on the ground, which meant something round in the grass, and Martha [a childhood friend] always understood. (p. 28.)

She familiarized herself with the outdoors, guided by her sense of smell. It is perhaps worth noting that this continued well after Sullivan’s arrival: They often studied outside (p. 43; cf. Sullivan’s comments on the significance for Keller of the sense of smell, pp. 293ff). She also, of course, had a sense of taste, learning thereby that the ocean was salty—a bit of commonsense knowledge that she lacked because no one thought to tell her such an obvious thing (p. 230)! She also had a “sense” of vibration, being able to sense when a door closed (p. 27). I am not sure whether to count this as part of her sense of touch, or as a remnant of a sense of hearing, which is, after all, a sensitivity to vibrations (cf. p. 208 on her ability to sense music).

3.2.4 Age 5

By the age of 5, she could perform rather complex tasks, such as folding and separating clean clothes (p. 27), and she knew that her father did something mysterious by holding a newspaper in front of his eyes. She imitated this, but it did not illuminate the mystery (p. 30).³³ Similarly, she knew that others communicated, not with signs, but by moving their lips; imitation of this, too, was not successful (pp. 27–28).

³³Cf. what I have called the “miracle of reading”: “When we read, we seemingly just stare at a bunch of arcane marks on paper, yet we thereby magically come to know of events elsewhere in (or out of!) space and time” (Rapaport 2003a, §6; a typo in the original is here corrected). Clifton

3.2.5 Age 6

Fortunately, others understood her signs (p. 28). When she was 6, she tried to teach her dog some of these (with no success, of course; p. 29), though Sullivan tells a similar story (p. 253) about Keller trying to teach *finger spelling* to her dog at about the same age or a bit later (20 March 1887, to be exact—*after* Keller had begun to learn words but *before* the well house) (p. 253). She certainly, at about this time, had a desire to express herself (p. 32).

3.2.6 Sullivan’s Arrival: Finger-Spelling in a Chinese Room

On 3 March 1887, Sullivan arrived at Keller’s home to become her teacher; Keller was now about 6;8. Almost immediately upon her arrival, Sullivan and Keller began to communicate with each other using signs and gestures (p. 245). The next day, Sullivan began teaching Keller finger spelling, presenting her with an object or action and finger-spelling its name: ‘doll’, ‘pin’, ‘hat’, ‘cup’, ‘sit’, ‘stand’, and ‘walk’ are the words Keller remembered. Sullivan cites ‘doll’, ‘cake’, and (sewing) ‘card’. ‘Cup’ is of some interest, since ‘mug’ was to give Keller a notorious difficulty a few weeks later.

To what extent did Keller understand these words? As we saw, she herself considered this to have been “monkey-like imitation” (p. 35): Finger spelling was an activity to be performed upon presentation of certain objects. It was a ritual, with a syntactic structure: There were right and wrong ways to perform it. But Keller did this “in . . . [an] uncomprehending way” (p. 35) and did not yet understand “that everything has a name” (p. 35). It certainly *seems* that she was in a Chinese Room.

Was Keller really so uncomprehending at this stage? Recall that she had already developed her own system of signs and gestures for communicating her needs and wants.³⁴ Surely, this is evidence of a semantic correspondence.

I suppose that it is remotely possible that even Keller’s early self-invented signs were ritual movements performed uncomprehendingly in certain circumstances, yet rituals that just happened to convey information to other people. In Robert Sheckley’s story “Ritual” (1954), creatures living on a remote planet perform a

Fadiman once observed that:

[W]hen I opened and read the first page of a book for the first time, I felt that this was remarkable: that I could learn something very quickly that I could not have learned any other way . . . [I] grew bug-eyed over the miracle of language . . . [viz.,] decoding the black squiggles on white paper. (Quoted in Severo 1999.)

Hofstadter 2001: 525 makes similar observations.

³⁴Leiber 1996 also notes that Keller had linguistic knowledge and abilities both before and immediately after her illness.

series of ritual “welcoming the gods” dances as a religious ceremony. The dance happens to consist of the preparations for the arrival of a spaceship. When a spaceship finally does arrive after centuries without a landing, the villagers perform their “dance”, which just happens to facilitate the spaceship landing.

But I doubt that Keller’s signs were such rituals. Had *all* of Keller’s gestures been such conveniently coincidental (“extensional”) rituals, she would not have been able to do the complex tasks she did, or to satisfy her needs, or to have the appropriate background knowledge that, eventually, was the basis for her language learning.

All that Sullivan was doing can be seen as offering Keller a new system for accomplishing her communicational goals. It is, of course, possible that Keller did not realize this, so that, for her, her *own* gestures for an object *did* constitute a semantic correspondence while Sullivan’s finger spellings did not. However, that Keller *was* able to associate finger spellings with objects and actions surely indicates that she had the means to promote these to *semantic* correspondences.

There is, in fact, evidence that she did so: The day that Sullivan arrived, she taught Keller ‘cake’, and the next day she taught her ‘card’. Keller . . .

...made the ‘c-a,” then stopped and thought, and making the sign for eating and pointing downward she pushed me [Sullivan] toward the door, meaning that I must go downstairs for some cake. The two letters ‘c-a,” you see, had reminded her of Friday’s ‘lesson”—not that she had any idea that *cake* was the name of the thing, but it was simply a matter of association, I suppose. (p. 246.)

I would argue that Keller *did* have the idea that ‘cake’ “was the name of the thing”—but that she had that idea *de re*, not *de dicto*: She did not yet have the *concept* of names for things. She could certainly associate words (i.e., finger spellings) with objects:

Then I [Sullivan] spelled “d-o-l-l” and began to hunt for it. She [Keller] follows with her hands every motion you make, and she knew that I was looking for the doll. She pointed down, meaning that the doll was downstairs. . . . [S]he ran downstairs and brought the doll . . . (pp. 246–247),

although not without a reward of some cake.

As Keller built up a vocabulary of finger-spelled words and made mental links between (her internal representations of) these and (her internal representations of) objects and actions, she was building a “semantic network” of associated representations that she could and did use in a language-like way. Searle would

argue that she did not *understand* language. Perhaps. I'd prefer to say that she did not understand language in a *de dicto* way—she *did* understand it *de re*, in the sense that she was using it, but did not realize that she was using it or how it worked. That is, perhaps she understood language, but did not understand that (or how) she understood it.

She was, in fact, at the same stage of language development as a normal child would have been at a much earlier age. Are we prepared to say that normal children at this stage do not understand language? Perhaps. But eventually they *do*, and eventually Keller did. Why not Searle-in-the-room or a computer? What is the crucial step (or steps) that must be taken to move from *this* level of understanding (or, if you prefer, from this level of *not* understanding) to the level that we adult speakers of language are at? We'll return to this in §4.

3.2.7 Ages 6;8,12–6;8,14

By 11 March 1887, Sullivan says that “Keller knows several words now, but has no idea how to use them, or that everything has a name” (p. 251). Yet two days later, Keller can associate words with objects: “when I give her the objects, the names of which she has learned, she spells them unhesitatingly” (p. 251).

3.2.8 Age 6;8,21

Around 20 March 1887, Keller reports that she was confused by ‘mug’ and ‘water’:
Keller and Sullivan

... had a tussle over the words “m-u-g” and “w-a-t-e-r.” Miss Sullivan had tried to impress upon me that “m-u-g” is *mug* and that “w-a-t-e-r” is *water*, but I persisted in confounding the two. (p. 36.)

Apparently, she confused the container with the contained; perhaps this was because they always appeared together. Although she may have had a *mug* by itself, perhaps the *water* was always *in* the mug.

Here is Sullivan's account (concerning a different liquid):

Helen has learned several nouns this week. “M-u-g” and “m-i-l-k” have given her more trouble than other words. When she spells “milk,” she points to the mug, and when she spells “mug,” she makes the sign for pouring or drinking, which shows that she has confused the words. She has no idea yet that everything has a name. (Sullivan, 20 March 1887, pp. 252–253.)

By ‘learning’ (as in ‘Keller has learned several nouns this week’), Sullivan must mean the ability to spell, to make the finger movements. ‘Mug’ and ‘milk’ (or ‘water’?) give Keller trouble, *not* in terms of the movements, but in terms of how to use them (what they refer to, or name). But that assumes that Keller knows that they have a *use*, which is plausible, as we’ve seen, though not altogether clear. Moreover, Sullivan’s interpretation of Keller’s actions is not the only one: Pointing to the mug *could* also be pointing to the milk (or water) *in* the mug, and making signs for pouring or drinking *could* refer to what one does with a *container* as well as with its *contents*. Keller’s own version suggests that she was not making *any* distinctions at all, rather than merely confusing the mug and the liquid.

3.2.9 Age 6;9,9

Sullivan’s second version of the confusion supports my interpretation that Keller was aware only of *events* considered as unanalyzed wholes:

... “mug” and “milk” had given Helen more trouble than all the rest. She confused the nouns with the verb “drink.” She didn’t know the word for “drink,” but went through the pantomime for drinking whenever she spelled “mug” or “milk”. (Sullivan, 5 April 1887, p. 256.)

I think it is significant that Sullivan reports the confusion as between ‘mug’ and ‘milk’, where Keller reports it as between ‘mug’ and ‘water’.³⁵ First, and most importantly (if only for Freudian reasons), Keller’s one remaining *spoken* word was, you will recall, ‘water’ (‘wah-wah’). Second, if Sullivan’s report is the one to be trusted, besides the semantic-domain confusion between container and contained, there might also have been a syntactic-domain confusion between two words beginning with ‘m’: Recall the earlier ‘confusion’ between ‘ca[ke]’ and ‘ca[rd]’.

3.2.10 Just Before the Well-House

There were a few days to go before the visit to the well house. What did Keller learn in those days between (a) her confusing the word for a mug with the word for its liquid contents and (b) her later epiphany? By 20 March, according to Sullivan, Keller knew 12 word-object combinations (p. 255) yet instinctively used her own signs—not finger-spelled words—to *communicate*. By 1 April, Sullivan

³⁵In an earlier autobiography, Keller also called this a ‘mug’/‘milk’ confusion (p. 364). And in Sullivan’s description of the well-house episode (see §3.3.1, below), she describes “w-a-t-e-r” as a “new word” for Keller (p. 257).

reports, Keller's vocabulary had increased to 25 nouns and 4 verbs³⁶—including, significantly, 'mug', 'milk', and 'water'. Yet, two days later, Sullivan says that Keller "has no idea what the spelling means" (p. 256). I take it that, from Sullivan's point of view, Keller's "knowledge" of these words was at least associative and probably even communicative, yet not "conscious". But not "conscious" in what sense? Keller apparently could ask for the finger spellings that corresponded to certain objects (the ones marked 'x' in note 36). What more could Sullivan want at this stage?

Searle, no doubt, would say that for real natural-language understanding, a lot more is wanted. I'd have to agree: Keller could not yet have passed a Turing test. So although imagining what Keller was like at this stage may give us an insight as to what Searle-in-the-room is like, there is a large gap between the two. Searle-in-the-room, remember, passes the Turing test.

Perhaps what Keller "knew" at this stage was an association of these words with certain complex, unanalyzed events, and what she learned at the well house was that the events have parts, each of which is associated with a word. If so, then what she learned was as much about the semantic domain as it was about the association between the two domains. Of course, she also presumably learned then that the *words* did not refer to complex events but only to *parts* of them. So she learned something about the syntactic domain, too.

3.3 Helen Keller and the "Miracle Worker" at the Well House.

3.3.1 Epiphany.

The magical day was 5 April 1887. Sullivan, having failed to clarify the difference between 'mug' and 'milk', took Keller for a walk to the well house.

... This morning, while she was washing, she wanted to know the name for "water." ... I spelled "w-a-t-e-r" ... [I]t occurred to me that with the help of this new word I might succeed in straightening out the "mug-milk" difficulty. We went out to the pump-house, and I made Helen hold her mug under the spout while I pumped. As the cold water gushed forth, filling the mug, I spelled "w-a-t-e-r" in Helen's free hand. The word coming so close upon the sensation of cold water rushing over her [other] hand seemed to startle her. She dropped the mug and stood as one transfixed. A new light came into her face. She spelled "water" several times. Then she dropped on the ground

³⁶"Doll, mug, pin, key, dog, hat, cup, box, water, milk, candy, eye (x), finger (x), toe (x), head (x), cake, baby, mother, sit, stand, walk. ... knife, fork, spoon, saucer, tea, paper, bed, and ... run" (p. 256). "Those with a cross after them are words she asked for herself" (p. 256).

and asked for its name and pointed to the pump and the trellis, and suddenly turning round she asked for my name. I spelled “Teacher.” Just then the nurse brought Helen’s little sister into the pump-house, and Helen spelled “baby” and pointed to the nurse. All the way back to the house she was highly excited, and learned the name of every object she touched, so that in a few hours she had added thirty new words to her vocabulary. . . .

. . . Helen got up this morning like a radiant fairy. She has flitted from object to object, asking the name of everything and kissing me for very gladness. Last night when I got in bed, she stole into my arms of her own accord and kissed me for the first time, and I thought my heart would burst, so full was it of joy. (Sullivan, 5 April 1887, pp. 256–257.)

A few observations on this passage and on the well-house episode are in order.

1. On “wanting to know the name for ‘water’ ”: Clearly, Keller wanted to know the name for water (the *stuff*), not for ‘water’ (the *word*); she did not want to know the name for a name. However, Sullivan is not to be blamed for this particular use-mention confusion! On the other hand, hasn’t Sullivan repeatedly told us that Keller did not know that things have names? Then why does she report Keller as asking for the *name* of water? Perhaps this needs to be taken *de re*: Note that it’s quite possible that what Keller wanted to know was the appropriate finger spelling for *washing*!
2. Sullivan’s comment about “straightening out the ‘mug-milk’ difficulty” can be interpreted as supporting my suggestion that the mug-milk confusion was one of container vs. contained or of unanalyzed events.
3. Note that, at the well house, there was little chance to “confound” two objects—there was a direct and simultaneous association of word with object. Although the mug in Keller’s hand *might* have caused some interference, Keller’s own account indicates that it did not:

We walked down the path to the well-house Some one was drawing water and my teacher placed my hand under the spout. As the cool stream gushed over one hand she spelled into the other the word *water*, first slowly, then rapidly. I stood still, my whole attention fixed upon the motions of her fingers. Suddenly I felt a misty consciousness as of something forgotten—a thrill of returning thought; and somehow the mystery of

language was revealed to me. I knew then that ‘w-a-t-e-r’ meant the wonderful cool something that was flowing over my hand. . . .

. . . As we returned to the house every object which I touched seemed to quiver with life. That was because I saw everything with the strange, new sight that had come to me. (p. 36.)

Moreover, if, indeed, it was ‘*milk*’—not ‘*water*’—that Keller had been confusing with ‘mug’, then the well-house experience was a controlled experiment, filling the mug with *water* instead of milk.

4. The finger-spelled word ‘w-a-t-e-r’ meant ‘the wonderful cool something that was flowing over my hand’: ‘W-a-t-e-r’ was antecedently meaningless; ‘the wonderful cool something . . .’ was antecedently understood. The semantic relation is asymmetric (Rapaport 1995); here we have the intensional, asymmetric equivalence of a definition.
5. Note that Keller did *not* say that ‘water’ meant H₂O: *Twin* Keller (living on Putnam’s (1975) Twin Earth, where ‘water’ refers, not to H₂O, but to XYZ) would have had the *same* experience, and ‘water’ would have meant exactly the same thing for her (modulo the essential indexical ‘my’), viz., ‘the wonderful cool something that was flowing over my hand’.
6. Keller’s post-well-house experiences of seeing “everything with the strange, new sight” should be the eventual experience of Searle-in-the-room, who would then have *semantic* methods for doing things in addition to purely syntactic ones (again, cf. syntactic vs. semantic proofs in logic and math; see §2.3). Crucial to promoting semantics-as-correspondence to semantics-as-meaning—semantics-as-understanding—is that the semantic domain must be antecedently understood. This, as we shall see shortly, was crucial for Keller’s progress.

3.3.2 Aftereffects.

Five days later, Sullivan reports Keller replacing her own signs by the corresponding finger-spelled words as soon as she learns them (p. 257). Clearly, Keller had realized the advantages of this new, more efficient and expressive code for communication. Equally crucially, as Sullivan observes (p. 258), Keller *understood* what the finger-spelled words referred to *before* she was able to ‘utter’ them: “The idea always precedes the word” (Sullivan, 8 May 1887, p. 260). As Sullivan noted later (pp. 291ff), Keller had her *own* signs for things before she had *words* for them, still using her signs when she had not yet learned the words

(p. 260), so she was using two codes. She had several ways to communicate her ideas, preferring one (words), but using whatever was at hand (so to speak).³⁷

Two other observations that Sullivan made are worth mentioning at this point. First, it was important for Keller to *generate* language, not merely to *understand* it, in order to help build her vocabulary (Sullivan, 16 May 1887, pp. 262ff); interactive conversation is crucial (cf. Rapaport 2000, §8; Rapaport 2003a).

Second,

Language grows out of life, out of its needs and experiences. . . . *Language* and *knowledge* are indissolubly connected; they are interdependent. Good work in language presupposes and depends on a real knowledge of things. As soon as Helen grasped the idea that everything had a *name*, and that by means of the manual alphabet these names could be transmitted from one to another, I proceeded to awaken her further interest in the *objects* whose names she learned to spell with such evident joy. *I never taught language for the PURPOSE of teaching it*; but invariably used language as a medium for the communication of *thought*; thus the learning of language was *coincident* with the acquisition of knowledge. In order to use language intelligently, one must have something to talk *about*, and having something to talk about is the result of having had experiences; no amount of language training will enable our little children to use language with ease and fluency unless they have something clearly in their minds which they wish to communicate, or unless we succeed in awakening in them a desire to know what is in the minds of others. (Sullivan, p. 317.)

Jerome Bruner has observed much the same thing:

So at the end of this first round of examining the simplest form of request—asking for objects—we are forced to a tentative conclusion. Language acquisition appears to be a by-product (and a vehicle) of culture transmission. Children learn to use a language initially (or its prelinguistic precursors) to get what they want, to play games, to stay connected with those on whom they are dependent. In so doing, they find the constraints that prevail in the culture around them embodied in their parents' restrictions and conventions. The engine that drives the enterprise is not language acquisition per se, but the need to get on with the demands of the culture. . . . Children begin to use language

³⁷All puns intended!

... not because they have a language-using capacity, but because they need to get things done by its use. Parents assist them in a like spirit: they want to help them become “civilized” human beings, not just speakers of the language. (Bruner 1983: 103–104.)

This is an insight that—beyond its evident importance for education *in general*—is of importance for *computational* natural-language understanding systems, too. It is not far from some of the insights of Hubert L. Dreyfus (1992). Whether it is something that *cannot* be accomplished with computers remains, however, an empirical and open question, suggesting a clear direction for research. Knowledge (especially “prior” or “background” knowledge) is also important for figuring out a meaning for an unfamiliar word “from context”. Indeed, the “context” from which this can be done must include not only the unfamiliar word’s *textual* context (i.e., the surrounding words, or “co-text”) but also the reader’s prior knowledge (Rapaport 2003b).

Keller’s language learning proceeded apace after the well house. Two months later, she wrote her first letter to a friend (p. 123). Her vocabulary learning was cyclic and recursive—each new encounter with a word serving to clarify and enhance what she already knew (p. 40).

Words for abstract concepts (e.g., ‘love’, ‘think’)—concepts that could not be “shown”, hence for which there was nothing apparent to associate them with—were harder for her to learn, but not impossible (for the details, see pp. 40f, 300). In April 1887, she learned prepositions by direct experience—standing *on* a chair or *in* her wardrobe (Sullivan’s account, p. 279). Keller’s own account of learning sentence structure is reminiscent of Russellian propositions: She would paste pieces of paper with words written on them “in raised letters” (p. 279) onto the things they named: She would put her doll on the bed, the doll labeled ‘doll’, the bed labeled ‘bed’, with labels for ‘is’ and ‘on’ placed near the doll, on the bed; or she would put the label ‘girl’ on herself, the labels for ‘is’, ‘in’, and ‘wardrobe’ on the wardrobe, and then she would stand in the wardrobe, thus labeled.

Over a year later, by which time her language was of Turing-test quality, she would, nonetheless, use some not-yet-understood words in “parrotlike” fashion (Macy, p. 134) until she learned how to use them properly (until she learned their meaning?). These included “words of sound and vision which express ideas outside of her experience” (Macy, pp. 134–135). I have argued that we do the same, with words like ‘pregnant’ used by a male (cf. Rapaport 1988b, 2003a). Evidently, though, much more of Keller’s knowledge is knowledge by description than is ours. (Of the paintings in an art gallery, she said, “I have at least the satisfaction of seeing them through the eyes of my friends”—p. 200.)

3.4 The Puzzle of the Well House.

But what really happened at the well house? The well-house association of ‘water’ with water was not different in kind from previous word-object associations that Keller had made and had used for communication. Sullivan was not trying to teach Keller something new; she was merely trying to reinforce something she had more or less successfully taught her before. Various incidental experiences—Keller’s mug/water-or-milk confusion, her memory of the spoken word ‘wah-wah’, and the perhaps unique “co-activation” of word and object (cf. Mayes 1991: 111)—no doubt contributed to making the well-house experience the significant event it was. But what exactly *did* she learn, and why was it so significant?

3.4.1 What Did Keller Learn?

Keller learned something she had not been taught. In her own and Sullivan’s words, she learned that things have “names”. But not just that, for merely knowing that ‘w-a-t-e-r’ or ‘d-o-l-l’ were the appropriate finger spellings to perform when in the presence of water or a doll—or, significantly, when *not* in the presence of water or a doll, but desiring one or (more importantly; see §4, below) merely wishing to convey one’s thoughts about water or a doll—could be described as knowing that those things had names.³⁸

What Keller learned was that some things in the world (viz., finger spellings) were names of other things in the world. More precisely, she learned the concept of a name, thereby learning a metalinguistic fact:³⁹ Her mental world became more than an unstructured associative network of concepts; it developed a syntactic-semantic structure, by means of which some of the entities in it (her internal representations of words) “named” others (her internal representations of objects, events, ideas, etc.).

3.4.2 A SNePS Analysis of What Keller Learned: Preliminaries

It may prove helpful to use SNePS to understand Keller’s accomplishment. To do this, however, we will first replace the `lex` arcs of Figures 1 and 2 by a different mechanism for relating words to things, namely, a “name-object” case frame.⁴⁰

³⁸Of course, ‘name’ (or ‘word’) might be overly simplistic. A simple (to us) finger-spelled “name” might be interpreted as a full sentence: Possibly, ‘d-o-l-l’ means ‘Please give me my doll.’ Cf. ‘Please machine give cash’ as the meaning of pushing a button on a cash machine; see §4.3.

³⁹David Wilkins pointed out the metalinguistic nature of the well-house episode to me.

⁴⁰In other SNePS writings, this has been called an `expression-expressed` case frame (Neal & Shapiro 1987, Shapiro et al. 1996) or a `word-object` case frame, but I will use ‘name-object’ instead of ‘word-object’ for consistency with Keller’s and Sullivan’s terminology.

Thus, the proposition that humans are mammals will now be represented as in Figure 3, instead of as in Figure 1.

In SNePSLOG, we can define a predicate `ISA-Thing-Named`:

```
define-frame ISA-Thing-Named (nil object name)
```

where

```
[[ISA-Thing-Named( o,n )]] = o is the concept associated with  
or expressed by) the lexical entry  
(i.e., the word, or ‘name’) n.
```

(Cf. the semantics of the function symbol `thing-called`, in §1.4, above.) The network of Figure 3 can then be constructed as follows:

```
ISA-thing-named(b1,humans).  
ISA-thing-named(b2,mammals).  
AKO(b1,b2).
```

(I.e., roughly—but only very roughly! (see below)—`b1` is a thing named ‘humans’; `b2` is a thing named ‘mammals’; and `b1` is a kind of `b2`—i.e., humans is a kind of mammals, or—more idiomatically— humans are mammals.)

In Figure 1, nodes `m1` and `m2` are functional terms representing what we have called “structured individuals” in previous SNePS writings. They have what philosophers might call “essential” features, namely, their lexical expression. In Figure 3, nodes `b1` and `b2` are base nodes (i.e., nodes with no arcs emanating from them), which are like Skolem constants, representing what philosophers might call “bare particulars”,⁴¹ with *no* essential features, only “accidental” features asserted about them, namely, their names—i.e., their lexical expression.⁴²

In Figure 1, Cassie can *use* the words ‘humans’ and ‘mammals’ but has no beliefs *about* them. In Figure 3, she can *both* use the words *and* has beliefs about them. For example, node `m4` represents a belief about something, viz., `[[b1]]` (which is, in fact, the class of humans) and a name (or word) for it, viz., ‘humans’.

However, just as the `lex-arc` representation is not expressible by Cassie, `[[m4]]` is a belief that is also not expressible easily, if at all. The lesson of the discussion that follows is that a cognitive agent can have names for things without realizing or

The replacement of `lex-arc`s with a `name-object` case frame is not essential to my point, but makes the exposition clearer.

⁴¹Sellars 1963: 282; Allaire 1963, 1965; Chappell 1964; Baker 1967.

⁴²Instead of using “base-node” labels `b1`, `b2`, etc. we could have used mnemonic (for *us*) English words, like `humans1` and `mammals1`, as is often done in AI. For present purposes, as well as for McDermott-1981-like reasons, these would be needlessly confusing.

being able to say that they are names. The latter is a greater cognitive achievement than the former. This is what “understanding that everything has a name” really means.

Only *m6* is easily expressible; to express it, Cassie uses the words at the heads of *m4*'s and *m5*'s name arcs. The SNePS Natural Language Processing System's algorithm for expressing a node such as *m6* that has a `subclass-superclass` case frame is to produce an expression for the node at the head of the `subclass` arc (*b1*), followed by the word ‘are’, followed by an expression for the node at the head of the `superclass` arc (*b2*). This (recursive) algorithm for expressing a node like *b1* (or *b2*) that is at the head of an `object` arc emanating from an asserted node that has a name arc emanating from *it* is to produce the lexeme at the head of the name arc that emanates from the asserted node (*m4*!, or *m5*!) that has the `object` arc to *b1* (or *b2*). Thus, *m6* would be expressed as ‘Humans are mammals’. (See Shapiro 1982 for details.)

But to express *m4* or *m5*, she would also have to use those words to express *b1* or *b2*—so how would she express the nodes at the heads of *m4*'s or *m5*'s name arcs? One (bad) possibility for expressing *m4* is to produce an expression for *b1*, as before, followed by the word ‘is’, followed by the word at the head of the name arc, viz., ‘humans’. But this leads to the awkward and ambiguous ‘Humans is humans’.

Another (poor) possibility is to express the base node *b1* with a demonstrative word like ‘this’, then follow this with the word ‘is’, followed by the word ‘humans’. However, if we apply this to Figure 4's node *m8*, we wind up saying “this is wet”, which is ambiguous between “this thing has the property of being wet” and “this thing is called ‘wet’”.⁴³ Actually, it would be better to express *m8* with something like “this is wetness”, but it is difficult to generate such “philosophical” nouns without further network information to signal to the natural-language generator when (and how) to use them.

Yet another possibility is to express Figure 3's *m4* by saying “something that is a subclass of mammals is humans”, i.e., to describe *b1* in terms of the

⁴³Fig. 4 can be created, using SNePSLOG, by defining:

```
define-frame Is (nil object property)
```

where

```
[[Is (o, p)]] = object o has property p
```

and then asserting:

```
ISA-thing-named(b3, water).
ISA-thing-named(b4, wet).
Is(b3, b4).
```

assertional information about it in the rest of the network (“something that is a subclass of mammals”), followed by the word ‘is’, followed by the word ‘humans’: “Something that is a subclass of mammals is humans.” This is better, but still suffers from an ambiguity in the word ‘is’: The first occurrence (in “something that is a subclass of mammals”) is the ‘is’ of predication; the second occurrence (in “is humans”) is really elliptical for something like ‘is called’ or ‘is named’. And *that* has the further problem that it suggests that Cassie has a belief about “names” or about “calling”, which, at this stage, she does not, for the following reason.

In SNePS, Cassie only has beliefs about entities represented by *nodes*; arcs and their labels are merely structuring devices. From a network point of view, arc labels serve to identify argument positions for an implicit (or unnamed) function symbol. Thus, for example, the `object-property` case frame of Figure 4 is an implicit, 2-place function-symbol, the order of whose arguments is irrelevant since each argument place has a label, viz., an arc label, in a Davidsonian fashion (Davidson 1967). SNePSLOG makes that function symbol explicit (here, we “name” it ‘Is’) and fixes the order of the arguments (so that the arc labels do not need to be visible). Just as in predicate logic, we cannot speak about *either* argument positions (arcs or their labels) *or* function symbols (whether implicit or explicit). We can only speak about terms (i.e., nodes). This is a linguistic analogue of Quine’s (1953) dictum that to be is to be the value of a bound variable; or, as Wittgenstein might have put it, whereof we cannot quantify over, thereof we must be silent. See §1.4, above, and Shapiro & Rapaport 1987 for further discussion.)

3.4.3 The SNePS Analysis: “Promotion” of a Structural Arc to a Conceptual Node

We are now in a position to see Keller’s accomplishment through the eyes of SNePS. In SNePS terms, *before* the well house, Keller’s belief that water is wet could be represented as in Figure 4. Figure 4 is very similar to Figure 3, the major difference being that, in Figure 3, we used the `subclass-superclass` case frame to represent the ‘is’ of class subsumption (“Humans *are*—in the sense of *is a subclass of*—mammals”), whereas, in Figure 4, we use an `object-property` case frame to represent the ‘is’ of predication (“Water *is*—in the sense of *has the property of being*—wet”). In both cases, we represent the principal concepts (*human*, *mammal*, *water*, *wet*) by base nodes—nodes with no structural information. As noted above, logically and linguistically, these are very much like Skolem constants; ontologically, they are very much like bare particulars. We are able to talk about them by giving them names, using the `name-object` case frame. However, although we (and Cassie) can express `b3`, say, by *using* the

‘name’ ‘water’, there is no explicit belief that ‘water’ *is a name for* b3.

After the well house, I suggest, Keller ‘promoted’ the arc-label name to a node, about which she can have an explicit belief, as shown in Figure 5. (From the SNePSLOG point of view, one *might* say that she ‘demoted’ a predicate to a term, but really she ‘promoted’ a (hidden) argument position to a term.) Here, we use the `object-rel-possessor` case frame to represent general possessive expressions of the form ‘*x is y’s z*’, as follows: object *x* stands in the *z* relation to possessor *y* (see Chun 1987 for details).⁴⁴ In SNePSLOG, we could define a predicate:

```
define-frame Possession (nil object possessor rel)
```

where

```
[[Possession(o,p,r)]]= object o is possessor p’s r,  
                        or: p’s r is o,  
                        or: object o stands in the r relation to  
                        possessor p.
```

Then Figure 5 is the result of asserting:

```
ISA-thing-named(b5 , name) .  
Possession(water , b3 , b5) .  
Possession(wet , b4 , b5) .
```

Thus, e.g., `[[m11]] = ‘Water’` is `[[b3]]’s name` (i.e., the object ‘water’ stands in the ‘name’ relation to possessor b3). Cassie (or Keller) can now *say* that ‘water’ is a name for water, because she has an explicit concept of a name, represented by a node, rather than an inexpressible arc label.

Note that the now-explicit concept of a name (`[[b5]]`) is itself expressed by the name ‘name’. At the risk of making the network of Figure 5 illegible, we should also show a node `m13` that represents the proposition that `[[b5]]’s name` is ‘name’. For the sake of clarity, this is shown separately in Figure 6.⁴⁵

⁴⁴Briefly, this case frame is the SNePS analogue of the English possessive constructions (*’s*, *of*, etc.), neutral as to whether the possession is that of whole to part, ownership, kinship, etc. E.g., ‘Bill is Frank’s father’ might be represented as: `Possession(Frank, Bill, father)`, understood as expressing the relationship that Frank *is the father of* Bill. If, in general, `Possession(o,p,r)`, it might follow that *o* is an *r*—e.g., Frank is a father—but only because, in this case, $\exists p$ [Frank is the father of *p*]. As Stuart C. Shapiro (personal communication, August 2006) observes, if one man’s meat is another’s poison, it doesn’t follow that the first person’s meat is meat *simpliciter*, but only meat *for him*, since it is also poison *for* someone else.

⁴⁵In SNePSLOG, node `m13` could be constructed by asserting

```
Possession(name , b5 , b5) .
```

Roger Brown (1973: 3) observed that ‘linguistic processes, in general tend to be invisible. The mind’s eye seeks the meaning and notices the medium as little as the physical eye notices its own aqueous humor’. The well-house experience made one crucial linguistic process visible to Keller. Keller learned more than how to *use* words or signs; she learned that certain things *were* words or signs. She learned more than how to ‘see through’ a sign to its meaning; she learned how to see the sign *as a sign*.

i.e., [[b5]]’s name is ‘name’. This is not so much self-referential as it is the base case of a recursion when rule (1) of §4.3.1, below, is applied to m10. (The first consequent of that rule would not build a new node (because of the Uniqueness Principle; Shapiro 1986); rather, it would return the already-existing node m10. The second consequent builds m13.

4 The Significance of Names.

4.1 Terrace's Theory of Names.

But why is it so significant to learn what a name is? A possible answer to this, consistent with Keller's post-well-house behavior, can be found in Herbert S. Terrace's essay, 'In the Beginning Was the 'Name' ' (1985). In this section, we'll look at Terrace's theory of why "naming" is important, whether his notion of "naming" is akin to Helen Keller's, the extent to which his theory is supported by more recent observations, and the relevance of all this to computational natural-language competence. We'll begin with a brief overview, and then look at his theory in detail.

It will be both useful and convenient to distinguish Terrace's terms from, say, Keller's. So, I will refer to Terrace's theory of names and naming as the theory of 'T-names' and 'T-naming'.

4.2 Overview of T-Naming.

In a letter to the editor of the *New York Review of Books*, Terrace summarized his theory as follows:

Before speculating about the origins of grammar, it is prudent to ponder the origins of the *referential* use of individual words. Unlike apes, children use individual words to comment about objects **for the sheer joy of communicating**. Adults do not reward a child with a tree when she points to one and then says *tree* By contrast, there is no evidence that apes communicate *about* things. As Lord Zuckermann observes, apes use language not as "... a way of conversing, but a game associated with pleasurable reward." Although the origins of human language are unclear, one contributing factor must be the adaptive value of communicating meanings that cannot be expressed in a single word (e.g., *the large tree* or *the single-tusked elephant at the large tree*). It appears, therefore, that the cognitive leap to language occurred in two stages: first, developing the lexical competence to use arbitrary symbols to refer to particular objects and events, and then the syntactic competence to combine and inflect those symbols systematically so as **to create new meanings**. (Terrace 1991: 53; italics in original, boldface mine.)

The game-like nature of language use by apes is reminiscent of Keller's *pre*-well-house use of language, and the child's use of words for objects "for the sheer joy

of communicating” clearly describes Keller’s *post*-well-house behavior. So, *prima facie*, T-naming might well be what Keller learned to do at the well house. It was there that, by her metalinguistic discovery, she “develop[ed] the lexical competence to use arbitrary symbols to refer to particular objects and events”.

Antonio Damasio (1989: 25) has observed that one of the “stages” of concept formation “is that of generating names that are pertinent to the stimulus and are usable to narrate the primary display when inserted in appropriate grammatical structures.” Thus, I perceive an object, which causes neuronal activity representing its features and structure. These, in turn, are linked to other neuronal structures that “generate names”, which allow me to communicate to you that I am thinking of an object and what it is.

Bruner makes a similar observation:

In object request the principal task is to incorporate reference into request. When the child finally masters nominals, he [sic] need no longer depend upon the interpretive prowess of his mother or the deictic power of his indexical signaling. The demands of dealing with displaced reference in requesting objects provide an incentive. (Bruner 1983: 114.)

Having (T-)names gives one power. Keller, apparently, lacked this ability before the well house.

One doesn’t, however, *create* new meanings (*pace* Terrace’s quote, above). Rather, one creates new (combinations of) symbols that are able to be associated with meanings that one had no words for before (cf. Elgin 1984 for a literary exploration of this theme). Zuckermann’s way of putting this is better: “the additional adaptive value of joining lexical items in ways that multiplied the meanings that they can convey” (Zuckermann 1991: 53).

So, for Terrace, syntax is built on top of lexical semantics, as it seems to have been for Keller, too. Bruner, after observing two children’s acquisition of language, concurs:

[R]equesting absent objects . . . requires a degree of specification not needed when an object is within reach or sight. An object out of sight requires the use of nominals for easy specification. . . .

Remote or displaced requests began at the landmark age of fourteen months in both children. (Bruner 1983: 97–98.)

T-naming enables conversation—the exchange of *information*, distant or displaced in space and/or time, with no immediate, practical goals (other than, perhaps, to

satisfy one's curiosity or to be sociable).⁴⁶

Let's now examine Terrace's claims in detail.

4.3 T-Naming.

Terrace considers "The ability to refer with names" to be "perhaps" the most "fundamental" and "uniquely human skill" (Terrace 1985: 1011). This referential ability appears to be akin to symbol grounding. It is the link between word and world (reminiscent of the notion of "alignment" in SNePS; see Shapiro & Ismail 2003). But reference, as Frege taught us, is not all there is to meaning: Is Terrace's notion of referring *Bedeutung*? Is it *Sinn*? What would he say about the referring use of a name like 'unicorn' or 'Santa Claus'? (Note that it is only under certain circumstances that one would use such "non-referring" names to request the named entity.)

4.3.1 SNePS Analysis of Learning to Name

How did Helen Keller learn to name? *What* did she learn? She learned the nature of the relationship between a name (i.e., a finger spelling) and an object. She learned the name of naming. So, is it possible for Cassie to learn to name? Given the network of Figure 4, we'd like her to be able to say, when asked, something like "'water' is the *name* of b3", or (more likely) "'water' is the name of water", or (even more likely) something along the lines of "'water' is the name of this stuff" (since [[b3]] is not otherwise characterized; cf. the earlier discussion of this in §3.4.2.). In any case, the point is for her to do something more than merely express b3 *using* 'water'. But—as we have seen—that would require a network like that of Figure 5.

Similarly, consider the network in Figure 7. This is one way to represent that something is a red, round ball (represented by nodes m19, m17, and m15, respectively).⁴⁷ In SNePSLOG, Figure 7 is produced by first defining a classic AI predicate for class membership:

```
define-frame ISA (nil member class)
```

and then asserting:⁴⁸

⁴⁶A humorous—and relevant—take on the nature of conversation appeared in a *New Yorker* cartoon showing a man, two women, and a chimp, all dressed in suits and ties, sitting in a bar at a table with drinks; the chimp thinks, "Conversation—what a concept!"

⁴⁷Ann Deakin pointed out to me that color is not a good example for Helen Keller! Perhaps taste or smell would be better? On the other hand, for *Cassie*, color might be more accessible than taste or smell (cf. Lammens 1994)!

⁴⁸Lines beginning with semicolons are comments.

```

;;; Something is a thing named "ball".
ISA-thing-named(b6,ball).

;;; Something is a ball.
ISA(b7,b6).

;;; Something is a thing named "round"
ISA-thing-named(b8,round).

;;; The ball is round.
Is(b7,b8).

;;; Something is a thing named "red"
ISA-thing-named(b9,red).

;;; The round ball is red.
Is(b7,b9).

```

Of course,, we would like Cassie to be able to say that red's name is 'red', etc. Again, this would require additions to the network of Figure 7 as shown in Figure 8, produced, using SNePSLOG, by asserting:

```

Possession(ball,b6,b5).
Possession(round,b8,b5).
Possession(red,b9,b5).

```

We might also like Cassie to be able to say, when asked, not merely that the round ball is red, but explicitly that red is a *property* of the round ball. So, we would need to be able to have Cassie answer the following sort of questions: Given a propositional node asserting that some object *a* has some property *F*, what is the (name of the) relationship between *a* and *F*?

In Figure 4, Cassie can *use* 'water' as the name of [[b3]] (i.e., she can call [[b3]] 'water'). However, without a node explicitly asserting (i.e., naming) the relationship between the node labeled 'water' and b3, she does not understand—*de dicto*—that 'water' is [[b3]]'s *name*. She has no theory of names or naming. Similarly, without a node explicitly asserting (i.e., naming) the relationship between b9 and b7 in Figure 7, she does not understand—*de dicto*—that b9 (i.e., red) is a *property* of b7. She has no theory of properties, either. (Most people other than cognitive scientists probably don't!)

Although these two situations are analogous, there is, perhaps, a slight advantage to the naming case over the property case (in addition to the fact that

most people who are not cognitive scientists *do* have a theory of naming!). For in order for Cassie to *utter* something about [[b3]] or [[b4]] (in Figure 4), she must use a word for it, whereas she would not need to use the word ‘property’ in uttering m17 or m19 (in Figure 7). (This may explain why theories of naming are probably more common than theories of properties. ‘Property’ (in this sense) is a technical term of philosophy, not a term of ordinary language.)

So Cassie could, perhaps, recognize that there is a relationship between her concept of an object and the word she uses when she says something about it (e.g., in Figure 4, between b3 and the node labeled *water*), though this is a task for empirical investigation. From this recognition, she could come to believe a proposition that would link these nodes with a node for the relationship, which, if she asked us, we could tell her was called ‘naming’. So she could assert node m11 in Figure 5. In any event, let us suppose that this can be done. It is, it would seem, what Keller did.

In SNePS, this could be done via rules such as the following:

1. If x is y (in the sense that x is named, or called, y), then y is x ’s name;⁴⁹

more precisely:

For any (two nodes) x and y , if object x has name y , then object y is possessor x ’s name,

or, in SNePSLOG:

```
all(x,y)(ISA-thing-named(x,y) =>
ISA-thing-named(b5,name) ,
Possession(y,x,b5)).
```

2. If x is y (in the sense that x has property y), then y is a property;⁵⁰

more precisely:

For any (two nodes) x and y , if object x has property y , then y is a member of the class of things called ‘property’;

or, in SNePSLOG:

```
all(x,y)(Is(x,y) =>
ISA-thing-named(b10,property) , ISA(y,b10)).
```

⁴⁹Note that, arguably, it does not necessarily also follow that y is a name *simpliciter*; see note 44. A similar rule appears in Shapiro 2003; cf. Shapiro & Ismail 2003.

⁵⁰Note that here it arguably *does* make sense to say that y is a property *simpliciter*, not merely that it is x ’s property; see note 49.

The trick would be to get Cassie to *learn* such rules, a worthy topic for future research.

4.3.2 Terrace's and Keller's Theories of Naming

Some of Terrace's claims about the linguistic abilities of apes are reminiscent of Keller's pre-well-house linguistic abilities:

...even though apes can learn substantial vocabularies of arbitrary symbols, there is no a priori reason to regard such accomplishments as evidence of human linguistic competence. After all, dogs, rats, horses, and other animals can learn to produce arbitrary "words" to obtain specific rewards. (Terrace 1985: 1012.)

Keller 'learn[ed] substantial vocabularies of arbitrary symbols', too. (Cf. Sullivan's use of expressions like 'Keller knew *n* words'.) But what kind of learning is this? Given the context of Terrace's paper, it does not seem to mean that apes (or Keller) could link the arbitrary symbols to objects. And given Terrace's belief in the logical and chronological priority of naming over syntax, it does not seem to mean that the apes (or Keller) knew the syntactic roles of the symbols.

Now, Keller, too, could 'produce arbitrary 'words' to obtain specific rewards'. So, by 'learning a symbol', Terrace must mean producing the symbol in return for a (not necessarily associated) reward, without any (semantic) linking of the symbol with the world.

It would be just as erroneous to interpret the rote sequence of pecks [by a pigeon], red → green → blue → yellow, as a sentence meaning, *Please machine give grain*, as it would be to interpret the arbitrary sequence of digits that a person produces while operating a cash machine as a sentence meaning *Please machine give cash*. In sum, a rote sequence, however that sequence might be trained, is not necessarily a sentence. (Terrace 1985: 1014.)

This sounds like Keller's pre-well-house use of language. But *why* aren't those rote sequences sentences with those meanings? Granted, perhaps, they lack the exact grammatical *syntax* of those *sentences*, but why say that they lack the same meaning? Propositions (meanings) can be *implemented* (i.e., expressed) in numerous different ways, even within the same language; after all, paraphrases exist. When I push a sequence of buttons on a cash machine, why am I *not* telling (asking) it to give me a certain amount of money? Isn't that what the input-output encoding scheme amounts to? Isn't there at least an aspect of meaning (or a kind of meaning) that is determined by how an expression is used or by the context that it

appears in (cf. Rapaport & Ehrlich 2000, Kalderon 2001)? Granted, perhaps what the symbols mean *to me* is not what the symbols mean (if anything) *to the machine*, but as long as we can communicate (so as to overcome misunderstandings), what's the difference? And, in any case, it *does* mean *something* to the machine—it has syntactic meaning (internal meaning).

A brief example of a symbol having two such different meanings (a meaning for the “speaker” or user, and another for the “hearer” or “receiver”) might be instructive. In my university library, when I push the button for the elevator on the ground floor, the button lights up. I have come to learn empirically that if the light stays on when I release the button, it means that the elevator is not on the ground floor. If the light immediately goes off, it means that the elevator *is* on the ground floor and that in a few seconds the door will open. The light's going off is a symbol that *I* interpret to mean “the elevator is on the ground floor; the door will open shortly”. This is its meaning *for me*: *my* interpretation of it. It is important to note that I have determined this meaning empirically (in particular, contextually; cf. Rapaport & Ehrlich 2000), and I could be wrong. If the light goes off and no elevator door opens within a few seconds (and it is in working order), I would have to revise my beliefs.

There is also its meaning *for the elevator system*: the role that the light going off plays in the electrical network that controls the elevator. A study of the wiring diagram might reveal, let us suppose, that when the button is pushed, a circuit is closed that lights the button and a test is conducted to determine the location of the elevator. If the elevator is on the ground floor, the circuit is opened, turning off the light, and, a short time later, another circuit is closed, opening the door; else, the circuit remains closed, keeping the light on, and another circuit is closed, sending the elevator to the ground floor, where the light-circuit is opened, turning off the light, followed by the door opening as before. The meaning of the light's going off, then—its role in that network—is correlated with the elevator's being on the ground floor. From the elevator's point of view, so to speak, the only way the light going off would *not* mean that the elevator is on the ground floor would be if the elevator were broken.

One thing missing from such behavioral uses of language is the *intention* to communicate an idea by using a certain word (Terrace 1985: 1017). Non-human animals who have been trained, behavioristically, to “use language” (which I place in scare quotes so as not to beg any questions about what it is they are actually doing) *seem* to communicate intentionally with each other. But, Terrace points out,

That would presuppose not only that Jill [one of the pigeons mentioned earlier] could discriminate each color from the others (when she clearly could) but that Jill also understood that (a) some

arbitrary communicable symbol described color_i, (b) she sought to communicate to Jack [the other pigeon] that the color she saw was color_i, and (c) Jack would interpret Jill's message as Jill intended it. There is no evidence to support any of these suppositions. (Terrace 1985: 1016.)

This, of course, does not affect experiments in computational linguistics that endeavor to provide a mechanism (based on the theory of speech acts) for implementing intentions to speak.⁵¹ Indeed, that's one of the advantages a computer has over non-human animals: The latter might not have intentions to communicate, but the former can be given them!

But, according to Terrace, even if it could somehow be shown that a non-human animal *intended* to use a certain word to communicate the idea that it wanted a specific object, that would not suffice to show that it was using the word as a *name* for the object. Why? Presumably because the animal might believe that using that word is the best way to cause the listener to give it the desired object. Roughly, the animal might be ascribed the belief, "If I make such and such a noise [or use such and such a sign], my trainer will bring me one of those sweet, red, round things to eat, so I'll make that noise [or use that sign]".

So, for Terrace, *naming* is a very special activity:

... the main function of such words [viz., the "use of a symbol as a name"] in the use of human language—[is] the transmission of information from one individual to another *for its own sake*. (Terrace 1985: 1016–1017; my italics.)

That is what I call "T-naming". Terrace goes on to say that

... a child will utter a name of an object, person, color, and so on, simply to indicate that she or he knows that the object she or he is attending to has a name and also to communicate the fact that she or he has noticed that object

... In many instances, the child refers to the object in question spontaneously and shows no interest in obtaining it. The child not only appears to enjoy sharing information with his or her parent but also appears to derive intrinsic pleasure from the sheer act of naming. (Terrace 1985: 1017.)

⁵¹Bruce 1975; Cohen & Perrault 1979; Allen & Perrault 1980; Cohen & Levesque 1985, 1990; Grosz & Sidner 1986; Haller 1994, 1995.

A mere *associative* link between an arbitrary symbol and a specific object⁵² is not sufficient for a *semantic* link, according to Terrace. What is also needed is *intending* to use the symbol for the object *for no purpose other than to indicate that you are thinking of the object*.

So, one difference between Keller’s pre- and post-well-house language might be this: *Before*, she didn’t have such intentions; *after*, she did. That is, *after*, she *intended* to refer to water by ‘water’. To do this, she needed to be able to think and talk about the naming relationship. But is it really the case that she lacked that intention before? Although the evidence is at best unclear, I think she did have the intention, but not the name of the naming relationship, so that her intentions were often frustrated.

Searle might say that Searle-in-the-room (or the Chinese Room itself, on the systems reply) *says* things but doesn’t *mean* them. What, then, does it mean to mean something by a word? In Cassie’s terms, it would be this: Cassie has a concept that she wants to communicate to Oscar.⁵³ She has a name for the concept. So, she utters that name, *assuming* that Oscar uses that word for the “same” (i.e., the corresponding) concept (in his mind), and *intending* that Oscar will think of that concept—that that concept will be “activated” in Oscar’s mind. As Terrace puts it,

In most human discourse, a speaker who utters a name expects the listener to interpret the speaker’s utterance as a reference to a jointly perceived (or imagined) object (Terrace 1985: 1017.)

“Jointly imagined” is where such entities as unicorns and the Hob-Nob witch⁵⁴ come in. So, T-naming is more appropriately a relationship between a name and a mental concept. That is, in answer to the question raised at the beginning of this section, it is more like *Sinn* than *Bedeutung* (actually, it is more Meinongian—cf. Rapaport 1981). So, in the Chinese-Room Argument, what’s missing from Searle’s description is the *intention* to communicate.⁵⁵ So one *could* argue that if Searle’s

⁵²Perhaps like that between the node labeled humans and node m1 in Fig. 1 or between the node labeled humans and node b1 in Fig. 3. (The latter associative link is represented by node m4.)

⁵³Oscar is the **O**ther **S**NePS **C**ognitive **A**gent **R**epresentation, first introduced in Rapaport, Shapiro, & Wiebe 1986.

⁵⁴“Hob thinks a witch has blighted Bob’s mare, and Nob wonders whether she (the same witch) killed Cob’s sow” (Geach 1967: 628).

⁵⁵I.e., the intention to communicate should be one of the features of computational natural-language competence in addition to those I cited in Rapaport 2000, §8. There, I said that a computational cognitive agent must be able to ‘take discourse (not just individual sentences) as input; understand all input, grammatical or not; perform inference and revise beliefs; make plans (including planning speech acts for natural-language generation, planning for asking and answering questions, and planning to initiate conversations); understand plans (including the speech-

Chinese Room is to be taken literally, then it *doesn't* understand, but that would only be because Searle hasn't fleshed out the full theory of computational natural-language competence.⁵⁶ And if it's to be taken as schematic for whatever would be a full theory of computational natural-language competence, then he's wrong.

As both Terrace and Bruner (whose Language Acquisition Support System augments the Chomskyan Language Acquisition Device—a LASS for a LAD; Bruner 1983; cf. Rapaport 2003a) point out—and as Anne Sullivan pointed out long before them—

... language draws upon certain kinds of nonlinguistic knowledge. For example, before learning to speak, an infant acquires a repertoire of instrumental behavior that allows her or him to manipulate and/or approach various objects. An infant also learns how to engage in various kinds of social interaction with her or his parents—for example, being able to look where the parent is looking or pointing. Eventually, the child learns to point to things that he or she would like the parent to notice. In short, the infant first masters a social and conceptual world onto which she or he can later map various kinds of linguistic expression. (Terrace 1985: 1018.)

So, *internal concepts* are learned first, via links between visual input and mental concepts.⁵⁷ Names are attached later.

This suggests, by the way, that the first rule in §4.3.1 need not incorporate a name right away; i.e., Cassie might have an unnamed concept. (Again, cf. Elgin 1984.) Thus, the rule might be:

If x is (named) y , then y is x 's *something-or-other*;

more precisely:

For all (nodes) x and y , if object x 's name is y , then object y is possessor x 's *something-or-other*,

or, in SNePSLOG:

act plans of interlocutors); construct a "user model" of its interlocutor; learn (about the world and about language); have lots of knowledge (background knowledge; world knowledge; commonsense knowledge; and practical, "how-to", knowledge ... and remember what it heard before, what it learns, what it infers, and what beliefs it revised ... And it must have effector organs to be able to generate language. In short, it must have a mind."

⁵⁶I.e., the ability to understand and to generate natural language; see Shapiro & Rapaport 1991.

⁵⁷Cf. PIC arcs connecting SNePS nodes representing concepts and SNePS nodes representing pictorial images of them: Rapaport 1988b; Srihari & Rapaport 1989, 1990; Srihari 1991ab, 1993, 1994. Also cf. anchoring or "alignment"; Shapiro & Ismail 2003.

```
all(x,y)(ISA-thing-named(x,y) =>
Possession(y,x,something-or-other))
```

Here, the “base” node (roughly, Skolem constant) labeled something-or-other might *later* be “given” a name by asserting a proposition to the effect that whatever object is represented by that node is expressed by the word ‘name’; in SNePSLOG:

```
ISA-Thing-Named(something-or-other,name).
```

(Here, ‘name’ appears embedded in the predicate, about which predicate we cannot speak; it is hidden as the label for an argument position in the definition of the predicate, about which position we also cannot speak; and it appears explicitly as a term, which *can* be spoken about.) Alternatively, we could have a “defeasible” rule (with three conjoined antecedents) to the following effect:

```
all(x,y,z,r)(Is(z,unknown),Possession(x,y,z),Possession(x,y,r)
&=> Equivalent(z,r)).
```

where⁵⁸

```
define-frame Equivalent (nil equiv)
```

and

```
[[Is(o,unknown)]] = object o has the property of being
[[unknown]]
[[Equivalent(x1,... ,xn)]] = mental entities x1,... ,xn
correspond to the same actual object.
```

I.e., if z is an unknown (or unfamiliar) concept, and if x is y 's z , and if x is (also) y 's r (where r is some antecedently *familiar* concept), then (defeasibly, of course!) z is (or: might be) equivalent to r . (The “defeasibility” is automatic in SNePS: Virtually all beliefs are revisable in the light of later beliefs, so there is very little “cost” to this somewhat overgenerous rule; in any case, the proposition about being “unknown” serves to limit its application somewhat.) So, if something-or-other is an unknown concept, and if x is y 's something-or-other, and if x is also y 's name, then (perhaps) a something-or-other is a name. This is, at least, a plausible hypothesis, subject to further modification (cf. Rapaport &

⁵⁸Note that the predicate “Equivalent” is defined in terms of a single arc (“equiv”) that can point to a *set* of nodes; this has the same effect as having a set of arcs with the same label, each of which points to a node. See Maida & Shapiro 1982; Shapiro & Rapaport 1987; Rapaport, Shapiro, & Wiebe 1997 for further discussion of the SNePS notion of equivalence.

Ehrlich 2000, Rapaport 2005a). (See the appendix for a computer demonstration of this and other computational aspects of our story.)

Keller, thus, was able to learn language once she grasped the concept of naming—“the conventions of using symbols and words that do the work of referring” (Terrace 1985: 1021). Apes, according to Terrace, lack this ability (pp. 1021, 1023–1024). More specifically, “an ape does not naturally refer to an object to which it attends *solely for the purpose of noting that object to a conspecific*. . . . [W]hatever referential skills an ape displays naturally seem to be in the service of some concrete end” (p. 1024, my italics). Keller’s post-well-house interest in the names of objects seems to have been for its own sake. Her own pre-well-house signs were always “in the service of some concrete end”.

5 Can Apes Speak for Themselves?

But is Terrace's emphasis on intentional but non-purposive naming a reasonable one? Why does Terrace think that using a sign in order to get a reward is *not* using it linguistically (cf. Terrace 1985: 1016–1017)? What is so important about T-naming? And how do we know that apes *don't* have intentions when they 'use language'? Finally, is there any evidence that apes *do* T-name?

5.1 Purposive Naming

Terrace has two reasons for believing that purposive-only use of names is *not* language. First, he claims that it is simply a matter of fact that apes *don't* use signs except when they want something, whereas human children at 18 months *do* (p. 1023). So, at least, what apes do is *preliminary* to full human-like language use.

Second—and this is what's important about T-naming, according to Terrace (1985: 1011)—without it, grammatical syntax would not develop:

... when there is a desire simply to communicate information about a relationship between one object or action and another, about some attribute of an object, or about past or future events ... ungrammatical strings of words would not suffice—hence the functional value of syntax. (Terrace 1985: 1026.)

His argument seems to be this: If I want something that is present in the environment containing you and me, I can use a name or make some (perhaps non-semantic) sign to get you to give it to me. If a single name is insufficient (say, because there are two things, and I want a specific one, e.g., the *large* banana), I can combine two (or more) names. But it doesn't really matter in what order I combine them (since we can both see the items), so grammar (beyond mere juxtaposition) is not necessary. If I don't know a name for the object, I can point to it. But for *absent* (displaced) objects, pointing won't work. And if I wanted to communicate about some *feature* of an (absent) object, then grammar facilitates my communication: If I wanted to talk about the color of the banana, it becomes crucial *which* of the previously juxtaposed signs is the color-term and which the noun; grammar enters upon the scene. Note that it is highly unlikely that I *want* the color; rather, I just want you to know that I'm *thinking* about the color—that's T-naming.

5.2 Intentions

As for intentions, we *don't* know that apes don't have them. I'd bet on evolutionary grounds that they do. We *can*, however, insure that a *computer's* use of language

would be intentional, by having the natural-language-competence program include speech-act or intention-and-action modules. Searle might object that that's just more syntax, not "real" intentions or desires. But what would be the difference? What is a "real" intention? Moreover, desires (and intentions) *can* be adequately incorporated. In our contextual-vocabulary-acquisition project, there might be times when the system, in order to settle on a definition, would need more information. That need—and why wouldn't it be a "real" need?—could prompt it to seek that information, to ask questions—and why wouldn't these be "real" desires for information or intentions to ask questions?⁵⁹

One might reply that such computational desires or intentions have no qualitative "feel" to them. Perhaps. Qualia may best be seen as a feature of the implementing medium (rather than what I call the Abstraction, as I argue in Rapaport 2005b). So, of course, the computational desires and intentions *might* have a "feel", depending on their implementation. Or they might not. But why would a lack of "feeling" disqualify them as "real" desires or intentions?

The thoughts, desires, and intentions of a language user that is not an ordinary human—an ape or a computer (or even, perhaps, a Helen Keller)—might be very different in kind from those of a normal human (as the cartoon described in note 46 suggests). They might very well depend on the language user's body and purposes. But they would be thoughts (or desires or intentions) nonetheless.

5.3 From Representation to Language

Is Terrace right about the inability of non-human primates to T-name? Several papers published *after* Terrace's deal with the issue specifically, which we look at in the next three sections.

Terrace claims that apes lack something that humans have that enables humans to have language. One researcher—Jacques Vauclair (1990)—is sympathetic to this, though it's not clear that that something is T-naming.

According to Vauclair (p. 312), both human and non-human primates have "basically similar ways of coding environmental stimuli in terms of cognitive organization"; i.e., they have "mental representations". But they do *not* share language. Why not? For Vauclair, it seems, it's partly by definition:

Representation is an individual phenomenon by which an organism structures its knowledge with regard to its environment. This knowledge can take two basic forms: either reference to internal substitutes (e.g., indexes or images) or use of external substitutes (e.g., symbols, signals, or words).

⁵⁹In Shapiro 1981, SNePS asks questions as a result of back-chaining.

Communication is a social phenomenon of exchanges between two or more **conspecifics** who use a code of specific signals usually serving to meet common adaptive challenges (reproduction, feeding, protection) and promote cohesiveness of the group.

(Vauclair 1990: 312; italics in original, my boldface.)

Since apes and humans are not conspecifics, they cannot, *by definition*, communicate with each other. Even if that restriction is lifted, it is not clear whether T-naming is Vauclairian communication (unless, perhaps, as a by-product, it ‘promote[s] cohesiveness of the group’—perhaps that’s the function of conversation (as the cartoon described in note 46 suggests).

Language is conceived as a system that is both communicational and representational: It is grounded in social conversation that attributes to certain substitutes (called *signifiers*) the power to designate other substitutes (called *referents*).

(Vauclair 1990: 313.)

So, apes and humans could never have a common language, because language is communicational (and, by definition, apes can’t communicate with humans). But how does this definition of language make it *human-specific*? Perhaps it is the social-communication aspect (cf. Bruner’s notion of ‘negotiation’, discussed in Rapaport 2003a). After all, apes and humans don’t share a common ‘society’.

The closest Vauclair gets to supporting Terrace’s theory is in claiming that two of the marks of language are its ability to deal with displacement *in* space and time (i.e., things not in the presence of the speaker) and its ability to deal with what might be called displacement *from* space and time (i.e., dealing with non-existents) (Von Glasersfeld 1977, cited in Vauclair 1990: 316). T-naming, however, is logically *independent* of this. For one could, in principle, be able to refer to something displaced in (or from) space or time either if one wanted it or if one merely wanted to talk about it ‘for its own sake’. And, clearly, one could be able to refer to something in one’s current environment for either of those reasons.

So several issues are still open: *Do* non-human primates T-name? (Terrace, of course, says ‘no’.) Do they use language to talk about displaced objects? One would expect Vauclair to delve into this. Instead, he locates the gap between ape and human language elsewhere:

I am convinced that apes display the most sophisticated form of representation in the animal kingdom . . . , but this phenomenon is insufficient in itself to qualify for linguistic status. To go beyond the 1-1 correspondence between the sign and the actual perceptual

situation, we need to introduce a third term. The relation between symbol and object is more than the simple correspondence between the two. Because the symbol is tied to a conception, we have a triangular connection among objects, symbols, and concepts: “It is the conceptions, not the things, that symbols directly mean” (Langer, quoted in von Glasersfeld, 1977). (Vauclair 1990: 320.)

Now, I am happy to agree with Langer; her claim is an aspect of syntactic semantics. But it’s not clear what that has to do with Vauclair’s point. He *seems* here to be saying that what’s missing is the concept: no concept, no language. Yet earlier he claimed that representation *required* concepts: Although ‘concept’ is not part of his *definition* of ‘representation’, on pp. 313ff he talks about ‘internal processing’, ‘internal representation’, ‘cognitive maps’, ‘internal coding’, and ‘internal substitutes’. What are these if not concepts?

Later (p. 321), he locates the gap ‘in the emergence in humans of verbal language’, but he is silent on what these emergent features are; perhaps it is T-naming. Or perhaps it is being intentional:

The specificity of human language is above all of functional order. First, this system uses representative stimuli that allow the sender to know the status of the sent message, to control it, and to endow it with intentions. (Vauclair 1990: 321.)

Of course, as we have seen, this won’t distinguish between human and *computer* use of language. Perhaps, however, this *was* something Helen Keller lacked before the well house.

The other thing that non-human primates lack is the social convention, the Brunerian negotiation of meaning (p. 322). This, however, seems irrelevant to T-naming. In any event, Helen Keller, arguably, had this *before* the well house, and computers certainly can have it (witness, e.g., the vocabulary-acquisition process described in Rapaport & Ehrlich 2000; also see Rapaport 2003a).

5.4 Orangutan Reference.

Can non-human primates T-name? H. Lyn White Miles’s (1990) work with the orangutan Chantek is suggestive of T-naming in such a primate.

Chantek was clearly capable of “displaced reference” (pp. 520–523), and four out of about 97 cited uses of *names* do not *appear* to involve wanting the object: making the signs (1) ‘car’ “as he passed [his] caregiver’s car on a walk”, (2) ‘time’ “when [his] caregiver looked at her pocket watch”, (3) ‘Coke drink’ “*after* finishing his Coke” (my italics), and (4) ‘time drink’ “when [his] caregiver looked at her

watch” (pp. 520–523). Each of these, however, could be interpreted otherwise: (1) Perhaps Chantek was tired of walking and wanted to ride in the car; (2) perhaps he wanted to know the time (though it’s hard to believe that he had the appropriate concepts for understanding time) or perhaps ‘time’ was also his sign for the watch itself (we are not told); (3) perhaps he wanted another Coke to drink; (4) perhaps he was thirsty. It is hard to know when a naming act is a *T-naming*. So T-naming may be an overly restrictive criterion.

On the other hand, those who are more sympathetic than Terrace to the view that apes can use language tend to have criteria that are overly permissive. Consider Miles’s three “elements” of “linguistic representation” (p. 524):

1. A sign must designate an element of the real world.
2. A shared cultural understanding about its meaning must exist.
3. The sign must be used intentionally to convey meaning.

The first element is surely too strong, since we can talk about nonexistents. Moreover, it would seem better to say that a sign must *be used by someone to* designate something (where ‘something’ is construed broadly, along Meinongian lines).

The second element seems to rule out interspecies linguistic representation and, perhaps, computer language. Elsewhere, I have discussed what I call “Winston’s Problem” (Rapaport 2003a, §10), which concerns what might happen if the knowledge-representation language (i.e., language of thought) of a computer system that can learn concepts differs significantly from that of humans. According to Patrick Henry Winston (1975/1985: 143) (and Ludwig Wittgenstein 1958: 223)—“If a lion could talk, we could not understand him.”), what would happen is that the two systems—computer and human (or lion and human)—would not be able to understand each other. However, in the various ape experiments, both subject and experimenter are using an artificial language, so they do have a shared cultural understanding, where the “culture” is that of the laboratory. Granted, the sign for Coke may have all sorts of connotations for the human but not for the chimp. But that’s no different from the fact that the word ‘Coke’ has all sorts of connotations for *me* but not for *you* because of our different experiences with it. The case of the computer is a bit easier, since *we* get to *give it* its cultural knowledge. Hence, insofar as the computer has a ‘mind’ (i.e., a knowledge base; cf. Rapaport 1995, §1.1.3), both it and we can have “shared cultural understanding”, *pace* Dreyfus and others (as long as we avoid Winston’s Problem).

Helen Keller’s pre-well-house uses of finger-spelled words seem in some cases to have designated in the sense of element (1) (e.g., some of her uses of ‘cake’ and

'doll'). Even her confused use of 'mug' and 'milk'/'water' might be taken to have designated the mug-plus-liquid complex. Clearly, before the well house, she could designate via her own signs. Arguably, her inability to clearly designate with finger spellings could be attributed to insufficiencies in her *shared* cultural understanding. She clearly shared in some cultural understanding—after all, she was a human, living with other humans. But, of course, she was blind and deaf, hence cut off from much that the rest of us share without even realizing it. Finally, though she used her own signs intentionally to convey meaning, most of her pre-well-house use of finger spellings was no doubt mere mimickry.

Again, Miles's criteria for referential use of words or signs is weaker than Terrace's:

first, that signs can be used to indicate an object in the environment; second, that signs are not totally context dependent; third, that signs have relevant semantic domains or realms of meaning; fourth, that signs can be used to refer to objects or events that are not present. (Miles 1990: 524.)

(I take the third criterion to mean that there is a systematic correlation between sign and referent.) All of these are necessary—but not sufficient—for T-naming. One of the essential aspects of T-naming is that there be no desire to *have* the object named—no ulterior motive.

However, Chantek showed some behavior that seems to be part of T-naming when he would show his caregivers some object (pp. 524–525). Since he already had the object, it would seem that he had no other purpose for showing it than to get his caregivers to understand that he was thinking about it. This behavior, when combined with displaced reference, surely lays the groundwork for eventual T-naming.

Is T-naming a significant mark either of human language development in particular or of language development *simpliciter*? Granted that Helen Keller exhibited it after (and apparently *only* after) the well house, it would seem that it is significant for humans (or, at least, for her). And if Chantek either could easily have exhibited T-naming, or in fact did exhibit it (on occasion), it might *not* be unique to human language. It certainly makes for more sophisticated use of language (the ability to tell stories, the ability to fabricate), and it does make language learning easier. Yet there's an awful lot of linguistic behavior that apes such as Chantek are capable of that makes one wonder why Terrace requires that, in order to T-name, the language user must *not* want the object. Chantek, for instance, learned labels for things he wanted, displayed displacement reference for things he wanted, and used language to deceive in order to get something (pp. 526–529). And Chantek,

apparently, was capable of a metalinguistic achievement that, again, could underlie eventual T-naming:

Chantek used the blades of a scissors instead of his hands to make the sign for biting By transferring the total shape of the sign, including configuration and movement, to another means of expression, he showed that he understood that the sign was an abstract representation in which the composite elements stood for something else. (Miles 1990: 530.)

Indeed, some of the beginnings of what looks like T-naming can be seen in the following passages:

The second stage of development, that of subjective representation . . . ranged from 2 years to almost $4\frac{1}{2}$ years of age In this stage, Chantek used his signs as symbolic representations, but his perspective remained subjective. He gave the first evidence of displacement . . . and developed proximal pointing, which indicated that he had mental representations. . . . He elaborated his deception and pretend play He showed evidence of planning through mental representations and signed to himself about objects not present. . . . For the first time he also used signs in his deceptions. (Miles 1990: 534–535.)

The third stage, nascent perspective taking, ranged from about $4\frac{1}{2}$ years to over 8 years of age, during which his vocabulary increased to 140 signs Chantek's representations became more objective and moved toward perspective taking, the ability to utilize the point of view of the other. . . . Most important, he was able to take the perspective of the other by getting the caregiver's attention and directing the caregiver's eye gaze before he began to sign.

It was at this point that he invented his own signs. . . . He clearly understood that signs were representational labels, and he immediately offered his hands to be molded when he wanted to know the name of an object. (Miles 1990: 535.)

How reminiscent of Helen Keller's post-well-house behavior, whether or not it is T-naming!

5.5 Against T-Naming.

Two arguments can be mounted *against* the significance of T-naming. The first, due to Patricia Marks Greenfield and E. Sue Savage-Rumbaugh (1990), is based on

possible biases on the part of researchers. Terrace's claim that apes don't T-name is apparently supported by evidence such as that "Kanzi [a pygmy chimpanzee] had a much smaller proportion of indicatives to statements (4%) in comparison with requests (96%), than would be normal for a human child" (Greenfield & Savage-Rumbaugh 1990: 568). But, as Greenfield and Savage-Rumbaugh point out, an alternative explanation is that this is an artifact of their artificial, human-controlled environment, in which they *must* request things. By contrast, "In the wild, a given animal might *state* his planned activity, rather than *requesting* it" (p. 568). They suggest that if we studied human language development without the assumption that children will eventually succeed in learning language, we might *not* ascribe T-naming to them at the analogous developmental stage at which we deny it to apes (p. 571).

The second, perhaps weaker, argument against T-naming focuses on just what it is that a speaker intends to communicate. T-naming certainly involves a desire to communicate—but to communicate what? For Terrace, it is the desire to communicate that the speaker is thinking of a distal object. The speaker is playing a sort of "guess what I'm thinking about" game, using a word that means *what he or she is thinking about*. But that notion—what the speaker is thinking about—is ambiguous between the actual object (a *de re* interpretation) and the concept in the speaker's mind (a *de dicto* interpretation). However, since the speaker can be thinking of an object that doesn't exist, or a proposition that may lack a truth value, the *de re* interpretation can fail. Only the *de dicto* interpretation can be consistently maintained in all cases (Rapaport 1976, 1978, 1981, 1985/1986, 1986a; Rapaport et al. 1997). As a consequence, *all* uses of names appear to turn out to be T-naming, i.e., the use of a name for something that one is merely thinking of.

However, whether or not non-humans are capable of it, T-naming by itself is probably not sufficient for full language use in conversation. As W. Tecumseh Fitch (2006: 370) puts it, "Chimpanzees certainly have important ingredients necessary for human language (for example, the ability to pair arbitrary signals with meanings), but they are still not discussing philosophy or even what they had for dinner yesterday."

6 Return to the Well House.

What was the significance of the well-house episode?

Negotiation is crucial to understanding language (Rapaport 2002, 2003a; Arrighi & Ferrario 2005). When a speaker uses a word in a conversation, all participants in the conversation must try to align the concepts that each finds or constructs in their minds. Often, an interlocutor has to *merge* two concepts (e.g., ‘Oh, John Williams the former conductor of the Boston Pops whom you’re talking about is John Williams the composer whom I’m thinking of!’) or to *split* one into two (e.g., ‘Oh, John Williams the classical guitarist whom you’re talking about isn’t John Williams the conductor and composer whom I’m thinking of!’; cf. Maida & Shapiro 1982). So, one thing that was significant about Keller’s experience at the well house was that two of *her* concepts merged or were aligned: *her* concept of water (previously linked to ‘wah-wah’) and *her* concept of *Sullivan’s* concept of water. Prior to the well house, Sullivan *thought* that these were aligned ideas, but, in fact, they weren’t.

Moreover, the well house itself played a significant role:

... a key feature of human referring acts ... [is that] [t]hey are highly context sensitive or deictic. Parties to a referring act infer its referent from an *utterance* in a *context*. ... John Lyons argues that deixis is the source of reference, that ‘locating in context’ rather than simply ‘tagging’ is the heart of reference ... (Bruner 1983: 69–70.)⁶⁰

Keller’s experience was significant because the context was extremely simple: water in one hand, ‘water’ in the other.⁶¹

One might reasonably expect to find, then, that the acquisition of referring procedures is heavily dependent on the ‘arranging’ and simplifying of contexts by the adult to assure that deictic demands be manageable for the child. (Bruner 1983: 70.)

The story of Helen Keller is a fascinating one. Not only teachers of language skills, but every teacher ought to read Sullivan’s letters and reports on her teaching methods. But, equally, Keller was an amazing pupil; one wonders what would have become of her had she *not* been blind and deaf! Consider, for example, the

⁶⁰I consider other aspects of Bruner’s book in Rapaport 2003a, §8. On the role of deixis in natural-language understanding, cf. Bruder et al. 1986; Rapaport, Segal, Shapiro, Zubin, Bruder, Duchan, Almeida et al. 1989; Rapaport, Segal, Shapiro, Zubin, Bruder, Duchan & Mark 1989; Duchan et al. 1995.

⁶¹Actually, as we saw, there was a mug in the water hand, but it seems to have been ignored. Cf. §3.3.1, observaton 3, above.

large number of syntactic systems with which she was familiar: finger spelling (the manual alphabet, *tactually* understood), lip reading (again, *tactually* understood), the typewriter, three varieties of Braille, the Roman alphabet (again, *tactually* understood), oral speech (her own)—Keller’s knowledge of speech is also akin to the Chinese Room; she had a ‘syntactic’ knowledge of speech, since she couldn’t hear herself (cf. Keller 1905: 327)—Morse code, English, French, German, Latin, and Greek (and probably the Greek alphabet in some form).

Now, Searle-in-the-room also knows a syntactic system—squiggles—which are known to others as Chinese writing. The task for Searle-in-the-room is to get ‘beyond’ the syntax. But what lies beyond? Ideas (mental entities?)? Objects? In general, of course, his task is to get to what the squiggles *mean*. How? Well, clearly, the more squiggles, the better. Note that *much* of Helen Keller’s learning was *book*-learning, which is purely syntactic (cf. Keller 1905: 30, 318; but cf. p. 317). But also Searle-in-the-room needs more experiences, even if only self-bodily ones. But, ultimately, *all* such experiences are internal(ly represented), just as are (the experiences of) the squiggles.

Ditto for Helen Keller. What she learned was not merely that everything has a name, but also that there is a naming relation, and she learned the name of naming. She was thus able to take a large network of mental entities, some of which were representatives of external objects (most of which, in turn, were internalized by the sense of touch) and some of which were representatives of external words (also internalized by the sense of touch), and partition it into two parts with explicit relations between them. She imposed a semantic structure on a domain hitherto only understood syntactically. When she was able to *organize* all her internal symbols such that some were names for others (and some of the ‘others’ were directly linked to her experiences), she began to get beyond the syntax to the meanings (cf. Keller 1905: 169). It was still a syntactic system, but now had semantic organization. *The organizing principle was discovered at the well house*. And that’s how Helen Keller escaped from her Chinese Room.

7 Concluding Remarks

In a footnote to an important essay on software engineering, David L. Parnas (1972: 331, fn. 1; my italics) said:

It should be clear that while we cannot afford to use natural language specifications [for “program segments”, p. 330] we cannot manage to do without natural language explanations. *Any formal structure is a hollow shell to most of us without a description of its intended interpretation.* ... [F]ormal specifications ... would be meaningless without a natural language description of the intended usage of the various functions and parameters. On the other hand, we insist that once the reader is familiar with the intended interpretation the specifications should answer all of his [sic] questions about the behavior of the programs without reference to the natural language text.

The italicized sentence in this passage is reminiscent of the Chinese Room. Note, however, that, although we may understand “hollow” code in terms of natural language, Parnas assumes that we understand natural language directly. Or perhaps we only understand natural language via a further semantic interpretation, but then it must be the case that we understand this semantics directly. The sequence of understanding one domain or language in terms of another must stop somewhere with a domain that we understand directly. And this “direct understanding” is what I have called “syntactic understanding” (Rapaport 1986b, 1995, 2000; cf. Harel’s (1998) notion of “symbolic reasoning”).

The third fundamental idea of syntactic semantics is that understanding is recursive—we understand one domain in terms of an antecedently understood one. The base case is a domain that is understood “directly”—i.e., syntactically—in terms of itself. In such domains, some elements of the base domain are understood in terms of other elements of the base domain—i.e., they are understood syntactically. (For a related argument, see Kalderon 2001.) In the case of language, linguistic elements (which are internal) are understood in terms of other internal, but non-linguistic (or “conceptual”), elements. This is how semantics can arise from syntax. And that is how natural-language understanding, by human or computer, is possible.

Appendix

This appendix contains a computer demo of the SNePS-related material discussed in the text, including the application of the rules enabling Cassie to “understand that everything has a name”. It was run using SNePS-2.6.2, and is slightly edited for space considerations and to add a few comments. Added comments are indicated by the word ‘COMMENT’; comments in the original demo are indicated by semicolons at the beginning of a line. The computer prompt is ‘>’; the Lisp prompt consists of a phrase like ‘cl-user(1)’ followed by a colon; the SNePSLOG prompt consists of a colon.

```
Script started on Wed Aug 02 13:50:07 2006
COMMENT: "acl" causes Allegro Common Lisp to be launched
> acl
International Allegro CL Enterprise Edition
8.0 [Solaris] (Jun 30, 2006 12:35)
Copyright (C) 1985-2005, Franz Inc., Oakland, CA, USA.
All Rights Reserved.
```

```
This development copy of Allegro CL is licensed to:
  [4549] University at Buffalo
```

```
COMMENT: We load SNePS-2.6.2:
cl-user(1): :ld /projects/shapiro/Sneps/sneps262
; Loading /projects/shapiro/Sneps/sneps262.cl
; Loading /projects/shapiro/Sneps/Sneps262/load-sneps.lisp
; Loading
; /projects/snwiz/Install/Sneps-2.6.1/load-logical-pathnames.lisp
Loading system SNePS...10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
SNePS-2.6 [PL:1a 2004/08/26 23:05:27] loaded.
Type '(sneps)' or '(snepslog)' to get started.
cl-user(2): (snepslog)
```

```
Welcome to SNePSLOG (A logic interface to SNePS)
```

```
Copyright (C) 1984--2004 by Research Foundation of
State University of New York. SNePS comes with ABSOLUTELY NO WARRANTY!
```

```
: demo "helenkeller.demo"
```

```
File /u0/faculty/rapaport/Papers/HelenKeller/Snepslog/helenkeller.demo
is now the source of input.
```

```
COMMENT: SNePSLOG's "mode 3" allows us to define our own case-frames
;;; use mode 3
set-mode-3
```

```
Net reset
In SNePSLOG Mode 3.
Use define-frame <pred> <list-of-arc-labels>.
```

```

:
;;; define frames

;;; Need a "lex arc" case frame.
;;; [[thing-called(w)]] = concept associated with word w in lexicon
define-frame thing-called (nil lex)
thing-called(x1) will be represented by {<lex, x1>}

:
;;; [[AKO(sub,sup)]] = sub is a subclass of superclass sup
define-frame AKO (nil subclass superclass)
AKO(x1, x2) will be represented by {<subclass, x1>, <superclass, x2>}

:
;;; Figure 1: humans are mammals

;;; This will be:
;;; AKO(thing-called(humans),thing-called(mammals)).
;;; i.e., a thing called 'humans' is a kind of thing called 'mammals'
AKO(thing-called(humans),thing-called(mammals)).

wff3!: AKO(thing-called(humans),thing-called(mammals))

:
;;; describe all relevant nodes (see Figure 1 in text)
%(describe m3)

(m3! (subclass (m1 (lex humans))) (superclass (m2 (lex mammals))))

wff3!: AKO(thing-called(humans),thing-called(mammals))

:
;;;-----
;;; Replace lex arc with "name-object" case frame
;;; -- this is a proposition
;;; Same as "expression-expressed" or "word-object" case frames
;;; -- but uses Helen Keller's terminology

;;; [[ISA-thing-named(o,n)]] = object o is named (expressed)
;;; by the word n in the lexicon

```

```

define-frame ISA-thing-named (nil object name)
ISA-thing-named(x1, x2) will be represented by {<object, x1>, <name, x2>}

:
;;; NB: This is a hard-to-express proposition
;;; -- Any attempt to express it either uses "name",
;;;    which shouldn't really be used, since it's an arc
;;;    or works for some cases but not others (no gen'l sol'n)
;;; -- All it does is provide a name (i.e., a word) for o

;;; Figure 3: humans are mammals
ISA-thing-named(b1, humans).

    wff4!:  ISA-thing-named(b1, humans)

: ISA-thing-named(b2, mammals).

    wff5!:  ISA-thing-named(b2, mammals)

: AKO(b1, b2).

    wff6!:  AKO(b1, b2)

:
;;; describe all relevant nodes (see Figure 3 in text)
%(describe m4 m5 m6)

(m6! (subclass b1) (superclass b2))
(m5! (name mammals) (object b2))
(m4! (name humans) (object b1))

    wff6!:  AKO(b1, b2)
    wff5!:  ISA-thing-named(b2, mammals)
    wff4!:  ISA-thing-named(b1, humans)

:
;;;-----
;;; [[Is(o,p)]] = object o has property p
define-frame Is (nil object property)
Is(x1, x2) will be represented by {<object, x1>, <property, x2>}

```

```

:
;;; Figure 4: water is wet
ISA-thing-named(b3,water).

wff7!: ISA-thing-named(b3,water)

: ISA-thing-named(b4,wet).

wff8!: ISA-thing-named(b4,wet)

: Is(b3,b4).

wff9!: Is(b3,b4)

:
;;; describe all relevant nodes
%(describe m7 m8 m9)

(m9! (object b3) (property b4))
(m8! (name wet) (object b4))
(m7! (name water) (object b3))

wff9!: Is(b3,b4)
wff8!: ISA-thing-named(b4,wet)
wff7!: ISA-thing-named(b3,water)

:
;;;-----
;;; [[Possession(o,p,r)]] = object o is possessor p's r
;;;                               or: p's r is o
;;;                               or: o stands in the r relation to p
define-frame Possession (nil object possessor rel)
Possession(x1, x2, x3) will be represented by
{<object, x1>, <possessor, x2>, <rel, x3>}

:
;;; Figure 5: "water" is b3's name; "wet" is b4's name
ISA-thing-named(b5,name).

```

```

wff10!:  ISA-thing-named(b5,name)

: Possession(water,b3,b5).

wff11!:  Possession(water,b3,b5)

: Possession(wet,b4,b5).

wff12!:  Possession(wet,b4,b5)

:
;;; describe all relevant nodes
%(describe m7 m8 m9 m10 m11 m12)

(m12! (object wet) (possessor b4) (rel b5))
(m11! (object water) (possessor b3) (rel b5))
(m10! (name name) (object b5))
(m9! (object b3) (property b4))
(m8! (name wet) (object b4))
(m7! (name water) (object b3))

wff12!:  Possession(wet,b4,b5)
wff11!:  Possession(water,b3,b5)
wff10!:  ISA-thing-named(b5,name)
wff9!:   Is(b3,b4)
wff8!:   ISA-thing-named(b4,wet)
wff7!:   ISA-thing-named(b3,water)

:
;;;-----
;;; Figure 6: Giving a name to "name".
;;; "name" is b5's name
Possession(name,b5,b5).

wff13!:  Possession(name,b5,b5)

:
;;; describe all relevant nodes
%(describe m10 m13)

```

(m13! (object name) (possessor b5) (rel b5))

(m10! (name name) (object b5))

wff13!:: Possession(name,b5,b5)

wff10!:: ISA-thing-named(b5,name)

:

;;;-----

;;; [[ISA(m,c)]] = m is a member of class c

define-frame ISA (nil member class)

ISA(x1, x2) will be represented by {<member, x1>, <class, x2>}

:

;;; Figure 7: Something is a red, round ball.

;;; Something is a thing named "ball".

ISA-thing-named(b6,ball).

wff14!:: ISA-thing-named(b6,ball)

:

;;; Something is a ball.

ISA(b7,b6).

wff15!:: ISA(b7,b6)

:

;;; Something is a thing named "round"

ISA-thing-named(b8,round).

wff16!:: ISA-thing-named(b8,round)

:

;;; The ball is round.

Is(b7,b8).

wff17!:: Is(b7,b8)

:

;;; Something is a thing named "red"

```

ISA-thing-named(b9,red).

    wff18!:  ISA-thing-named(b9,red)

:
;;; The ball (which is round) is red.
Is(b7,b9).

    wff19!:  Is(b7,b9)

:
;;; describe all relevant nodes
%(describe m14 m15 m16 m17 m18 m19)

(m19! (object b7) (property b9))
(m18! (name red) (object b9))
(m17! (object b7) (property b8))
(m16! (name round) (object b8))
(m15! (class b6) (member b7))
(m14! (name ball) (object b6))

    wff19!:  Is(b7,b9)
    wff18!:  ISA-thing-named(b9,red)
    wff17!:  Is(b7,b8)
    wff16!:  ISA-thing-named(b8,round)
    wff15!:  ISA(b7,b6)
    wff14!:  ISA-thing-named(b6,ball)

:
;;;-----
;;; Figure 8:  "ball" is b6's name;
;;;           "round" is b8's name;
;;;           "red" is b9's name

Possession(ball,b6,b5).

    wff20!:  Possession(ball,b6,b5)

: Possession(round,b8,b5).

```

```

wff21!: Possession(round,b8,b5)

: Possession(red,b9,b5).

wff22!: Possession(red,b9,b5)

:
;;; describe all relevant nodes
%(describe m10 m20 m21 m22)

(m22! (object red) (possessor b9) (rel b5))
(m21! (object round) (possessor b8) (rel b5))
(m20! (object ball) (possessor b6) (rel b5))
(m10! (name name) (object b5))

wff22!: Possession(red,b9,b5)
wff21!: Possession(round,b8,b5)
wff20!: Possession(ball,b6,b5)
wff10!: ISA-thing-named(b5,name)

:
;;;-----
;;; If Cassie/Keller is to learn the name of naming,
;;; then she needs a rule like this:

;;; Rule (1):
;;; If x is named y, then ..?.. & y is x's name
;;; -- where "x is named y" is the hard-to-express proposition providing
;;; a way of uttering x, and where "..?.." is the hard-to-express
;;; proposition providing Cassie/Keller with a name for naming

;;; So we assert this rule and do forward inference with it
;;; (indicated by the exclamation mark punctuation):

all(x,y)(ISA-thing-named(x,y) => {ISA-thing-named(b5,name),
                                Possession(y,x,b5)}!)

I wonder if p1: ISA-thing-named(x,y)
holds within the BS defined by context default-defaultct

```

```

I know wff4!: ISA-thing-named(b1,humans)

Since wff23!: all(y,x)({ISA-thing-named(x,y)}
                    v=> {Possession(y,x,b5),ISA-thing-named(b5,name)})
and wff4!: ISA-thing-named(b1,humans)
I infer wff24: Possession(humans,b1,b5)

COMMENT: The rest of the "trace" of the inference, part of which is
        shown above, is edited out to save space, as are all
        subsequent inferenceinference traces.

:
;;; So, she now believes wff25 and wff24:
%(describe m24 m25)

(m25! (object mammals) (possessor b2) (rel b5))
(m24! (object humans) (possessor b1) (rel b5))

wff25!: Possession(mammals,b2,b5)
wff24!: Possession(humans,b1,b5)

:
;;; And to remind you what b1, b2, b5 were:
%(describe m4 m5 m10 m24 m25)

(m25! (object mammals) (possessor b2) (rel b5))
(m24! (object humans) (possessor b1) (rel b5))
(m10! (name name) (object b5))
(m5! (name mammals) (object b2))
(m4! (name humans) (object b1))

wff25!: Possession(mammals,b2,b5)
wff24!: Possession(humans,b1,b5)
wff10!: ISA-thing-named(b5,name)
wff5!: ISA-thing-named(b2,mammals)
wff4!: ISA-thing-named(b1,humans)

:
;;; And to remind you how they are related:
%(describe m4 m5 m6 m10 m24 m25)

```

```

(m25! (object mammals) (possessor b2) (rel b5))
(m24! (object humans) (possessor b1) (rel b5))
(m10! (name name) (object b5))
(m6! (subclass b1) (superclass b2))
(m5! (name mammals) (object b2))
(m4! (name humans) (object b1))

wff25!:: Possession(mammals,b2,b5)
wff24!:: Possession(humans,b1,b5)
wff10!:: ISA-thing-named(b5,name)
wff6!:: AKO(b1,b2)
wff5!:: ISA-thing-named(b2,mammals)
wff4!:: ISA-thing-named(b1,humans)

:
;;; And, to learn the name and concept of "property"
;;; she'd need a rule like this:

;;; Rule (2):
;;; If x is (i.e., has property) y,
;;; then ..?.. & y is a property
;;; -- This provides a concept of property and a name for it
all(x,y)(Is(x,y) => {ISA-thing-named(b10,property), ISA(y,b10)})!

:
;;; She now believes:
%(describe m28 m29 m30 m31)

(m31! (class b10) (member b9))
(m30! (class b10) (member b8))
(m29! (object property) (possessor b10) (rel b5))
(m28! (class b10) (member b4))

wff31!:: ISA(b9,b10)
wff30!:: ISA(b8,b10)
wff29!:: Possession(property,b10,b5)
wff28!:: ISA(b4,b10)

:

```

```

;;;-----
;;; Learning the name for the concept "name":

;;; Need a rule like this:
;;; If x "is a thing named" y, then y is x's something-or-other
;;; NB: "something-or-other" is a base node with an English-like
;;;      node identifier, which, by McDermott's Fallacy, has no
;;;      significance whatever (except some convenience to us).

;;; Here, we should pretend that this rule is applied *before*
;;; Cassie learns the name of naming. Once the rule is applied,
;;; then we can tell her that things have names, and she can
;;; infer that her unknown "something-or-other" must be the
;;; relation of naming. However, to be consistent with the
;;; order in which the material is presented in the text,
;;; we are doing this demo "out of order", so to speak.

;;; We assert and do forward inference with:

all(x,y)(ISA-thing-named(x,y) => Possession(y,x,something-or-other))!

:
;;; Now she believes:
%(describe m33 m34 m35 m36 m37 m38 m39 m40 m41)

(m41! (object property) (possessor b10) (rel something-or-other))
(m40! (object red) (possessor b9) (rel something-or-other))
(m39! (object round) (possessor b8) (rel something-or-other))
(m38! (object ball) (possessor b6) (rel something-or-other))
(m37! (object name) (possessor b5) (rel something-or-other))
(m36! (object wet) (possessor b4) (rel something-or-other))
(m35! (object water) (possessor b3) (rel something-or-other))
(m34! (object mammals) (possessor b2) (rel something-or-other))
(m33! (object humans) (possessor b1) (rel something-or-other))

wff41!: Possession(property,b10,something-or-other)
wff40!: Possession(red,b9,something-or-other)
wff39!: Possession(round,b8,something-or-other)
wff38!: Possession(ball,b6,something-or-other)
wff37!: Possession(name,b5,something-or-other)

```

```

wff36!: Possession(wet,b4,something-or-other)
wff35!: Possession(water,b3,something-or-other)
wff34!: Possession(mammals,b2,something-or-other)
wff33!: Possession(humans,b1,something-or-other)

:
;;;-----
;;; To figure out what a "something-or-other" is (or: what its name is?)
;;; need a "defeasible" rule like this:

;;; all(x,y,z,r)({Is(z,unknown),Possession(x,y,z),Possession(x,y,r)}
;;;               => Equivalent({z,r}))

;;; where
;;; [[Is(z,unknown)]] = z has the property of being unknown
;;;                   (e.g., by Cassie)

;;; This is probably not the best way to represent a concept being
;;; unknown or unfamiliar, but it will suffice for present purposes

;;; We also need to define:
define-frame Equivalent (nil equiv)
Equivalent(x1) will be represented by {<equiv, x1>}

:
;;; where:
;;; [[Equivalent({x1,...,xn})]] = mental entities x1, ..., xn
;;;                               correspond to the same actual object.

;;; And we have to assert that "something-or-other" is unknown:
Is(something-or-other,unknown).

wff42!: Is(something-or-other,unknown)

:
;;; Now we can assert and do forward inference with:

all(x,y,z,r)({Is(z,unknown),Possession(x,y,z),Possession(x,y,r)}
              => Equivalent({z,r}))!

```

:
COMMENT: Now, Cassie has come to believe (i.e., she has "learned"):
%(describe m10 m13 m42 m43 m44)

(m44! (equiv something-or-other b5))
(m43! (forall v10 v9 v8 v7)
(&ant (p9 (object v7) (possessor v8) (rel v10))
(p8 (object v7) (possessor v8) (rel v9))
(p7 (object v9) (property unknown)))
(cq (p10 (equiv v10 v9))))
(m42! (object something-or-other) (property unknown))
(m13! (object name) (possessor b5) (rel b5))
(m10! (name name) (object b5))

wff44!: Equivalent({something-or-other,b5})
wff43!: all(r,z,y,x)({Possession(x,y,r),Possession(x,y,z),
Is(z,unknown)} &=> {Equivalent({r,z})})
wff42!: Is(something-or-other,unknown)
wff13!: Possession(name,b5,b5)
wff10!: ISA-thing-named(b5,name)

:
End of
/u0/faculty/rapaport/Papers/HelenKeller/Snepslog/helenkeller.demo
demonstration.

script done on Wed Aug 02 13:50:56 2006

Acknowledgments

An ancestor of this essay was originally written around 1992, was discussed in my seminar in Spring 1993 on “Semantics, Computation, and Cognition” at SUNY Buffalo, and first appeared in an unpublished technical report (Rapaport 1996). I am grateful to Albert Goldfain, Frances L. Johnson, David Pierce, Stuart C. Shapiro, and the other members of the SNePS Research Group for comments.

References

- Allaire, Edwin B. (1963), "Bare Particulars", *Philosophical Studies* 14; reprinted in Michael J. Loux (ed.), *Universals and Particulars: Readings in Ontology* (Garden City, NY: Anchor Books, 1970): 235–244.
- Allaire, Edwin B. (1965), "Another Look at Bare Particulars", *Philosophical Studies* 16; reprinted in Michael J. Loux (ed.), *Universals and Particulars: Readings in Ontology* (Garden City, NY: Anchor Books, 1970): 250–257.
- Allen, James F., & Perrault, C. Raymond (1980), "Analyzing Intentions in Utterance", *Artificial Intelligence* 15: 143–17; reprinted in Barbara J. Grosz, Karen Sparck Jones, & Bonnie L. Webber (eds.), *Readings in Natural Language Processing* (Los Altos, CA: Morgan Kaufmann, 1986): 441–458.
- Anderson, Michael L. (2003), "Embodied Cognition: A Field Guide", *Artificial Intelligence* 149(1): 91–130.
- Arahi, Kenji, & Momouchi, Yoshio (1990), "Learning of Semantic Concept in Copular Sentence" (in Japanese), *IPSJ SIG Reports*, Vol. 90, No. 77.
- Arrighi, Claudia, & Ferrario, Roberta (2005), "The Dynamic Nature of Meaning", in Lorenzo Magnani & Riccardo Dossena (eds.), *Computing, Philosophy, and Cognition: Proceedings of the European Computing and Philosophy Conference (ECAP 2004)* (London: [King's] College Publications): 295–312.
- Baker, Robert (1967), "Particulars: Bare, Naked, and Nude", *Noûs* 1(2): 211–212.
- Berners-Lee, Tim, & Fischetti, Mark (1999), *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor* (New York: HarperCollins).
- Berners-Lee, Tim; Hendler, James; & Lassila, Ora (2001), "The Semantic Web", *Scientific American* (17 May); online at:
[http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21]
- Brown, Roger (1973), *A First Language: The Early Stages* (Cambridge, MA: Harvard University Press).
- Bruce, Bertram C. (1975), "Generation as a Social Action", *Theoretical Issues in Natural Language Processing-1* (Morristown, NJ: Association for Computational Linguistics): 64–67; reprinted in Barbara J. Grosz, Karen Sparck Jones, & Bonnie L. Webber (eds.), *Readings in Natural Language Processing* (Los Altos, CA: Morgan Kaufmann, 196): 419–422.
- 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).
- Bruner, Jerome (1983), *Child's Talk: Learning to Use Language* (New York: W.W. Norton).
- Castañeda, Hector-Neri (1980), *On Philosophical Method* (Bloomington, IN: Noûs Publications); ISBN 0-96041-86-0-1.
- Castañeda, Hector-Neri (1984), "Philosophical Refutations", in J.H. Fetzer (ed.), *Principles of Philosophical Reasoning* (Totowa, NJ: Rowman & Allenheld): 227–258.

- Ceusters, Werner (2005, October 27), "Ontology: The Need for International Coordination" [<http://ncor.buffalo.edu/inaugural/ppt/ceusters.ppt>].
- Chappell, V.C. (1964), "Particulars Re-Clothed", *Philosophical Studies* 15; reprinted in Michael J. Loux (ed.), *Universals and Particulars: Readings in Ontology* (Garden City, NY: Anchor Books, 1970): 245–249.
- Chrisley, Ron (2003), "Embodied Artificial Intelligence", *Artificial Intelligence* 149(1): 131–150.
- Chun, Soon Ae (1987), "SNePS Implementation of Possessive Phrases", *SNeRG Technical Note 19* (Buffalo: SUNY Buffalo Department of Computer Science) [<http://www.cse.buffalo.edu/sneps/Bibliography/chun87.pdf>].
- Clark, Andy, & Chalmers, David J. (1998), "The Extended Mind", *Analysis* 58: 10–23.
- Cohen, Philip R., & Levesque, Hector J. (1985), "Speech Acts and Rationality", *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (University of Chicago)* (Morristown, NJ: Association for Computational Linguistics): 49–60.
- Cohen, Philip R., & Levesque, Hector J. (1990), "Rational Interaction as the Basis for Communication", in Philip R. Cohen, Jerry Morgan, & Martha E. Pollack (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press): 221–256.
- Cohen, Philip R., & Perrault, C. Raymond (1979), "Elements of a Plan-Based Theory of Speech Acts", *Cognitive Science* 3: 177–212; reprinted in Barbara J. Grosz, Karen Sparck Jones, & Bonnie L. Webber (eds.), *Readings in Natural Language Processing* (Los Altos, CA: Morgan Kaufmann, 196): 423–440.
- Cullingford, Richard (1981), "SAM", in Roger C. Schank & Christopher K. Riesbeck (eds.), *Inside Computer Understanding: Five Programs Plus Miniatures* (Hillsdale, NJ: Lawrence Erlbaum Associates): 75–119.
- Damasio, Antonio R. (1989), "Concepts in the Brain", in "Forum: What is a Concept?", *Mind and Language* 4: 24–27.
- Davidson, Donald (1967), "The Logical Form of Action Sentences", in Nicholas Rescher (ed.), *The Logic of Decision and Action* (Pittsburgh: University of Pittsburgh Press); reprinted in Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980): 105–148.
- Dehaene, Stanislas (1992), "Varieties of Numerical Abilities", *Cognition* 44: 1–42.
- Dresner, Eli (2002), "Holism, Language Acquisition, and Algebraic Logic", *Linguistics and Philosophy* 25: 419–452.
- Dreyfus, Hubert L. (1992), *What Computers Still Can't Do: A Critique of Artificial Reason* (Cambridge, MA: MIT Press).
- Duchan, Judith F.; Bruder, Gail A.; & Hewitt, Lynne E. (eds.) (1995), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates).
- Ehrlich, Karen (1995), "Automatic Vocabulary Expansion through Narrative Context", *Technical Report 95-09* (Buffalo: SUNY Buffalo Department of Computer Science).
- Ehrlich, Karen (2004), "Default Reasoning Using Monotonic Logic: Nutter's Modest Proposal Revisited, Revised, and Implemented", *Proceedings of the 15th Midwest Artificial Intelligence and Cognitive Science Conference (MAICS 2004, Roosevelt University)*: 4–54 [<http://www.cse.buffalo.edu/~rapaport/CVA/ehrllich-maics-sub.pdf>].

- Ehrlich, Karen, & Rapaport, William J. (1997), "A Computational Theory of Vocabulary Expansion", *Proceedings of the 19th Annual Conference of the Cognitive Science Society (Stanford University)* (Mahwah, NJ: Lawrence Erlbaum Associates): 205–210.
- Ehrlich, Karen, & Rapaport, William J. (2004), "A Cycle of Learning: Human and Artificial Contextual Vocabulary Acquisition", *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (Mahwah, NJ: Lawrence Erlbaum Associates, 2005): 1555.
- Elgin, Suzette Haden (1984), *Native Tongue* (New York: DAW Books).
- Fitch, W. Tecumseh (2006), "Hypothetically Speaking", *American Scientist* (July-August): 369–370.
- Galbraith, Mary, & Rapaport, William J. (guest eds.) (1995), *Where Does I Come From? Special Issue on Subjectivity and the Debate over Computational Cognitive Science, Minds and Machines* 5(4): 513–620.
- Geach, Peter T. (1967), "Intentional Identity", *Journal of Philosophy* 64(20): 627–632.
- Giere, Ronald N. (2002), "Distributed Cognition in Epistemic Cultures", *Philosophy of Science* 69: 637–644.
- Goldfain, Albert (2004), "Using SNePS for Mathematical Cognition: A SNeRE-Based Natural-Language Algorithm for Computing GCD" [<http://www.cse.buffalo.edu/~rapaport/CVA/GCD/ag33-ProgressReport.pdf>].
- Goldfain, Albert (2006), "A Computational Theory of Early Mathematical Cognition" [<http://www.cse.buffalo.edu/~ag33/DissertationProposal.pdf>].
- Greenfield, Patricia Marks, & Savage-Rumbaugh, E. Sue (1990), "Grammatical Combination in *Pan Paniscus*: Processes of Learning and Invention in the Evolution and Development of Language", in S.T. Parker & K.R. Gibson (eds.), *Language and Intellect in Monkeys and Apes* (Cambridge, UK: Cambridge University Press): 540–577.
- Grosz, Barbara J., & Sidner, Candace L. (1986), "Attention, Intentions, and the Structure of Discourse", *Computational Linguistics* 12:175–204.
- Haller, Susan M. (1994), "Interactive Generation of Plan Descriptions and Justifications", *Technical Report 94-40* (Buffalo, NY: SUNY Buffalo Department of Computer Science).
- Haller, Susan M. (1995), "Planning Text for Interactive Plan Explanations", in E.A. Yfantis (ed.), *Intelligent Systems: 3rd Golden West International Conference* (Dordrecht, Holland: Kluwer Academic Publishers): 61–67.
- Harel, Guershon (1998), "Two Dual Assertions: The First on Learning and the Second on Teaching (or Vice Versa)", *American Mathematical Monthly* 105(6) (June-July): 497–507.
- Harman, Gilbert (1987), "(Nonsolipsistic) Conceptual Role Semantics", in Ernest Lepore (ed.), *New Directions in Semantics* (London: Academic Press): 55–81.
- Harnad, Stevan (1990), "The Symbol Grounding Problem", *Physica D* 42: 335–346.
- Hofstadter, Douglas R. (2001), "Epilogue: Analogy as the Core of Cognition", in Dedre Gentner, Keith J. Holyoak, & Boicho N. Kokinov (eds.), *The Analogical Mind: Perspectives from Cognitive Science* (Cambridge, MA: MIT Press): 499–538.
- Hutchins, Edwin (1995a), *Cognition in the Wild* (Cambridge, MA: MIT Press).

- Hutchins, Edwin (1995b), "How a Cockpit Remembers Its Speeds", *Cognitive Science* 19: 265–288.
- Ismail, Haythem O.; & Shapiro, Stuart C. (2000), "Two Problems with Reasoning and Acting in Time", in A.G. Cohn, F. Giunchiglia, & B. Selman (eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the 7th International Conference (KR 2000)* (San Francisco: Morgan Kaufmann): 355–365.
- Jackendoff, Ray (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution* (Oxford: Oxford University Press).
- Kalderon, Mark Eli (2001), "Reasoning and Representing", *Philosophical Studies* 105: 129–160.
- 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).
- Keller, Helen (1903), *Optimism: An Essay* (New York: Crowell).
- Keller, Helen (1905), *The Story of My Life* (Garden City, NY: Doubleday, 1954).
- Kibby, Michael W.; Rapaport, William J.; Wieland, Karen M.; & Dechert, Debra A. (forthcoming), "CSI: Contextual Semantic Investigation for Word Meaning", in Lawrence A. Baines (ed.), *Multisensory Learning* (tentative title); [<http://www.cse.buffalo.edu/~rapaport/CVA/CSI.pdf>].
- Lammens, Johan (1994), "A Computational Model of Color Perception and Color Naming", *Technical Report 94-26* (Buffalo: SUNY Buffalo Department of Computer Science) [<http://www.cse.buffalo.edu/sneps/Bibliography/lammens94.pdf>].
- Leiber, Justin (1996), "Helen Keller as Cognitive Scientist", *Philosophical Psychology* 9(4): 419–440.
- Maida, Anthony S., & Shapiro, Stuart C. (1982), "Intensional Concepts in Propositional Semantic Networks", *Cognitive Science* 6: 291–330; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 169–189.
- Maloney, J. Christopher (1987), "The Right Stuff," *Synthese* 70: 349–372.
- Maloney, J. Christopher (1989), *The Mundane Matter of the Mental Language* (Cambridge, UK: Cambridge University Press).
- Martins, João, & Shapiro, Stuart C. (1988), "A Model for Belief Revision", *Artificial Intelligence* 35: 25–79.
- Mayes, A.R. (1991), Review of Hanna Damasio & Antonio R. Damasio, *Lesion Analysis in Neuropsychology (inter alia)*, *British Journal of Psychology* 2: 109–112.
- McDermott, Drew (1981), "Artificial Intelligence Meets Natural Stupidity", in John Haugeland (ed.), *Mind Design: Philosophy, Psychology, Artificial Intelligence* (Cambridge, MA: MIT Press): 143–160.
- Meinong, Alexius (1904), "Über Gegenstandstheorie", in Rudolf Haller (ed.), *Alexius Meinong Gesamtausgabe*, Vol. II (Graz, Austria: Akademische Druck- und Verlagsanstalt, 1971): 481–535; English translation ("The Theory of Objects") by Isaac Levi et al., in Roderick M. Chisholm (ed.), *Realism and the Background of Phenomenology* (New York: Free Press, 1960): 76–117.
- Miles, H. Lyn White (1990), "The Cognitive Foundations for Reference in a Signing Orangutan", in S.T. Parker & K.R. Gibson (eds.), *'Language' and Intellect in*

- Monkeys and Apes* (Cambridge, UK: Cambridge University Press): 511–539.
- Morris, Charles (1938), *Foundations of the Theory of Signs* (Chicago: University of Chicago Press).
- Nagel, Thomas (1986), *The View from Nowhere* (New York: Oxford University Press).
- Neal, Jeanette G., & Shapiro, Stuart C. (1987), “Knowledge-Based Parsing”, in L. Bolc (ed.), *Natural Language Parsing Systems* (Berlin: Springer-Verlag): 49–92.
- Papineau, David (1998), “Get a Grip”, review of Wilson 1998, *The New York Times Book Review* (19 July): 9;
[<http://www.nytimes.com/books/98/07/19/reviews/980719.19papinet.html>]
- Parnas, David L. (1972), “A Technique for Software Module Specification with Examples”, *Communications of the ACM* 15(5) (May): 330–336.
- Proudfoot, Diane (2002), “Wittgenstein’s Anticipation of the Chinese Room”, in John Preston & Mark Bishop (eds.), *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (Oxford: Oxford University Press): 167–180.
- Putnam, Hilary (1975), “The Meaning of ‘Meaning’ ”, reprinted in *Mind, Language and Reality* (Cambridge, UK: Cambridge University Press): 215–271.
- Quine, Willard Van Orman (1953), “On What There Is” and “Logic and the Reification of Universals”, in *From a Logical Point of View: 9 Logico-Philosophical Essays; 2nd Edition, Revised* (New York: Harper & Row), Chs. I, VI.
- Quine, Willard Van Orman (1960), *Word and Object* (Cambridge, MA: MIT Press).
- Quine, Willard Van Orman (1969), “Ontological Relativity”, in W.V. Quine, *Ontological Relativity and Other Essays* (New York: Columbia University Press): 26–68.
- Rapaport, William J. (1976), *Intentionality and the Structure of Existence*, Ph.D. dissertation (Bloomington: Indiana University Department of Philosophy).
- Rapaport, William J. (1978), “Meinongian Theories and a Russellian Paradox”, *Noûs* 12: 153–10; errata, *Noûs* 13 (1979) 125.
- Rapaport, William J. (1981), “How to Make the World Fit Our Language: An Essay in Meinongian Semantics”, *Grazer Philosophische Studien* 14: 1–21.
- Rapaport, William J. (1982), “Unsolvable Problems and Philosophical Progress”, *American Philosophical Quarterly* 19: 289–298.
- Rapaport, William J. (1985), “Machine Understanding and Data Abstraction in Searle’s Chinese Room”, *Proceedings of the 7th Annual Meeting of the Cognitive Science Society (University of California at Irvine)* (Hillsdale, NJ: Lawrence Erlbaum Associates): 341–345.
- Rapaport, William J. (1985/1986), “Non-Existent Objects and Epistemological Ontology”, *Grazer Philosophische Studien* 25/26: 61–95; reprinted in Rudolf Haller (ed.), *Non-Existence and Predication* (Amsterdam: Rodopi, 1986).
- Rapaport, William J. (1986a), “Logical Foundations for Belief Representation”, *Cognitive Science* 10: 371–422.
- Rapaport, William J. (1986b), “Searle’s Experiments with Thought”, *Philosophy of Science* 53: 271–279.
- Rapaport, William J. (1988a), “To Think or Not to Think” *Noûs* 22: 585–609.
- Rapaport, William J. (1988b), “Syntactic Semantics: Foundations of Computational Natural-Language Understanding”, in James H. Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht, Holland: Kluwer Academic Publishers): 1–131; reprinted

- in Eric Dietrich (ed.), *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines* (San Diego: Academic Press, 1994): 225–273.
- Rapaport, William J. (1990), “Computer Processes and Virtual Persons: Comments on Cole’s ‘Artificial Intelligence and Personal Identity’ ”, *Technical Report 90-13* (Buffalo: SUNY Buffalo Department of Computer Science, May 1990); [<http://www.cse.buffalo.edu/~rapaport/Papers/cole.tr.17my90.pdf>].
- Rapaport, William J. (1991a), “Predication, Fiction, and Artificial Intelligence”, *Topoi* 10: 79–111.
- Rapaport, William J. (1991b), “Meinong, Alexius I: Meinongian Semantics”, in Hans Burkhardt & Barry Smith (eds.), *Handbook of Metaphysics and Ontology* (Munich: Philosophia Verlag): 516–519.
- Rapaport, William J. (1993), “Because Mere Calculating Isn’t Thinking: Comments on Hauser’s ‘Why Isn’t My Pocket Calculator a Thinking Thing?’ ”, *Minds and Machines* 3: 11–20.
- Rapaport, William J. (1995), “Understanding Understanding: Syntactic Semantics and Computational Cognition”, in James E. Tomberlin (ed.), *Philosophical Perspectives, Vol. 9: AI, Connectionism, and Philosophical Psychology* (Atascadero, CA: Ridgeview): 49–88; reprinted in Josefa Toribio & Andy Clark eds. (1998), *Language and Meaning in Cognitive Science: Cognitive Issues and Semantic Theory, Artificial Intelligence and Cognitive Science: Conceptual Issues*, Vol. 4 (New York: Garland): 73–88.
- Rapaport, William J. (1996), *Understanding Understanding: Semantics, Computation, and Cognition, Technical Report 96-26* (Buffalo: SUNY Buffalo Department of Computer Science); [<http://www.cse.buffalo.edu/tech-reports/96-26.ps>].
- Rapaport, William J. (1998), “How Minds Can Be Computational Systems”, *Journal of Experimental and Theoretical Artificial Intelligence* 10: 403–419.
- Rapaport, William J. (1999), “Implementation Is Semantic Interpretation”, *The Monist* 2: 109–130.
- Rapaport, William J. (2000), “How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition”, *Journal of Logic, Language, and Information*, 9(4): 467–490; reprinted in James H. Moor (ed.), *The Turing Test: The Elusive Standard of Artificial Intelligence* (Dordrecht: Kluwer, 2003): 161–14.
- Rapaport, William J. (2002), “Holism, Conceptual-Role Semantics, and Syntactic Semantics”, *Minds and Machines* 12(1): 3–59.
- Rapaport, William J. (2003a), “What Did You Mean by That? Misunderstanding, Negotiation, and Syntactic Semantics”, *Minds and Machines* 13(3): 397–427.
- Rapaport, William J. (2003b), “What Is the ‘Context’ for Contextual Vocabulary Acquisition?”, in Peter P. Slezak (ed.), *Proceedings of the 4th International Conference on Cognitive Science/7th Australasian Society for Cognitive Science Conference (ICCS/ASCS-2003; Sydney, Australia)* (Sydney: University of New South Wales), Vol. 2, pp. 547–552.
- Rapaport, William J. (2005a), “In Defense of Contextual Vocabulary Acquisition: How to Do Things with Words in Context”, in A. Dey et al. (eds.), *Proceedings of the 5th International and Interdisciplinary Conference on Modeling and Using Context (Context-05)* (Berlin: Springer-Verlag Lecture Notes in Artificial Intelligence 3554):

396–409.

- Rapaport, William J. (2005b), “Implementation Is Semantic Interpretation: Further Thoughts”, *Journal of Experimental and Theoretical Artificial Intelligence* 17(4): 385–417.
- Rapaport, William J. (2005c), Review of Shieber’s (2004) *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*, in *Computational Linguistics* 31(3): 407–412.
- Rapaport, William J. (2005d), “The Turing Test”, in *Encyclopedia of Language and Linguistics, 2nd Edition* (Oxford: Elsevier), Vol. 13, pp. 151–159.
- Rapaport, William J. (2006), Review of John Preston & Mark Bishop (eds.), *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, in *Australasian Journal of Philosophy* 84(1) (March): 129–133.
- Rapaport, William J., & Ehrlich, Karen (2000), “A Computational Theory of Vocabulary Acquisition”, in Łucja M. Iwańska & Stuart C. Shapiro (eds.), *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (Menlo Park, CA/Cambridge, MA: AAAI Press/MIT Press): 347–375.
- Rapaport, William J., & Kibby, Michael W. (2002), “Contextual Vocabulary Acquisition: A Computational Theory and Educational Curriculum”, in Nagib Callaos, Ana Breda, and Ma. Yolanda Fernandez J. (eds.), *Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2002; Orlando, FL)* (Orlando: International Institute of Informatics and Systemics), Vol. II: Concepts and Applications of Systemics, Cybernetics, and Informatics I, pp. 261–266.
- Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; Zubin, David A.; Bruder, Gail A.; Duchan, Judith F.; Almeida, Michael J.; Daniels, Joyce H.; Galbraith, Mary M.; Wiebe, Janyce M.; & Yuhan, Albert Hanyong (1989), “Deictic Centers and the Cognitive Structure of Narrative Comprehension”, *Technical Report 89-01* (Buffalo: SUNY Buffalo Department of Computer Science).
- Rapaport, William J.; Segal, Erwin M.; Shapiro, Stuart C.; Zubin, David A.; Bruder, Gail A.; Duchan, Judith F.; & Mark, David M. (1989), “Cognitive and Computer Systems for Understanding Narrative Text”, *Technical Report 89-07* (Buffalo: SUNY Buffalo Department of Computer Science).
- 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); revised version published as Rapaport, Shapiro, & Wiebe 1997.
- Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1997), “Quasi-Indexicals and Knowledge Reports”, *Cognitive Science* 21: 63–107; reprinted in Francesco Orilia & William J. Rapaport (eds.), *Thought, Language, and Ontology: Essays in Memory of Hector-Neri Castañeda* (Dordrecht: Kluwer Academic Publishers, 1998): 235–294.
- Rogers, Hartley, Jr. (1959), “The Present Theory of Turing Machine Computability”, *Journal of the Society for Industrial and Applied Mathematics* 7(1): 114–130.
- Santore, John F.; & Shapiro, Stuart C. (2004), “Identifying Perceptually Indistinguishable Objects”, in S. Coradeschi & A. Saffiotti (eds.), *Anchoring Symbols to Sensor Data, Papers from the AAAI Workshop, Technical Report WS-04-03* (Menlo Park, CA:

- AAAI Press): 1–9.
- Searle, John R. (1980), “Minds, Brains, and Programs”, *Behavioral and Brain Sciences* 3: 417–457.
- Searle, John R. (1993), “The Failures of Computationalism”, *Think* (Tilburg, The Netherlands: Tilburg University Institute for Language Technology and Artificial Intelligence) 2 (June): 68–71.
- Searle, John R. (2002), “Twenty-One Years in the Chinese Room”, in John Preston & Mark Bishop (eds.), *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (Oxford: Oxford University Press): 51–69.
- Sellars, Wilfrid (1963), *Science, Perception and Reality* (London: Routledge & Kegan Paul).
- Severo, Richard (1999), “Clifton Fadiman, a Wordsmith Known for His Encyclopedic Knowledge, Is Dead at 95”, *The New York Times* (21 June): B5; abstract online at: [<http://select.nytimes.com/gst/abstract.html?res=F20A15FF3B5F0C728EDDAF0894D1494D81>]
- Shapiro, Stuart C. (1979), “The SNePS Semantic Network Processing System”, in Nicholas Findler (ed.), *Associative Networks: Representation and Use of Knowledge by Computers* (New York: Academic Press): 179–203.
- Shapiro, Stuart C. (1981), “COCCI: A Deductive Semantic Network Program for Solving Microbiology Unknowns”, *Technical Report 173* (Buffalo: SUNY Buffalo Department of Computer Science).
- Shapiro, Stuart C. (1982), “Generalized Augmented Transition Network Grammars for Generation from Semantic Networks”, *American Journal of Computational Linguistics* 8: 12–25.
- Shapiro, Stuart C. (1986), “Symmetric Relations, Intensional Individuals, and Variable Binding”, *Proceedings of the IEEE* 74(10): 1354–1363.
- Shapiro, Stuart C. (1989), “The CASSIE Projects: An Approach to Natural Language Competence”, in João P. Martins & Ernesto M. Morgado (eds.), *EPIA 89: 4th Portuguese Conference on Artificial Intelligence, Proceedings (Lisbon)* (Berlin: Springer-Verlag Lecture Notes in Artificial Intelligence 390): 362–380.
- Shapiro, Stuart C. (1992), “Artificial Intelligence”, in Stuart C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence, 2nd Edition* (New York: John Wiley & Sons): 54–57.
- Shapiro, Stuart C. (1998), “Embodied Cassie”, *Cognitive Robotics: Papers from the 1998 AAAI Fall Symposium, Technical Report FS-98-02* Menlo Park, CA: AAAI Press): 136–143.
- Shapiro, Stuart C. (2000), “SNePS: A Logic for Natural Language Understanding and Commonsense Reasoning”, in Łucja M. Iwańska & Stuart C. Shapiro (eds.), *Natural Language Processing and Knowledge Representation: Language for Knowledge and Knowledge for Language* (Menlo Park, CA/Cambridge, MA: AAAI Press/MIT Press): 175–195.
- Shapiro, Stuart C. (2003), “FevahrCassie: A Description and Notes for Building FevahrCassie-Like Agents”, *SNeRG Technical Note 35* (Buffalo, NY: SUNY Buffalo Department of Computer Science and Engineering) [<http://www.cse.buffalo.edu/~shapiro/Papers/buildingFevahrAgents.pdf>].
- Shapiro, Stuart C., & Ismail, Haythem O. (2003), “Anchoring in a Grounded Layered Architecture with Integrated Reasoning”, *Robotics and Autonomous Systems* 43: 97–

108.

- Shapiro, Stuart C.; Ismail, Haythem O.; & Santore, John F. (2000), "Our Dinner with Cassie", *Working Notes for the AAAI 2000 Spring Symposium on Natural Dialogues with Practical Robotic Devices* (Menlo Park, CA: AAAI Press): 57–61.
- Shapiro, Stuart C., & Rapaport, William J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network", in Nick Cercone & Gordon McCalla (eds.), *The Knowledge Frontier: Essays in the Representation of Knowledge* (New York: Springer-Verlag): 262–315; shorter version appeared in *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86, Philadelphia)* (Los Altos, CA: Morgan Kaufmann): 278–283; a revised shorter version appears as "A Fully Intensional Propositional Semantic Network", in Leslie Burkholder (ed.), *Philosophy and the Computer* (Boulder, CO: Westview Press, 1992): 75–91.
- Shapiro, Stuart C., & Rapaport, William J. (1991), "Models and Minds: Knowledge Representation for Natural-Language Competence", in Robert Cummins & John Pollock (eds.), *Philosophy and AI: Essays at the Interface* (Cambridge, MA: MIT Press): 215–259.
- Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family", *Computers and Mathematics with Applications* 23: 243–275; reprinted in F. Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 243–275.
- Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narratives", in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.
- Shapiro, Stuart C.; Rapaport, William J.; Cho, Sung-Hye; Choi, Joongmin; Feit, Elissa; Haller, Susan; Kankiewicz, Jason; & Kumar, Deepak (1996), "A Dictionary of SNePS Case Frames" [<http://www.cse.buffalo.edu/sneps/Manuals/dictionary.pdf>].
- Shapiro, Stuart C., & the SNePS Research Group (2006), "SNePS" [<http://en.wikipedia.org/wiki/SNePS>].
- Sheckley, Robert (1954), "Ritual", in Robert Sheckley, *Untouched by Human Hands* (New York: Ballantine Books): 155–165.
- Shieber, Stuart M. (2004), *The Turing Test: Verbal Behavior as the Hallmark of Intelligence* (Cambridge, MA: MIT Press).
- 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.
- Spärck Jones, Karen (1967), "Dictionary Circles", *Technical Memorandum TM-3304* (Santa Monica, CA: System Development Corp.)
- 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)* (Menlo Park, CA: AAAI Press/MIT Press): 80–85.
- Srihari, Rohini K. (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)*.
- Srihari, Rohini K. (1993), "Intelligent Document Understanding: Understanding Photos

- with Captions”, *Proceedings of the International Conference on Document Analysis and Recognition (ICDAR-93, Tsukuba City, Japan)*: 664–667.
- Srihari, Rohini K. (1994), “Use of Collateral Text in Understanding Photos in Documents”, *Proceedings of the Conference on Applied Imagery and Pattern Recognition (AIPR/SPIE, Washington, DC)*: 186–199.
- 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.
- Srihari, Rohini K., & Rapaport, William J. (1990), “Combining Linguistic and Pictorial Information: Using Captions to Interpret Newspaper Photographs”, in Deepak Kumar (ed.), *Current Trends in SNePS—Semantic Network Processing System*, Lecture Notes in Artificial Intelligence, No. 437 (Berlin: Springer-Verlag): 5–96.
- Swan, Jim (1994), “Touching Words: Helen Keller, Plagiarism, Authorship”, in Martha Woodmansee & Peter Jaszi (eds.), *The Construction of Authorship: Textual Appropriation in Law and Literature* (Durham, NC: Duke University Press): 57–100.
- Tarski, Alfred (1969, June), “Truth and Proof”, *Scientific American*: 63–70, 75–77.
- Taylor, John G. (2002), “Do Virtual Actions Avoid the Chinese Room?”, in John Preston & Mark Bishop (eds.), *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (Oxford: Oxford University Press): 269–293.
- Terrace, Herbert S. (1985), “In the Beginning Was the ‘Name’ ”, *American Psychologist* 40: 1011–1028.
- Terrace, Herbert S. (1991), Letter to the Editor, *New York Review of Books*, Vol. 3, No. 15 (10 October 1991): 53.
- Thomason, Richmond H. (2003), “Dynamic Contextual Intensional Logic: Logical Foundations and an Application”, in P. Blackburn et al. (eds.), *CONTEXT 2003*, Lecture Notes in Artificial Intelligence 2680 (Berlin: Springer-Verlag): 328–341.
- Turing, Alan M. (1950), “Computing Machinery and Intelligence”, *Mind* 59: 433–460; reprinted in Shieber 2004.
- Vauclair, Jacques (1990), “Primate Cognition: From Representation to Language”, in S.T. Parker & K.R. Gibson (eds.), *‘Language’ and Intellect in Monkeys and Apes* (Cambridge, UK: Cambridge University Press): 312–329.
- Von Glasersfeld, E. (1977), “Linguistic Communication: Theory and Definition”, in Duane M. Rumbaugh (ed.), *Language Learning by a Chimpanzee: The LANA Project* (New York: Academic Press): 55–71.
- Weizenbaum, Joseph (1966), “ELIZA—A Computer Program for the Study of Natural Language Communication between Man and Machine”, *Communications of the ACM* 9: 36–45; reprinted, with correspondence from *CACM* 17 (1974): 543, in *Communications of the ACM* 26(1) (1983): 22–28.
- Wilson, Frank R. (1998), *The Hand: How Its Use Shapes the Brain, Language, and Human Culture* (New York: Pantheon Books).
- Winston, Patrick Henry (1975), “Learning Structural Descriptions from Examples”, in Patrick Henry Winston (ed.), *The Psychology of Computer Vision* (New York: McGraw-Hill): 157–209; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann,

- 1985): 141–168 (page references are to this reprint).
- Wittgenstein, Ludwig (1958), *Philosophical Investigations*, 3rd edition, trans. by G.E.M. Anscombe (New York: Macmillan).
- Woods, William A. (1975), “What’s in a Link: Foundations for Semantic Networks”, in Daniel G. Bobrow & Alan Collins (eds.), *Representation and Understanding: Studies in Cognitive Science* (New York: Academic Press): 35–82; reprinted in Ronald J. Brachman & Hector J. Levesque (eds.), *Readings in Knowledge Representation* (Los Altos, CA: Morgan Kaufmann, 1985): 217–241. Also reprinted in Allan Collins and Edward E. Smith (eds.), *Readings in Cognitive Science* (San Mateo, CA: Morgan Kaufmann, 1988).
- Zuckermann, Lord (1991), Letter to the Editor, *New York Review of Books*, Vol. 3, No. 15 (10 October 1991): 53.

Figures

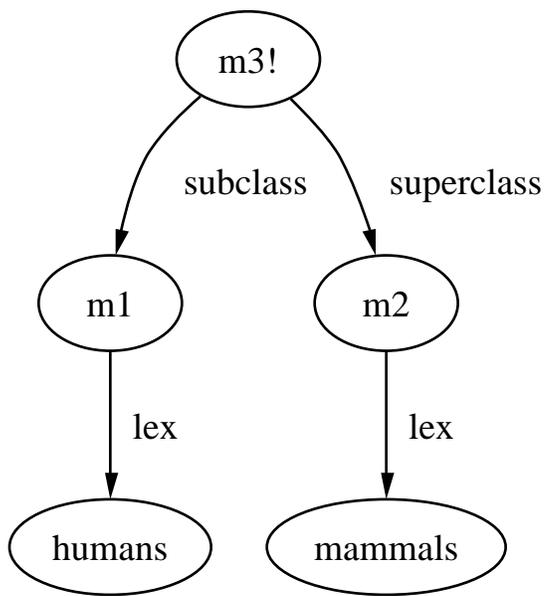


Figure 1: A simple SNePS network. (See text, §1.4, for an explanation. Note that the graphical package that drew the networks for this essay places all arc labels to the *right* of their arc.)

1. はじめに

自然言語理解システムの構築はマン・マシン・インタフェースおよび電子化された文書データの高度利用という観点から非常に重要な課題である。しかし、自然言語を理解する仕組みの構築には、汎用性や実用性の観点から多くの問題が依然として残されている¹⁾。この問題に対して、我々は自然言語理解システム構築のための一つの基礎として、日本語文に典型的に出現する名詞述語文の理解について研究を進めている。ここで、名詞述語文とは、主語と述語の対立の中で、述語が名詞で作られる文のことである²⁾。名詞述語文については、日本語学の観点から種々の考察が行われている^{2, 3, 4)}。しかし、計算言語学的観点から十分研究がなされているとはいえない。このような立場から我々は、従来より名詞述語文「<名詞句1>は<名詞句2>である。」を対象として、名詞句間の意味関係の学習と解析を行う手法の基礎的な考察を行っている⁵⁾。

本稿では、名詞句間の意味関係を学習・解析する手法の概要とその有効性を評価するため本手法に基づく実験システムを作成して行った実験結果について述べる。

2. 名詞句間の意味関係

名詞述語文における<名詞句1>と<名詞句2>の間の基本的な意味関係を以下に示す。

- (1) 下位・上位関係
 <名詞句2>が<名詞句1>の上位概念を表わす。
 例：鯨はほ乳類です。
- (2) 同一関係
 <名詞句1>と<名詞句2>が同一概念を表わす。
 対象レベルと概念レベルでの同一関係がある。
 例1：富士山は日本一の山です。
 (対象レベルでの同一関係)
 例2：正方形は直角正四辺形です。
 (概念レベルでの同一関係)
- (3) 対象・属性関係
 <名詞句1>が対象で<名詞句2>が<名詞句1>の属性値を表わす。
 例：太郎は医者です。
- (4) 対象・事象関係
 <名詞句1>が対象で<名詞句2>が<名詞句1>の関与する事象を表わす。
 例：父が帰宅です。
- (5) 要素・集合関係
 <名詞句1>が<名詞句2>の表わす集合の要素となっている。
 例：雄は性別です。
- (6) うなぎ文関係⁶⁾
 「太郎は鯉です。」に代表される文で文脈に依して、例えば「太郎は鯉を食べます。」という意味に解釈される。これは、太郎という名前の鯉がいて文字どおりそれが鯉であるという解釈も可能である。
- (7) 比喩関係
 <名詞句1>の属性に<名詞句2>の顕著な属性を重ねる隠喩を表わす。
 例：太郎は鯉です。
 (太郎は鯉のようにぬらりくらりと言いつをする。)
- (8) 同語反復同一関係
 <名詞句1>と<名詞句2>に同じ名詞句を置き、文脈(状況)の中で修辭的解釈が行なわれる。

例：太郎は太郎です。自分で考え、自分で行動すべきです。

ここでは、文字どおりの意味である(1)~(5)を対象とし、意味関係の学習と解析を行う。文字どおり以外の意味である(6)~(8)については、今後研究を進め別の機会に報告したい。

3. 知識表現

我々は、知識表現として、S.C.Shapiroらによって提案された意味ネットワークSNePS^{7, 8)}に基づく表現を用いている。図1に「人間はほ乳類です。」のSNePS表現の例を示す。SNePSでは節(nodes)が、命題、対象、属性、関係などを表わし、弧(arcs)はそれらの間の構造的、意味的つながりを表わす。ここで、ある概念からある概念までの弧の連なりを二つの概念間のパス(経路)と呼ぶ。図1で<lex>につながる<人間>と<ほ乳類>は語彙であり、<m1>と<m2>は<人間>と<ほ乳類>の概念であり、<m3>は<m1>と<m2>が下位・上位関係であるという概念である。

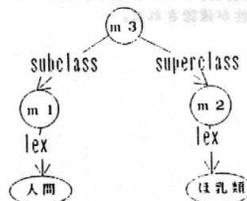


図1 SNePSの例

ここで用いる概念生成モデルは、事物の概念である具象概念を事物の性質に対応する属性概念によって制限することにより新たな概念である派生概念を生じ、それが具象概念あるいは属性概念として知識の中に組み込まれてゆくというものである。したがって、概念は具象概念、属性概念、派生概念のいずれかに分類される。これは、学習の段階でヒューリスティックスを用いて行われる。

ここで、2章で述べた(1)~(5)までの意味関係をネットワーク上での弧で表現する際の表記について述べる。ここで、(1)、(3)、(4)、(5)については、それぞれ<名詞句1>を始点とするか<名詞句2>を始点とするかで2通り考えられるので、実際には以下のような9つの関係が考えられる。

- (1) 下位・上位関係 : <-subclass-:-superclass->
 上位・下位関係 : <-superclass-:-subclass->
- (2) 同一関係 : <-equiv-:-equiv->
- (3) 対象・属性関係 : <-object-:-property->
 属性・対象関係 : <-property-:-object->
- (4) 対象・事象関係 : <-object-:-event->
 事象・対象関係 : <-event-:-object->
- (5) 要素・集合関係 : <-member-:-class->
 集合・要素関係 : <-class-:-member->

4. 概要

Figure 2: A Japanese-speaking computational cognitive agent (Arahi & Momouchi 1990: 2).

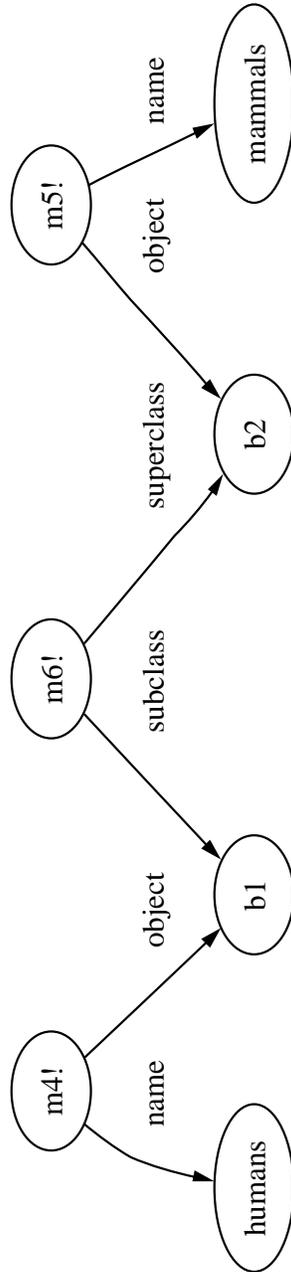


Figure 3: $[[m6]] = \text{Humans are mammals}$. Here, the name-object case frame is used instead of the `lex` arcs of Fig. 1. Nodes `b1` and `b2` correspond to nodes `m1` and `m2` of Fig. 1. The nodes labeled “humans” and “mammals” represent lexical entries exactly as the corresponding nodes of Fig. 1 do. Nodes `m4` and `m5` represent hard-to-express propositions relating entities to names for them (see text).

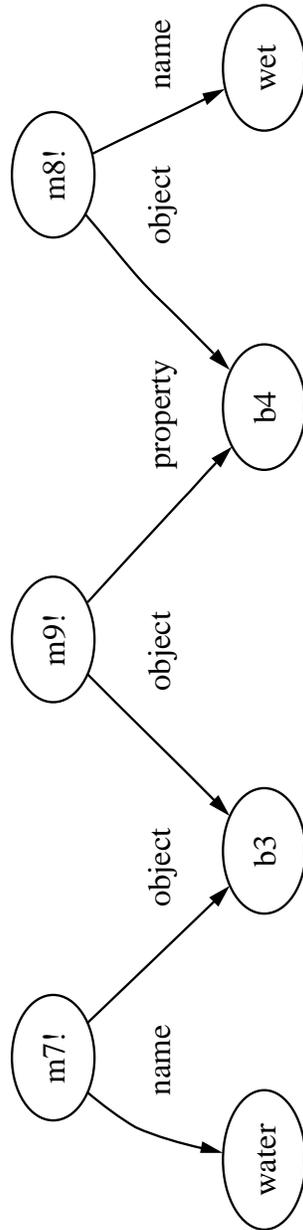


Figure 4: $[[m9]] = \text{Water is wet}$. Here, the proposition is analyzed as representing that an object $[[b3]]$ —i.e., something expressed by the “name” ‘water’—has the property $[[b4]]$ —i.e., something expressed by the “name” ‘wet’.

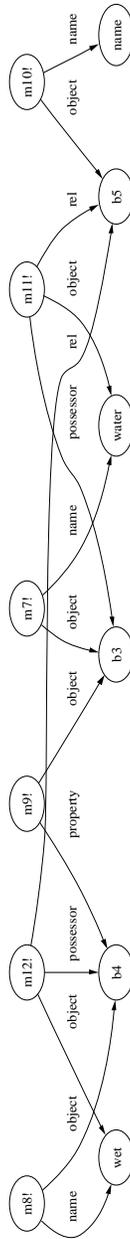


Figure 5: A network containing the network of Fig. 4 as a subnetwork. Here, $[[m11]]$ = Water's name is 'water'; i.e., $[[b3]]$'s $[[b5]]$ is 'water'. $[[m12]]$ = Wet's name is 'wet'; i.e., $[[b4]]$'s $[[b5]]$ is 'wet'. $[[m10]]$ = the inexpressible proposition asserting the name of $[[b5]]$ —i.e., the concept of a name—to be the lexical entry 'name'.

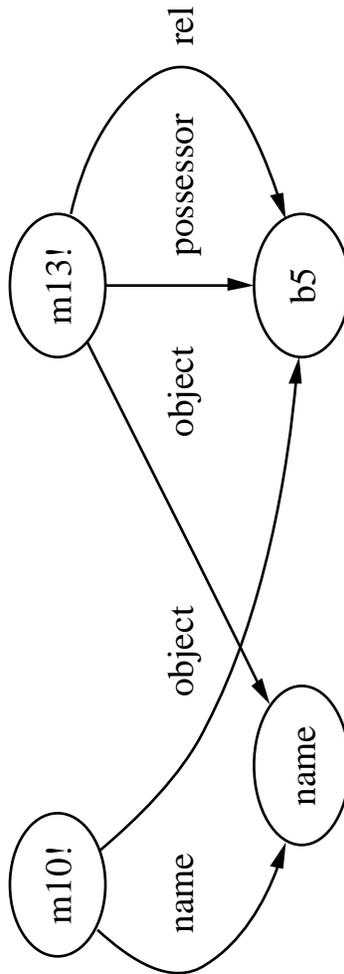


Figure 6: Node m10 is the hard-to-express node that represents that Cassie can *use* the name ‘name’ to talk about [[b5]]. Node m13 represents the proposition that allows Cassie to talk about the naming relationship between [[b5]] and ‘name’, viz.: [[m13]] = ‘name’ is [[b5]]’s name. I.e., the node labeled name represents [[b5]]’s [[b5]]; i.e., it represents [[b5]]’s name; i.e., it represents the name ‘name’.

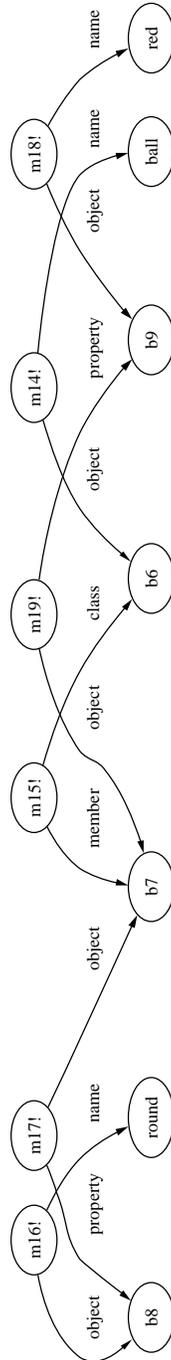


Figure 7: $[[m15]]$ = Something (viz., $[[b7]]$) is a ball (i.e., is a member of the class $[[b6]]$, which is expressed by the lexical entry ‘ball’).
 $[[m17]]$ = It is round (i.e., it has the property expressed by ‘round’).
 $[[m19]]$ = It is red (i.e., it has the property expressed by ‘red’).
 N.B.: These nodes are *added* to the networks of Figures 5 and 6.

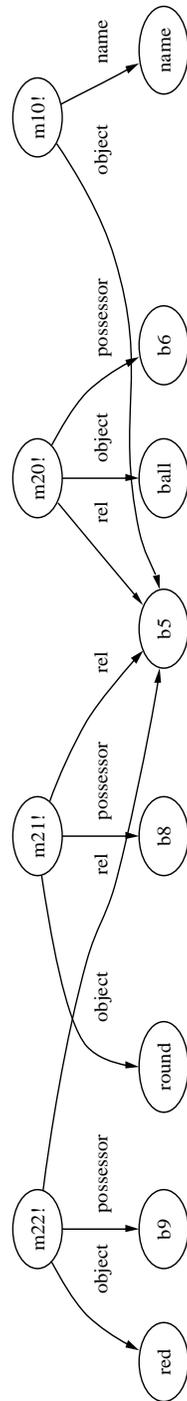


Figure 8: Nodes added to the networks of Figs. 5–7 after Cassie learns the concept of “names for things”. (Recall from Figs. 5 and 6 that $[[b5]]$ = the concept of “a name”.

$[[m20]]$ = $[[b6]]$ ’s name is ‘ball’.

$[[m21]]$ = $[[b8]]$ ’s name is ‘round’.

$[[m22]]$ = $[[b9]]$ ’s name is ‘red’.