CONTEXTUAL VOCABULARY ACQUISITION:
Development of a Computational Theory and Educational Curriculum

William J. Rapaport, Department of Computer Science & Engineering and Center for Cognitive Science
Michael W. Kibby, Department of Learning & Instruction and Center for Literacy & Reading Instruction
State University of New York at Buffalo, Buffalo, NY 14260

Section A: Project Summary. CONTEXTUAL VOCABULARY ACQUISITION

No doubt you have on occasion read some text containing an unfamiliar word, but you were unable or unwilling to find out from a dictionary or another person what the word meant. Nevertheless, you might, consciously or not, have figured out a meaning for it. Suppose you didn't, or suppose your hypothesized meaning was wrong. If you never see the word again, it may not matter. However, if the text you were reading were from science, mathematics, engineering, or technology (SMET), not understanding the unfamiliar term might seriously hinder your subsequent understanding of the text. If you do see the word again, you will have an opportunity to revise your hypothesis about its meaning. The more times you see the word, the better your definition will become. And if your hypothesis development were deliberate, rather than "incidental", your command of the new word would be stronger.

We propose (a) to extend and develop algorithms for computational contextual vocabulary acquisition (CVA): learning, from context, meanings for "hard" word: nouns (including proper nouns), verbs, adjectives, and adverbs, (b) to unify a disparate literature on the topic of CVA from psychology, first- and second-language (L1 and L2) acquisition, and reading science, in order to help develop these algorithms, and (c) to use the knowledge gained from the computational CVA system to build and to evaluate the effectiveness of an educational curriculum for enhancing students' abilities to use deliberate (i.e., non-incidental) CVA strategies in their reading of SMET texts at the middle-school and college undergraduate levels: teaching methods and guides, materials for teaching and practice, and evaluation instruments. The knowledge gained from case studies of students using our CVA techniques will feed back into further development of our computational theory. This project falls within **Quadrant 2** (fundamental research on behavioral, cognitive, affective and social aspects of human learning) and **Quadrant 3** (research on SMET learning in formal and informal educational settings).

The computational aspects of this project build upon our previous work on the development of a computational CVA theory and software system called 'Cassie'. Cassie consists of the SNePS knowledge-representation and reasoning system and a knowledge base of background information representing the knowledge that a reader (e.g., Cassie) brings to the text containing the unknown term. Input to the system consists, in part, of information from the text being read, which is parsed and incorporated directly into the knowledge-representation formalism. Cassie's other input is questions asked about the material being read. The question "What does [word] mean?" triggers a deductive search of the knowledge base, now consisting of background information plus information from the text, all marked with its "degree" of immunity from revision (roughly, a measure of the trustworthiness of the source of information). Output consists of a report of Cassie's current definition of the word in its context, or answers to other queries.

"Hard" words might be novel to the reader ('brachet'), familiar but misunderstood (does 'smite' mean to kill by hitting hard, or merely to hit hard?), or familiar but being used in a new sense (might "dressing" a sword mean to put clothes on it?). 'Context' includes the prior and immediately surrounding text, grammatical information, and the reader's background knowledge, but no access to a dictionary or other external source of information (including a human).

Our theory is that the meaning of such a word (1) *can* be determined from context, (2) can be *revised* upon further encounters, (3) "*converges*" to a dictionary-like definition if enough context has been provided and if there have been enough encounters, but (4) is always subject to further revision. Each encounter yields a definition (a hypothesis about meaning) and provides an opportunity to revise it in light of new evidence. The revision is *unsupervised*: There is no (human) "trainer" and no "error-correction" techniques. Finally, no *subject-matter-(or "domain")-specific* antecedent background information is required for the development and revision of the hypothesized definition (with the exception of the word's part of speech). The domain-independence of our system can make it more difficult to develop a good definition quickly, but is intended to model the typical reader of an arbitrary text. Clearly, the more background knowledge, including specialized knowledge, that the reader brings to the text, the more efficiently the unknown term can be learned.

In addition to developing the grammar and improving the search algorithms, we plan to adapt the system to be an aid to students who are learning explicit CVA techniques. Extant educational programs for CVA give only vague suggestions (often just to the teacher, not the student) for how to infer meaning from context, assume reading materials with words already semantically "clustered", or offer texts that always include a sentence explicitly defining the unknown term. Instead, we will use actual reading material (or develop our own as needed), from different grade levels, of the sort that a student would normally encounter.

It is not our intent to use a computer to teach students but, rather, to apply the knowledge gained from programming Cassie to do CVA to the creation of instructional methods to teach *students* to do CVA in regular classroom instructional milieus. We are not building a teaching machine, but rather teaching a machine to see if what we learn in teaching it can help us to teach students better.

Section C: Project Description
CONTEXTUAL VOCABULARY ACQUISITION:
Development of a Computational Theory and Educational Curriculum.

1 OVERVIEW.

No doubt you have on occasion read some text containing an unfamiliar word, but you were unable or unwilling to find out from a dictionary or another person what the word meant. Nevertheless, you might, consciously or not, have figured out a meaning for it. Suppose you didn't, or suppose your hypothesized meaning was wrong. If you never see the word again, it may not matter. However, if the text you were reading were from science, mathematics, engineering, or technology (SMET), not understanding the unfamiliar term might seriously hinder your subsequent understanding of the text. If you do see the word again, you will have an opportunity to revise your hypothesis about its meaning. The more times you see the word, the better your definition will become. And if your hypothesis development were deliberate, rather than "incidental" (Nagy 1997), your command of the new word would be stronger.

Our project has two major goals. (1) Our first goal is to extend and develop algorithms for computational contextual vocabulary acquisition (CVA) by a computer, i.e., programming a computer to learn from context the meanings of "hard" words (nouns (including proper nouns), verbs, adjectives, and adverbs). The development of these computational algorithms will be informed by two research processes. First (a) we will unify the disparate literature on the topic of CVA from psychology, first- and second-language (L1 and L2) acquisition, and reading science, and second (b) we will conduct several case studies of students using CVA techniques while reading SMET texts. We will use the knowledge gained from the literature and case studies to inform the development of our computational algorithms. (2) Our second major goal is to use the knowledge gained from the application of the computational algorithms to develop a CVA curriculum (i.e., teaching methods and guides, materials for teaching and practice, and evaluation instruments) to help teachers teach students to use CVA techniques for identifying hard words in SMET texts and to evaluate this curriculum in actual SMET middle-grade and college-level classrooms. This project falls within **Quadrant 2** (fundamental research on behavioral, cognitive, affective and social aspects of human learning) and **Quadrant 3** (research on SMET learning in formal and informal educational settings).

The computational aspects of this project build upon our previous work on the development of a computational CVA theory and software system called 'Cassie' (Ehrlich 1995, Ehrlich & Rapaport 1997, Hunt & Koplas 1998, Rapaport & Ehrlich 2000). Cassie consists of the SNePS knowledge-representation and reasoning system (see §3.1) and a knowledge base of background information representing the knowledge that a reader (e.g., Cassie) brings to the text containing the unknown term. Input to the system consists, in part, of information from the text being read, which is parsed and incorporated directly into the knowledge-representation formalism. Cassie's other input is questions asked about the material being read. The question "What does [word] mean?" triggers a deductive search of the knowledge base, now consisting of background information plus information from the text, all marked with its "degree" of immunity from revision (roughly, a measure of the trustworthiness of the source of information). Output consists of a report of Cassie's current definition of the word in its context, or answers to other queries.

"Hard" words might be novel to the reader ('brachet'), familiar but misunderstood (does 'smite' mean to kill by hitting hard, or merely to hit hard?), or familiar but being used in a new sense (might "dressing" a sword mean to put clothes on it?). 'Context' includes the prior and immediately surrounding text, grammatical information, and the reader's background knowledge, but no access to a dictionary or other external source of information (including a human).

Our theory is that the meaning of such a word (1) *can* be determined from context, (2) can be *revised* upon further encounters, (3) "*converges*" to a dictionary-like definition if enough context has been provided and if there have been enough encounters, but (4) is always subject to further revision. Each encounter yields a definition (a hypothesis about meaning) and provides an opportunity to revise it in light of new evidence. The revision is *unsupervised*: There is no (human) "trainer" and no "error-correction" techniques. Finally, no *subject-matter-(or "domain")-specific* antecedent background information is required for the development and revision of the hypothesized definition (with the exception of the word's part of speech). The domain-independence of our system can make it more difficult to develop a good definition quickly, but is intended to model the typical reader of an arbitrary text. Clearly, the more background knowledge, including specialized knowledge, that the reader brings to the text, the more efficiently the unknown term can be learned.

Evidence for this comes from psychology (Elshout-Mohr & van Daalen-Kapteijns 1987, Johnson-Laird 1987, Sternberg 1987), L1 and L2 vocabulary-acquisition (see Rapaport 2000c for a bibliography), and (informal) protocols that we have taken from subjects who were asked to reason out loud about their definition-forming and revision procedures when shown passages containing unknown words. These same passages have served as input to Cassie, the computational system that develops and revises definitions in ways similar to the human subjects.

In addition to developing the grammar and improving the search algorithms, we plan to adapt the system to be an aid to students who are learning explicit CVA techniques. Extant educational programs for CVA give only vague suggestions (often just to the teacher, not the student) for how to infer meaning from context, assume reading materials with words already semantically "clustered" (Marzano et al. 1991), or offer texts that always include a sentence explicitly defining the unknown term (Mueser 1984, Mueser & Mueser 1984). Instead, we will use actual reading material (or develop our own as needed), from different grade levels, of the sort that a student would normally encounter (§4.2.1(1)).

It is not our intent to use a computer to teach students but, rather, to apply the knowledge gained from programming Cassie to do CVA to the creation of instructional methods to teach *students* to do CVA in regular classroom instructional milieus. We are not building a teaching machine, but rather teaching a machine to see if what we learn in teaching it can help us to teach students better.

The system is being developed in the SNePS Research Group (SNeRG)[1] of the SUNY Buffalo Department of Computer Science and Engineering and the Center for Cognitive Science,[2] and the system and curriculum will be tested in The Center for Literacy and Reading Instruction[3] of the SUNY Buffalo Department of Learning and Instruction. This Center serves three major purposes: training, community service, and research. It provides internships in a clinical facility where graduate students complete diagnostic assessments of children thought to have reading or learning problems, and also extends one-to-one remedial instruction to reading-disabled children. As a community service, the Center offers diagnostic and remedial treatment to children in the Western New York area. The staff and director of the Center also regularly conduct and publish research related to diagnostic and remedial procedures for children with reading difficulties.

2 SIGNIFICANCE. Other *computational* theories of CVA[4] either (1) make recourse to online dictionaries or human informants (e.g., Zernik & Dyer 1987), whereas ours does not; (2) only map unknown terms onto known concepts (e.g., Hastings 1994; Hastings & Lytinen 1994ab), whereas our system is capable of concept *formation*, i.e., *constructing*, rather than merely *finding*, meanings of unknown terms; (3) are concerned with "correctness" of word meaning (Hastings), whereas for us such a notion does not apply; or (4) are more concerned with modeling L1 acquisition during childhood (e.g., Siskind 1992, 1994, 1996), whereas our focus is on relatively mature readers who already know a large part of their language and are (merely) expanding their vocabulary.

Nor do we know of other *curriculum* programs like ours. Sternberg et al. 1983 described a computer-based tutoring system for CVA that was never built.[5] Aist 2000 "focus[es] on developing computerized methods to help children learn (as opposed to developing a computer program that learns)"[6] and as opposed to devising a classroom curriculum (as we are), and is "adding extra information to textual contexts",[6] whereas we are investigating the use of *natural* texts. Moreover, we will be basing our curricular research on case studies with students actually doing CVA.

The importance of our project stems from the need (a) for natural-language-understanding systems that can operate independently of human assistance and (b) to improve both the teaching of reading and students' reading ability (especially in SMET). Our project is also distinctive in its proposed use of mutual feedback between the development of the computational theory and the educational curriculum.

2.1 Computational Significance. We want natural-language-understanding (a.k.a. "text understanding", "message processing", or "information extraction") systems to be (as) robust (as humans): They should not break down just because they have encountered an unknown expression. This is especially the case for systems that use unconstrained input text and must operate independently of human intervention. For example, "intelligent agents" ought to be able to figure out a human user's instructions without necessarily stopping to ask what each new word means. No assumption of a "fixed complete lexicon" can be made (Zernik & Dyer 1987): It could not be manually encoded, nor could it contain neologisms or new meanings given to old words in new contexts. Similarly, a system designed to locate "interesting" news items from an online information server should not be limited to keyword searches; e.g., if the user is interested in news items about dogs, and the filter detects items about "brachets" (a term not in its lexicon), it should

[1] [http://www.cse.buffalo.edu/sneps]

[2] [http://www.cogsci.buffalo.edu/]

[3] [http://www.readingcenter.buffalo.edu/]

[4] E.g.: Granger 1977, 1983; Kiersey 1982; Berwick 1983; Haas & Hendrix 1983; Carbonell & Hayes 1984; Pustejovsky 1987; Carter 1989; Asker et al. 1992; Hearst 1992. For discussion, see Rapaport & Ehrlich 2000; for a more complete list, see Rapaport 2000c.

[5] Sternberg, personal communication, 1/26/2000.

[6] Aist, personal communication, 11/10/2000.

deliver those items as soon as it decides that a brachet is a kind of dog. And, of course, a student who comes upon an unfamiliar term in a text should be able to figure out what it means, at least to the extent of being able to comprehend the rest of the text, if not to give a formal, precise definition.

Two key features of our system mesh nicely with these desiderata; they can be summarized as the advantages of learning over being told: (1) Being told requires human intervention, whose availability cannot be guaranteed. *Our system operates (i.e., learns meanings) independently of a human teacher or trainer* (with one exception that we propose to overcome), and thus can model a student learning CVA techniques. (2) One cannot predict all information that might be needed to understand unconstrained, domain-independent text; hence, the system must be able to learn on its own. *Our system does not constrain the subject matter (the "domain") of the text.* Although our research so far has been primarily concerned with narrative text, our techniques are perfectly general; here, we are proposing to extend them to SMET texts. Assuming the availability of an appropriate grammar, we are developing (and propose to elaborate on) algorithms for producing definitions independent of domain. However, the definitions *are* dependent on the system's background knowledge: The more background knowledge the system has, the better the definitions will be, and the more quickly they will "converge".

On the prescriptive approach to lexicography, definitions are "correct"; any other use is, by definition, incorrect. On the descriptive approach, a word means what it is used to mean by those who use it. For example, the prescriptive definition of the phrase '*à la mode*' is "in the current fashion", but a descriptive definition of the phrase as observed in American contexts would be something like "with ice cream". Our system produces descriptive definitions; i.e., we are *not* proposing a system that develops "correct" definitions (cf. Wilks & Fass's (1992) theory of "preference semantics"). Rather, our system develops dictionary-like definitions that enable the reader to continue with the task of understanding the text. Sometimes, an author uses a word incorrectly. However, since we are taking a descriptive approach to the definition of words and are interested in allowing Cassie to determine what is meant by a word in a given context, we have not considered any means to allow her to conclude that a word has been misused. If the usage does not agree with previous understanding, our approach is to develop a secondary definition for it.

Although both the computational and educational literature talk about "vocabulary" acquisition, it is really a *concept* that is acquired, for it is the *meaning* of the unknown word that is being learned. Sometimes, what the reader learns when seeing an unknown word is how to read it. The word might actually be a familiar word that the reader has never seen in print before, or it might be a new word that falls into one of the other two categories below. In any case, the reader is not learning what the word means; we are not concerned with this kind of "sight-vocabulary" acquisition. Other times, the unknown word is (merely) a synonym for a known word or a term for a known concept (as in Hastings's work). Although our system is capable of determining this, we feel that the more interesting and important situation is that in which the unknown word names a *new* concept: See the discussions of 'brachet', 'smite', and 'dress', in §3. In this third situation, "vocabulary" acquisition is really a form of *ontology learning*.

2.2 Educational Significance. Successfully programming Cassie to derive a word's meaning from context and to tell us how she did it is no small feat and is good basic research. However, this multidisciplinary proposal combines basic and applied research. Its significance would not be in its mere accomplishment, but in whether the knowledge gained by developing such a successful program can be applied to teaching CVA strategies to students so that they are able to use them successfully when they encounter hard words in their regular reading of SMET texts.

Helping students to gain the meanings of hard words from the contexts they are reading has been on the "to do" list of teachers for decades. No textbook in teaching reading or content areas (e.g., SMET, social studies) fails to include admonitions and suggestions to help students in CVA. But where the experts in reading and content areas fall short is in specifying: the specific forms of context available for CVA, the degree of helpfulness of a given context, the parameters of the role of prior (or background) knowledge in using available context, the cognitive/logical processes that must be applied to these contexts and prior knowledge, how to evaluate one's confidence in the derived meaning, or how often a word must be encountered to be confidently learned. In spite of this, humans learn huge numbers of words and concepts from the oral and written contexts they encounter daily. We need to conduct more systematic studies of CVA, not only to understand how context does and does not operate, but, more importantly, to teach students to be more efficient in their use of context. We have an unusual opportunity here to blend the talents of cognitive computer scientists and reading scientists to explore these issues.

Educators and psychologists argue that the vast majority of words that we know and use every day are not learned from direct instruction, but from oral and written contexts (Anderson & Freebody 1981; Beck et al. 1983; Sternberg 1987; Nagy & Herman 1987; Nagy 1988; Stahl 1999). Given the number of words in the average adult's reading, listening, speaking, and writing vocabularies, it is easy to agree that only a small portion of these words were taught directly to us as children, students, or adults. We are not saying that humans do not already use context to learn words or that schools are not already teaching context skills. However, the specific context clues that are recommended for instruction are not the result of scientific study of actual CVA, but the result of logical analyses of language usage and traditional notions about context clues (Deighton 1978, Fisher 1986, Nagy 1988, Harris & Sipay 1990, Friedland 1992, Blachowicz & Fisher 1996, Lipson & Wixon 1997, Stahl 1999). There has not been sufficient systematic study of the actual processes required in gaining word meanings from natural printed contexts. We do not really know how humans integrate context, prior knowledge, and cognitive/logical operations in CVA, nor do we know how many times these operations must be applied to a given word in context for that word to be confidently learned. We propose to conduct exhaustive tests of algorithms for maximizing the interplay of the text's language and meaning clues and the student's prior knowledge and logical thinking for the further development of our theory and software system.

We propose two strategies for studying how humans integrate context, prior knowledge, and cognitive/logical operations in CVA: The first is for cognitive computer scientists to test systematically those known (and to be discovered; see next strategy) forms of contextual support for CVA with varying degrees of prior knowledge and varying algorithms. We know that context can be both facilitating (Gipe 1980; Beck et al. 1983) as well as misleading (Schatz & Baldwin 1986), but we do not know the specific parameters of what makes a context helpful, not helpful, or misleading (§4.2.2(1c)). A second strategy will be case studies, conducted by the reading scientists, of students applying CVA techniques. The knowledge gained from these case studies will become part of the information the cognitive computer scientists will program into Cassie for further testing. We will also conduct, in regular teaching environments, pilot applications of those CVA strategies generated by Cassie.

Determining how humans do (or *should* do) CVA is a research problem well suited for artificial intelligence. If this application of artificial intelligence results in new knowledge about how humans do CVA, then it is possible to create better curricula and instructional materials for teaching CVA strategies in the reading of SMET texts.

3 CURRENT IMPLEMENTATION.

3.1 Software Environment. The technology we employ is the SNePS-2.5 knowledge-representation-and-reasoning system (Shapiro & Rapaport 1987, 1992, 1995; Shapiro et al. 1999; Shapiro 2000). SNePS has been and is being used for a number of research projects in natural-language understanding.[7] Each node in a SNePS network represents a concept or mental object (possibly built of other concepts), with labeled arcs linking the concepts. All information, including propositions, is represented by nodes, and propositions about propositions can be represented without limit. Arcs merely form the underlying syntactic structure of SNePS. Paths of arcs can be defined, allowing for path-based inference, including property inheritance within generalization hierarchies. There is a 1–1 correspondence between nodes and represented concepts. This uniqueness principle guarantees that nodes will be shared whenever possible and that nodes represent "intensional" objects, i.e., concepts, propositions, properties, and such objects of thought as fictional entities (e.g., Sherlock Holmes), non-existents (e.g., unicorns), and impossible objects (e.g., the round square) (Shapiro & Rapaport 1987, 1991). This wide representational ability is especially appropriate for CVA from arbitrary texts, whose subject matter could range from factual science to science fiction. After all, objects about which we can think, speak, or write need not exist; this includes not only unicorns, but possibly black holes, bosons, or other theoretical entities of contemporary SMET.

SNePS's inference package allows rules for both deductive and default reasoning. In the presence of a contradiction, the SNeBR belief-revision package allows for the removal of one or more of the propositions from which the contradiction was derived, as well as the conclusions that depended on it (Martins & Shapiro 1988). This mechanism is used to revise definitions that are inconsistent with a word's current use. We have developed algorithms for *partially* automating the identification and removal or modification of the offending premise, based on SNePSwD, a default belief-revision system that enables automatic revision (Martins & Cravo 1991, Cravo & Martins 1993). Members of SNeRG are exploring techniques for more *fully* automating it (Johnson & Shapiro 2000ab). SNePS also has an English lexicon, morphological analyzer/synthesizer, and a generalized augmented-transition-network parser-generator that, rather than building an intermediate parse tree, translates the input English directly into the propositional semantic-network knowledge-representation system (Shapiro 1982, 1989; Shapiro & Rapaport 1995). Members of SNeRG are exploring the use of the LKB computational language/grammar environment (Copestake 1999) as an interface to SNePS, which would provide a much wider range of grammatical coverage.

[7] E.g.: Shapiro 1982, 1989; Rapaport 1986, 1988; Wiebe & Rapaport 1986, 1988; Neal & Shapiro 1987; Peters & Shapiro 1987ab; Peters, Shapiro, & Rapaport 1988; Peters & Rapaport 1990; Wiebe 1990, 1991, 1994; Shapiro & Rapaport 1991; Ali & Shapiro 1993; Haller 1993, 1995; Almeida 1995; Yuhan & Shapiro 1995; Rapaport, Shapiro, & Wiebe 1997; McRoy et al. 2000.

"Cassie", our CVA system, consists of SNePS-2.5 (including SNeBR and the augmented-transition-network parser-generator), SNePSwD, and a knowledge base of background information. Currently, the knowledge base is hand-coded, because it represents Cassie's antecedent knowledge; *how* she acquired this knowledge is irrelevant. We begin with what some might call a "toy" knowledge base, but each of our tests so far has included all previous information, so the knowledge base grows as we test more words. Cassie's input consists, in part, of information from the text being read. Currently, this, too, is coded directly in the knowledge-representation formalism. However, a major part of our proposed research is the updating and further development of our grammar in order to automate the transduction of sentences from the text into information in the knowledge base. The initial grammars that we write will be specific to the texts that we are using, to ensure that they can be handled by the system; eventual use of the LKB grammar will enable the system to be more robust.

This system provides a software laboratory for testing and experimenting with the theory. It is a component of an interdisciplinary, cognitive-science project to develop a computational cognitive model of a reader of narrative text (Duchan et al. 1995). The proposed research will be an important contribution to this and similar projects, since to fully model a reader, it is important to model the ability to learn from reading, in particular, to expand one's vocabulary in a natural way while reading, without having to stop to ask someone or to consult a dictionary.

3.2 Algorithms. Our implemented theory at least partially tests the thesis that symbol manipulation (syntax) suffices for natural-language understanding (Rapaport 1981, 1988, 1995, 2000b). Humans understand one another by interpreting the symbols they read or hear. This interpretation is a mapping from the speaker's (or writer's) syntax to the hearer's (or reader's) concepts (semantics). We take the meaning of a word (as understood by a cognitive agent) to be the position of that word in a highly interconnected network of words, propositions, and other concepts. That is, a word's meaning is its (syntactic) relation to other words, (Quillian 1968, 1969). We thus adopt Quine's (1951) view that the beliefs held by a cognitive agent form an interconnected web, where a change or addition to some portion of the web can affect other portions that are linked to it. Such an agent's understanding of natural-language input will, therefore, be a part of such a web or semantic network composed of internal (mental) objects. If we take these mental objects to be symbols, then the interpretation of linguistic input is a syntactic operation, and formal symbol manipulation is sufficient for determining meanings of words.

In this (idiolectic) sense, the meaning of a word for a cognitive agent is determined by idiosyncratic experience with it. But a word's dictionary definition usually contains less information than the idiolectic meaning. Since the contextual meaning described above includes a word's relation to every concept in the agent's mind, it is too unwieldy to be of much use. To limit the connections used to provide the definition, we select for particular *kinds* of information. Not all concepts within a given subnetwork need be equally salient to a dictionary-style definition of a word. Agent *A* may have a direct mental connection from 'bachelor' to 'John', while agent *B* may never have heard of him, yet *A* and *B* may be able to agree on a definition of 'bachelor'. In the attempt to understand and be understood, people abstract certain conventional information about words and accept this information as a definition.

When a new word is encountered, people begin to hypothesize a definition. Applying the fundamental principle that the meaning of a term is its location in the network of background information and story information, our algorithms for hypothesizing a definition operate by deductively searching the network for information appropriate to a dictionary-like definition (see Rapaport & Ehrlich 2000 for the algorithms themselves). We assume that our grammar has been able to identify the unknown word as a noun or a verb. We have developed algorithms for hypothesizing and revising meanings for nouns and verbs that are unknown, mistaken, or being used in a new way. (Work on other parts of speech is part of the proposed project.)

Cassie was provided with background information for understanding the King Arthur stories in Malory 1470. In one test, when presented with a sequence of passages involving the hitherto unknown noun 'brachet', Cassie was able to develop a theory that a brachet was a dog whose function is to hunt and that can bay and bite. (*Webster's Second* (1937) defines it as "a hound that hunts by the scent".) However, based on the first context in which the term appeared (viz., "Right so as they sat, there came a white hart running into the hall with a white brachet next to him, and thirty couples of black hounds came running after them with a great cry."), the initial hypothesis was merely that a brachet was a physical object that may be white.

Each time the term appeared, Cassie was asked to define it. To do so, she deductively searched her background knowledge base, together with the information she had read in the narrative to that point, for information concerning (1) direct class inclusions (especially in a basic-level category), (2) general functions of brachets (in preference to those of individuals), (3) the general structure of brachets (if appropriate, and in preference to those of individuals), (4) acts that brachets perform (partially ordered in terms of universality: probable actions in preference to possible actions, actions attributed to brachets in general in preference to actions of individuals, etc.), (5) possible ownership of brachets, (6) part/whole relationships to other objects, (7) other properties of brachets (when structural and functional description is possible, the less salient "other properties" of particular brachets are not reported, although we do report any properties that apply to brachets in general), and (8) possible synonyms for 'brachet' (based on similarity of the above attributes). Some of these are based on psycholinguistic studies of CVA (Elshout-Mohr & van Daalen-Kapteijns 1987, Johnson-Laird 1987, Sternberg 1987). In the absence of some or all of this information, or in the presence of potentially inconsistent information (e.g., if the text says that one brachet hunts and another does not), Cassie either leaves certain "slots" in her definitional framework empty, or includes information about particular brachets. Such information is filled in or replaced upon further encounters with the term.

To define a verb (*V*), we currently report its predicate structure, a categorization of its arguments, and any causal or enablement information we can find. To categorize an argument of *V*, we look for, in order of preference, its membership in a basic-level category, membership in a subclass of animal, or membership in some other known category. Clearly, our verb-definition algorithm is not as elaborate as our noun-defining algorithm. We propose to remedy this.

In another test, Cassie was told that 'to smite' meant "to kill by hitting hard" (a mistaken belief actually held by one of the PIs before reading Malory 1470. Passages in which various characters were smitten but then continued to act triggered SNeBR, which identifies several possible "culprit" propositions in the knowledge base to remove in order to block inconsistencies. The reader then decides which belief to revise. Although the decision about which proposition (representing an incorrect definition) to withdraw and which *new* proposition (representing a revised definition) to add has been partially automated (using SNePSwD), this remains the one area still occasionally requiring human intervention. Automating this will be a major focus of our research.

A third case is exemplified by 'to dress', which Cassie antecedently understood to mean "to put clothes on something". This is a well-entrenched meaning, which should *not* be rejected. However, upon reading that King Arthur "dressed" his sword, SNeBR detects an inconsistency. Rather than *rejecting* the prior definition, we *add* to it. In this case, Cassie decides that to dress is *either* to put clothes on *or* to prepare for battle.

3.3 Revision. When readers encounter a discrepancy between the way a word is used and their previous understanding of it, they must either assume that the word is used incorrectly or revise their previous understanding. When Cassie encounters a contradiction derived from combining story information with background knowledge, she must decide which of the premises leading to the contradiction should be revised. To do this, we tag each assertion in the knowledge base and the story with a "knowledge category" (`kn_cat`). These are ordered by certainty of belief, so that the system can choose a belief for revision from premises believed with the least certainty. The hierarchy of `kn_cat`s, from greatest certainty of belief to least, is:

`kn_cat` *intrinsic*: Facts about language, including simple assertions and rules; very basic or fundamental background knowledge. Found in the knowledge base, not (usually) in stories. For example, the temporal relation "before" is transitive; containment of an item in a class implies containment in superclasses.

`kn_cat` *story*: Information present in the story being read, including stated propositions and propositions implicit in the sentence (necessary for parsing it). For example, "Sir Gryflette left his house and rode to town" contains the following story facts: Someone is named Sir Gryflette. That someone left his house. That someone rode to town.

`kn_cat` *life*: Background knowledge expressed as simple assertions without variables or inference. For example, taxonomies (e.g., dogs are a subclass of animals), assertions about individuals (e.g., Merlin is a wizard).

`kn_cat` *story-comp*: "Story-completion" information inferred by the reader to make sense of the story (Segal 1995). Using the example from `kn_cat` *story*, story completion facts include: Sir Gryflette is a knight; Sir Gryflette mounted his horse between leaving his house and riding to town.

`kn_cat` *life-rule.1*: Background knowledge represented as rules for inference (using variables) reflecting common, everyday knowledge. For example, if *x* bears young, then *x* is a mammal; if *x* dresses *y*, then *y* wears clothing.

`kn_cat` *life-rule.2*: Background knowledge represented as rules for inference (using variables) reflecting specialized, non-everyday information. For example, if *x* smites *y*, then *x* kills *y* by hitting *y*.

`kn_cat` *questionable*: A rule that has already been subjected to revision because its original form led to a contradiction. This is a temporary classification while Cassie looks for confirmation of her revision of a *life-rule.2*. Once she settles on a particular revision, the revised rule is tagged as a *life-rule.2*.

In case of contradiction, if only one belief has the highest level of uncertainty in the conflict set, it will be revised. If several alternatives exist with the same (highest present) `kn_cat`, Cassie looks for a verb in the antecedent: Humans more readily revise beliefs about verbs than about nouns (Gentner 1981). If this is still insufficient to yield a single culprit, then, in the current implementation, a human "oracle" chooses one (although this hasn't been needed in our tests to date). Ideally, Cassie would use discourse information to make the decision between possible culprits at the same level of certainty. For example, in the case of dressing a sword before fighting, the rule about what it means to dress something might be selected for revision because it is unrelated to the topic of fighting, whereas swords are closely associated with the topic. We will explore refinements of this hierarchy as part of the proposed research. Members of SNeRG are working on this in a related project; we hope to be able to use their results as well as provide them with cases for testing (Johnson & Shapiro 2000ab).

3.4 An Example. Here, we sketch Cassie's handling of 'smite', with this background information in the knowledge base: *There is a king named King Arthur. There is a king named King Lot. There is a sword named Excalibur. Excalibur is King Arthur's sword. Horses are animals. Kings are persons. Knights are persons. Dukes are persons. "Person" is a basic-level category. "Horse" is a basic-level category. "Before" and "after" are transitive relations. If* x *is dead at time* t, x *can perform no actions at* t *or at any subsequent time. If* x *belongs to a subclass of person,* x *is a person. If a person acts, the act performed is an action. If an agent acts on an object, and there is an indirect object of the action, then the action is bitransitive. If an agent acts on an object, then the action is transitive. If an agent acts on itself, then the action is reflexive. If* x *is hurt at time* t*, then* x *is not dead at* t. *If* x *is* not *dead at time* t*, then* x *was not dead at any prior time. If* x *smites* y *at time* t*, then* x *hits* y *at* t*, and* y *is dead at* t*, and the hitting caused the death.* (Note that the last is the only information about 'smite' in the knowledge base.)

Cassie then reads a sequence of passages containing 'smite' (Malory 1470: 13ff) interspersed with questions and requests for definitions (**P***n* = passage #*n*; **D***n* = the definition created after **P***n*; **Q***n* = a question asked after **P***n*; **R***n* = the answer.)

P1: King Arthur turned himself and his horse. He smote before and behind. His horse was slain. King Lot smote down King Arthur.
D1: A person can smite a person. If *x* smites *y* at time *t*, then *x* hits *y* at *t*, and *y* is dead at *t*.
Q1: What properties does King Arthur have?
R1: King Arthur is dead.
P2: King Arthur's knights rescued him. They sat him on a horse. He drew Excalibur.
Q2: When did King Arthur draw?

The inference required to reply to **Q2** triggers SNeBR, which reports that King Arthur's drawing (i.e., acting) is inconsistent with his being dead. Cassie automatically removes the proposition reporting her belief that smiting entails killing, which is *automatically* replaced with two beliefs: that although smiting entails hitting, it only possibly entails killing (`kn_cat` *questionable*), and that if smiting results in a death, then the hitting is the cause of death. These rules are *not* built in; rather, they are *inferred* by a set of general rules for replacing discarded definitions with revised definitions, which is being developed.

D2: A person can smite a person. If *x* smites *y* at time *t*, then *x* hits *y* at *t* and possibly *y* is dead at *t*.
P3: Two of King Claudas's knights rode toward a passage. Sir Ulfyas and Sir Brastias rode ahead. Sir Ulfyas smote down one of King Claudas's two knights. Sir Brastias smote down the other knight. Sir Ulfyas and Sir Brastias rode ahead. Sir Ulfyas fought and unhorsed another of Claudas's knights. Sir Brastias fought and unhorsed the last of Claudas's knights. Sir Ulfyas and Sir Brastias laid King Claudas's last two knights on the ground. All of King Claudas's knights were hurt and bruised.

The information that the knights were hurt was added in forward-chaining mode to allow Cassie to notice that they were still alive at the time that they were hurt and therefore could not have died earlier at the time that they were smitten. Cassie has now heard of two cases in a row (King Arthur, and the two knights) where a smitee has survived being smitten, with no intervening cases of death by smiting, yielding:

D3: A person can smite a person. If *x* smites *y* at time *t*, then *x* hits *y* at *t*.

Further encounters with 'smite' cause no further revisions. Not only has the definition stabilized ("converged"), but it has done so in a manner similar to our human protocols.

4 PROPOSED RESEARCH: Extending and Applying the Theory. The proposed project has two parallel research "streams": computational and educational. Both will last through the full 3 years of the project; results from each will inform and support the other. Cutting across these two streams, we propose 4 tasks: (1) We will improve the coverage and behavior of our computational CVA system. As part of this, Cassie will be modified to output not only a definition of the unknown term, but also an explanation of how she inferred its meaning. Since Cassie will always give an appropriate meaning for the term, she can be used (2) to help develop curriculum materials in CVA techniques. (3) The new curriculum will be tested in field research with students reading real SMET texts. (4) Results from the curriculum-development case studies and field tests will be used to further develop the computational theory and system.

4.1 Computational Stream: Extending the Theory.

4.1.1 Improve the coverage of the grammar that transduces the natural-language text into SNePS. Our current toy grammars have the capacity to add new words to the lexicon when they are encountered (Hunt & Koplas 1998). This means only that a combination of morphological analysis and grammatical constraint on the parse of the sentence is used to determine the new word's part of speech. This suffices to allow the parse of sentences that contain the word, and to permit Cassie to decide whether to use her noun- or verb-definition algorithm when we ask what the word means. (Such a determination does not actually require a word's presence in the lexicon as opposed to the knowledge base: Its position in the network can identify its part of speech. But lexical look-up is faster than searching the network.) We do not intend as part of this project to write a comprehensive and robust grammar for English. Rather, we plan to write grammars that can handle the particular texts we will be working with. As indicated elsewhere in this document (§3.1), an independent, though related, research project in SNeRG concerns the adaptation of the LKB grammar to SNePS. Certainly, if that project is successful, we intend to use it. In addition, we intend our new grammar to build appropriate representations of causal and hyponym-hypernym (Hearst 1992) relations between clauses.

4.1.2 Improve and elaborate the algorithms for defining verbs, modifiers, and proper names. We plan to elaborate our theory of verb acquisition in the light of research on verb meanings (Gleitman 1990, 1994). It could also be improved by developing synonym-finding procedures and by the use of some scheme for classifying verbs of different types. One such scheme already available is Conceptual Dependency's primitive acts (Schank & Rieger 1974). Another possibility would be to attempt to develop a classification of verbs analogous to Rosch's superordinate, basic-level, and subordinate categories (Rosch 1978). If such a classification is possible, the concepts associated with Schank's primitive acts might be seen as superordinate categories, as would English words such as 'move' or 'go'. 'Walk' might be a basic-level verb, while 'amble', 'pace', 'stroll', 'stride', and 'hobble' might be considered subordinate level. If we can establish such a structure for the classification of (at least some types of) verbs, then, just as we defined 'brachet' as a type of dog, we might define 'striding' as a type of walking.

Determining the class to which a particular verb belongs, as well as determining synonyms, will depend largely on our ability to analyze the results of actions and events. To date, we have paid relatively little attention to the flow of a narrative. In defining nouns, the propositional content of individual sentences usually suffices. In defining the verbs 'dress' and 'smite', it was necessary to note somewhat the sequence of events. Dressing a sword was reported to enable fighting, because such activities were observed to occur immediately before fighting. The definition of 'smite' had to be revised because the sequence of events in the story had Arthur acting after Lot smote him, which would have been impossible had the smiting resulted in his death. To be generally effective at defining verbs, however, Cassie's capacity for inferring causal or enabling relations between actions and events should be expanded.

Looking for textual markers is an obvious step in deducing causality from narrative. 'So [modifier] that' is a proverbial cause/effect marker, as are 'therefore', 'because', and 'since'. Other such markers exist. Our grammar should recognize at least some of these, and build the appropriate relations between the propositions of the surrounding clauses. Temporal terms like 'until' can also be clues. If one activity or state of affairs continues until some event occurs (or other state obtains), the new event or state may cause the termination of the previous activity or state. In the case of a terminated activity, it may end because its goal is accomplished, or because the activity is no longer possible. The related work of SNeRG members Almeida (1995) on time in narratives and Ismail & Shapiro (2000) on first-person aspect may be of some use here, as would Cohen's (1987, 1990) work on textual markers and Mann & Thompson's work on rhetorical structure theory.

Sequence of action may be considered in terms of whether it suggests a causal or enabling link. In general, an act by one agent followed by the act of a different agent is not apt to be causal or enabling unless a textual marker is present. In defining 'dress', we used the heuristic that, when the second verb in a time sequence is a second action by the same agent, the first act may enable the second, or it may enable a larger activity of which the second act is a part. The difficulty with this heuristic is that a list of actions may be effectively simultaneous, if they are done to enable something else.

To a large extent, inferring causal or enabling relations in narrative is a matter of having sufficient background knowledge, and does not require changes in our definition algorithms. However, determining which verbs (and how many) form a sequence will require some type of discourse analysis that Cassie currently lacks. A better understanding of a story can only improve the understanding of the words in it.

The three major types of open-class words are nouns, verbs, and modifiers. We need to abstract definitions for adjectives or adverbs, and consider methods for defining proper names. For some adjectives, morphological information might be of use (see §4.1.3). Definitions of color terms might be facilitated using the techniques from related work on color recognition and naming (Lammens 1994). Proper names, of interest in applications to reading news items, ought to be handleable by extending our noun-definition algorithm (perhaps to include slots for city/state/country, function, occupation, etc.).

4.1.3 Develop algorithms to deal with synonyms and etymological/morphological clues for definitions. Readers sometimes look for clues to the meaning of a new word in the word itself. Some such clues can be very misleading (e.g., 'brachet' rhymes with 'latchet' so perhaps it refers to some sort of hardware or fastener), but etymological clues may be quite relevant. In some cases, a familiarity with Greek, Latin, or other roots can allow a reasonable interpretation of a new word even if the context is not very informative. If combined with an informative context, such etymological knowledge can permit very rapid acquisition of a word's meaning. SNePS's morphological analyzer gives some of this information, and we will explore adapting it for this use.

4.1.4 Modify the reasoning system to allow it to induce generalizations from instances and to fully automate the belief-revision mechanism. We plan to investigate getting Cassie to *generalize* in situations such as that encountered with 'dress' (§3.2, above): Since getting dressed is also a form of preparation (as is the use of salad "dressing"), Cassie should be able to *induce* a more general meaning (while maintaining such everyday meanings as putting clothes on). The automation of the belief-revision mechanism has already been discussed (§3.3, above).

4.1.5 Decide whether, when, and how to store newly learned definitions. It appears that humans do not store inferred definitions (Johnson-Laird 1987). Rather, we link various aspects of meaning to words, and different portions of a word's meaning "come to mind" depending on the context in which they are used. The exception is when we expect a need to produce that definition, as with a student who must memorize definitions for an exam. Our decision not to build definitions into the knowledge base seems, therefore, to be both cognitively valid and practical (since Cassie need not withdraw "belief" in an early stage of a definition once she learns more). However, every time we ask her for a definition, she must search out the information afresh even if she has just told us the same thing many times in a row. We currently have no mechanism for compiling a definition, or even of keeping track of how many times a definition has been reported. We are considering the possibility of allowing Cassie to come to believe that certain portions of her knowledge constitute a definition. It seems likely that we would want both completeness and stability before storing a definition. The stability criterion, however, would require the development of some method for keeping track of how often a particular definition has been presented.

4.1.6 Foundational issues to be addressed. (1) How much assistance does context actually give? And how does it actually facilitate vocabulary acquisition? Clearly, some contexts will be more useful than others; perhaps some will be either useless or misleading. Our techniques are primarily developed with the useful contexts in mind. But they should work with useless (more technically, "minimal" or "empty") contexts, too: Consider the sentence "Tommy broke a **weti**", with unknown word 'weti'. With some reasonable, but minimal, background knowledge, we might theorize that a weti is a breakable physical object. But even with no background knowledge or other contextual information (other than grammatical structure), we could theorize that a weti is something that Tommy broke, by "solving" the sentence for its unknown term, as in algebra (Higginbotham 1989). More importantly, errors inferred from misleading contexts will eventually be corrected as more occurrences of the word are encountered. More generally, our research is aimed at answering questions about the value and nature of contextual assistance to vocabulary acquisition; our algorithms should be considered as (implementing) a *theory* of how context operates.

(2) Does context operate differently in different subject areas ("domains")? This important *empirical* question can only be answered by testing our system on texts from different subject areas. Our research to date has been primarily on literary texts. The current proposal focuses on SMET texts. As discussed in §4.2.1(1), SMET authors should, in

principle, write in a way that facilitates CVA, whereas literary authors might not be expected to (though, clearly, our results to date suggest that sometimes they do). We hope that one outcome of our research will be insight into the nature of such different contexts, resulting, perhaps, in advice to authors who wish to write CVA-facilitative texts (see Project 2061 (1999, 2000) for discussion of some problems with recent SMET textbook writing).

4.2 Educational Stream: Curriculum Development.

We have endeavored to incorporate into Cassie, or be consistent with, research in psycholinguistics on language acquisition, notably the work already cited in §§3.2, 3.3. The goal is not to limit Cassie's CVA strategies to just those strategies already taught by teachers or used by students; no significant educational benefit results from this. The goal is for Cassie to use not only those strategies commonly taught and used, but to create new, previously unused or unknown algorithms. Successfully programming Cassie for CVA will be a worthy accomplishment, but the significance is not Cassie's adequate performance, but whether the CVA strategies Cassie generates benefit the teaching of CVA so students are better able to "compute" word meanings when they independently read SMET materials. Accordingly, the parallel educational stream will have three phases:

4.2.1 Phase 1: Instructional methods and case studies. Three tasks are required in Phase 1 of the educational stream: (1) finding and/or developing SMET texts for CVA; (2) conducting case studies of good and poor readers to analyze in depth the forms of context, prior knowledge, and the logical operations used or applied during CVA; and (3) extending Cassie's computational approaches and background knowledge in the analysis of SMET texts found (task 1) with the findings, data, and insights gained from the case studies (task 2).

1. Development of instructional materials for guided instruction and application. SMET texts are especially appropriate for deliberate CVA techniques: First, they cannot (or, at least, should not) be read quickly and passively, but slowly and actively, a technique that requires each sentence to be understood before the next one is read (Rapaport 2000a). A sentence with an unknown term (or a familiar term used in an unfamiliar, technical sense) that is not initially understood will require conscious, deliberate, and explicit thinking and reasoning on the part of the reader to develop at least an initial hypothesis about its meaning. Consistent with our theory, further encounters with the term will force the reader to revise the hypotheses and ultimately determine *a* meaning (not necessarily "the" meaning) that will foster overall comprehension. Nagy 1997 has advocated the importance of comprehension as the ultimate goal of reading over merely determining the "correct" meaning of a term. A related issue to be explored is student motivation: Poorer readers tend to skip over unknown words (Reed 1953). We believe that the methods of slow and active reading can help such readers notice such words and decide whether (and, of course, how) to figure out their meaning.

Second, authors of SMET texts, unlike the literary texts of our research so far, have the intention of helping the reader learn new words. Of course, they do not always live up to these good intentions; many authors use words presuming that the reader already knows their meaning, rather than facilitating the reader's learning of their meaning. One anticipated outcome of our research is guidelines for writing SMET texts that are conducive to CVA techniques.

A major task in Phase 1 will be finding, adapting, and/or developing suitable SMET texts: i.e., text that contains a limited number of hard words, but uses those hard words in a variety of contexts numerous times throughout the text. These texts will be used both by Cassie for development of CVA algorithms, will become part of the CVA curriculum to be developed, and will be used by teachers and students for instruction and application. The two primary sources of texts are: (a) textbooks and (b) other reading materials used in SMET instruction. As a backup and only when necessary, a third source could be (c) materials written especially for this project covering the same content as (a) and (b). The major purpose of this project is to enhance students' CVA abilities as they read the instructional materials regularly used in their SMET instruction, not to prepare special texts designed to facilitate CVA. But it may be the case that a given textbook provides too little context for CVA; in such cases, we will need to prepare additional texts that do provide additional context for CVA and that will be used in SMET instruction as a supplement to the regular classroom text.

Our plan is to use SMET texts for grades 5–8 and introductory computer science texts for college students. Readers at grade levels 5–8 are the most likely to come across new words, whereas adult readers seldom do (Durkin 1978/1979, Nagy 1988, Harris & Sipay 1990). Besides their intrinsic importance, we have chosen these subjects and ages for convenience, since the PIs have access to, and have experience teaching, these subjects to these ages. We have examined one of the science texts (Cooney et al. 2000) currently used by a cooperating 6th-grade teacher. This text presents hard science words repeatedly throughout a unit, and is ideally suited for this project. We are continuing to investigate other SMET texts for grades 5–8.

2. Case studies for developing CVA strategies. The literature of cognitive psychology, computational linguistics, and vocabulary research will inform the development of Cassie's algorithms, but we contend that this process should also be informed by the strategies and skills real students apply when trying to gain a word's meaning from context. It will be the junction of logic and theory with practice and application that will lead to developing a new perspective on the nature of context for CVA, the role of prior knowledge in this learning, the mental processes required to use context, and the number of encounters with a word in context required to learn that word.

Students of varying reading and intellectual abilities will be taught CVA strategies in one-to-one and small-group settings. Phase 1 will involve three sets of students:

- 4–6 students in grades 5–8, with reading difficulties who are receiving one-to-one remedial reading instruction in the Center for Literacy and Reading Instruction. These students will score at least two grade-equivalent levels below their grade placement on the Gates-MacGinitie Reading Achievement Test, 4th Ed. (Gates et al. 2000) on both the Comprehension and Vocabulary subtests. This is the standard definition for a reading difficulty for children of average intellectual ability (Harris & Sipay 1990). To be sure that each student has average or better listening vocabularies, each will have a score on the Peabody Picture Vocabulary Test, 3rd Ed (PPVT-III) (Dunn & Dunn 1997) not below the 46th percentile level.

- 2–4 students in grades 5–8 who are excellent readers will receive one-to-one instruction in their school. On both the Gates-MacGinitie and the PPVT-III, each will score at or above the 80th percentile-insuring that all are excellent readers with strong listening vocabularies.

- 6–10 students in grades 5–8 with scores at or above the 80th percentile level on both the Gates-MacGinitie and the PPVT-III will be taught in small groups of 3–5.

Data in the Phase 1 stage will be gained in the case studies by four methods:

(a) Observations of teaching. One-to-one or small group instruction will be videotaped as teachers teach CVA strategies to students and assist students in applying these techniques in texts. Observations and follow-up interviews will focus on: (1) specific contexts teachers point out as useful for CVA; (2) cognitive operations teachers guide students in using to deduce a meaning for a hard word; (3) prior knowledge necessary for deducing a meaning of a hard word; and (4) fix-up strategies teachers employ when students do not derive a plausible meaning for the hard word (e.g., redirecting student, providing information presumed known by student but not known, assistance given in applying cognitive operations).

(b) Structured interviews of teachers. Following instruction, as researcher and teacher review the videotape of the lesson, the teacher will be given a structured interview (Seidman 1991; Preissle & LeCompte 1993) about her or his perceptions on the forms of the contexts used, the cognitive operations applied, the role of students' prior knowledge in CVA, and insights on improving these components.

(c) Observations of students. During instruction, students will use Kucan & Beck's (1997) think-aloud technique as they apply CVA strategies to hard words in texts. The focus here will be students' uses, comments, and insights on varying forms of context, prior knowledge they relied upon, logical operations used, and their confidence in the correctness of the meaning they deduced for the hard word.

(d) Structured Interviews of Students. The purpose and format of the student interviews parallel those for the teacher interviews, but focusing on students' perceptions and observations. Varying information might be obtained from students participating in a one-on-one structured interview and students being interviewed in a small group; which is why both one-on-one and small-group instruction and interviews with excellent students will be employed.

Case-Study Procedures. All instruction will be conducted by experienced reading clinicians/reading specialists in the Center for Literacy and Reading Instruction as well as in public-school classrooms in Western New York. All structured interviews will be conducted by a researcher or research assistant. The materials used in the instruction will be adapted from prior research (e.g., Beck et al. 1983; Friedland 1992), published materials for teaching context skills (e.g., Johnson & Pearson 1984; Mueser 1984; Mueser & Mueser 1984), materials currently used in the students' classrooms, and drafts of materials being developed (described in prior section). Parental consent forms (as approved by SUNY Buffalo's Human Subjects Review Board) will be obtained for all students included at this and every stage of the research. All teachers will be provided workshops on theories of vocabulary learning, research on the teaching of CVA strategies, and specific context and context teaching methods that have been found useful (Deighton 1978, Beck et al. 1983, Johnson & Pearson 1984, Fisher 1986, Nagy 1988, Friedland 1992, Kibby 1995, Stahl 1999). Code sheets for observations and structured interviews will be created. The major foci of these code sheets will be to note in detail the varying forms of context used, the parameters of prior information needed to use the context, and the thinking or cognitive processes students used to form a meaning for a hard word. This information will not only inform the cognitive computer scientists for their programming of Cassie (the major purpose of this part of Phase 1), but will also inform the preparation of curricular methods and materials (Phase 2).

3. Extending Cassie's computational approaches and background knowledge. A major purpose of the in-depth case-study observations and interviews of excellent and poor readers learning and applying CVA techniques is to gain data and insights to inform the cognitive computer scientists who are programming Cassie about (1) the forms of context teachers point out and students use, (2) the prior knowledge required to use this context and what teachers do to assist students with this prior knowledge, and (3) the cognitive or logical operations taught by teachers and applied by students. (These data will also inform the reading scientists in developing the CVA curriculum.)

Cassie will have to be supplied with suitable background knowledge to enable her to understand (all but the unknown terms of) these texts (Nagy 1997), and this will be another central task of our project. Many education researchers only pay lip service to the need for prior knowledge, but artificial-intelligence researchers in the fields of computational linguistics and knowledge representation are well aware of the importance of background knowledge for understanding. As Kibby 1995 points out, prior knowledge, suitably organized, of the concepts expressed by unknown words is of immense help; conversely, learning new words can help (re-)organize the reader's background knowledge. Thus, a student's failure to learn vocabulary contextually might reflect, not so much a failure to correctly apply CVA inference techniques, but a gap in background knowledge or knowledge that is not organized in a way conducive to reasoning. Cassie's explanations, then, will have to include not only strategic advice, but also, where necessary, factual information to fill in (or to show the student how to fill in) any gaps in a usefully organized way. The case studies will provide data on what prior knowledge students used and what prior knowledge teachers were required to teach in order for the students to apply CVA techniques successfully. These data will inform the cognitive computer scientists in programming Cassie.

Similarly, Cassie will need to develop a repertoire cognitive or logical operations for using available context and available prior knowledge for CVA. Linguistic theory and logical analyses of language and texts will inform these processes, but, in this study, the cognitive computer scientists will also be informed by the variety of logical operations gained from the case studies.

4.2.2 Phase 2: Curriculum development and evaluation. There are three major tasks for Phase 2 of the educational stream: (1) developing the CVA curriculum and instructional materials, (2) developing CVA evaluation instruments, and (3) clinical and field trials of the curricular materials.

1. Development of teaching methods and instructional guidelines. After finding suitable texts and having Cassie apply CVA techniques to those texts, the next step will be to coordinate these texts and the teaching methods into a curriculum for teaching CVA techniques. This curriculum will be designed to help teachers identify the prior knowledge students will need to derive a given word's meaning from context as well as methods for teaching students to identify and apply contextual clues found in natural texts. Our review of the research yields a number of factors that affect students' CVA that must be addressed in the design of our curriculum and instructional materials:

(a) Noting hard words in the text. Students are not always cognizant that they have even encountered a hard word in a text (Reed 1953). Most previous studies of CVA (Friedland 1992) have used some form of typographical cue to be sure subjects noted the hard words in texts (Friedland 1992). Therefore, most of the instructional and evaluation materials developed will use a typographic cue that ensures the reader affixes attention to the hard word. To determine if students apply CVA strategies to hard words in natural texts, a few materials will not contain these typographical cues.

(b) Identifying context clues. A reader may not attempt to search for context that might provide a hard word's meaning. To determine what textual clues and prior knowledge do or do not facilitate CVA, it is necessary to control the degree of effort students must expend in applying CVA techniques. Therefore, when confronting a hard word in a text, the students in this study will be required to identify those contexts that might provide some degree of meaning for each hard word they encounter.

(c) Evaluating helpfulness of context clues. Context does not always facilitate CVA (Beck et al. 1983; Schatz & Baldwin 1986). It is important that students monitor or judge the helpfulness of the context for CVA as a way of evaluating the utility of context and monitoring their confidence in the meaning derived (see (f), below). Beck et al. (1983) have described four levels of "helpfulness" of context in providing information for CVA to the reader: directive (very helpful), generally directive (somewhat helpful), nondirective (not helpful at all), and misdirective (the context leads reader to an incorrect definition of the hard word). In this study, students will be required to rate the helpfulness of the context .

(d) Prior knowledge. Even if a reader notes hard words and looks for the context clues to gain a meaning, the reader may or may not have the prior knowledge to use that context. Therefore, every student will have a criterion level of prior knowledge. Those students who do not already possess the appropriate prior knowledge will be provided that knowledge by the teacher or other instructional materials.

(e) Attempting to define hard word. Even if the reader recognizes that the text contains hard words, the reader may have gained some comprehension from the text, and thus feels no need to try to gain the word's meaning from context (August, Flavell, & Clift 1984). Therefore, we must require all students to make an attempt to supply or pick a meaning for every hard word encountered.

(f) Monitoring appropriateness of derived meaning. Stahl (1999: 29) states that one needs to "sensitize children to the importance of learning words from context", and he suggests that students be required to monitor or rate their confidence in the meaning derived for a hard word from context. Blachowicz & Fisher (1996) and Dale & O'Rourke (1986) suggest using a knowledge rating checklist for this rating. Therefore, after students derive a meaning for a hard word from context, we will require them to rate their confidence in that meaning.

(g) Number of encounters with the hard word in context. No one expects anyone to gain a hard word's "appropriate meaning" in one encounter. Beck et al. (1983) say at least 12 encounters are needed. Therefore, all experimental and control students must have multiple (e.g., 12–16) encounters with each hard word.

In sum, our teaching methods and practice materials will be administered with certain limitations, i.e., we will: force students to note the hard word, require students to identify contexts that might help define the hard word, require students to evaluate the helpfulness of the context for CVA, assure that each student has the prior knowledge needed to use the available context and understand the hard word, require students to produce some meaning for the hard word, require them to evaluate their confidence in that meaning, and require students to encounter the use of the hard word in multiple texts.

2. Development of evaluation instruments.

(a) Measures of the dependent variables. In the experimental phase of this study (Phase 3), students in grades 5–8 studying SMET and college undergraduates studying introductory computer science will serve as the subjects in our experiments, and they will or will not receive experimental CVA program. At the end of the experimental treatments, it is predicted that experimental students taught CVA strategies will exceed control students on seven criteria: (1) number of the hard words noted in the texts they are reading, (2) number of segments of available context found useful for CVA, (3) judging the helpfulness of the available context, (4) number of hard words defined, (5) judging confidence in the accuracy of the derived meaning, (6) number of word definitions that closely approximate correctness, and (7) number of accurately defined words. Dependent variables 6 and 7 are the major dependent variables of the experimental studies. Test instruments and structured interview techniques will be developed to measure each of these outcome variables. The administration of these measures is described in §4.2.3(2).

(b) Measures of transfer. Students applying CVA strategies under the instructional conditions we use here may or may not apply them on their own at a later time. Our research aims to determine if students who use CVA strategies in experimental environments continue to apply these techniques to hard words encountered in their natural reading environments. Therefore, we will administer follow-up transfer tests to see if our students continue to apply CVA strategies in their required reading.

3. Refinement of Curriculum and Evaluation Instruments. There will be three field-study try-outs of curriculum and evaluation instruments. The try-outs will be with individual students in the Center for Literacy and Reading Instruction and with middle-grade and college students. None of the students in these try-outs will be a part of the controlled experiments, though it is our plan that the teachers might be the same. At the same time, Cassie will be continuing to evaluate our texts and to help us refine our teaching methods and materials.

4.2.3 Phase 3: Controlled Experiments.

1. Subjects and Treatments. Two pools of students will be the Ss for evaluating the effectiveness of the CVA techniques developed with input from Cassie: middle-grade students (grades 5–8) in regular classrooms in elementary or middle schools in two Western New York public-school districts, and college students (SUNY Buffalo undergraduates in an introductory computer science course).

Normal instruction in any SMET content area demands that teachers discuss the hard words in the text in some manner. Our instruction tries to coincide with normal teaching of SMET as much as possible. To have a control or treatment group in which hard words were never discussed in class would be completely detached from normal teaching; therefore, all instructional treatments will include some form of discussion of new or hard vocabulary. There will be three instructional treatments: experimental CVA strategies, dictionary/glossary, and control. In the experimental treatment, the regular classroom instructors will provide instruction and practice in applying the CVA strategies developed in Phases 1 and 2 of the study. In the dictionary/glossary methods, teachers will remind (but not require) students to use a dictionary or textbook glossary for definitions of hard words in their SMET texts. In the control group, no special instructions, teaching, or materials will be given to students for learning the hard words. In all three groups, teachers and students will orally discuss any hard word whenever students ask about a word's meaning. Systematic reviews of CVA strategies will be carried out during the implementation of the CVA treatment, and systematic reminders to use dictionaries and glossaries will be given in the dictionary/glossary treatment.

2. Procedures. The experiment will have four phases: (a) teacher training, (b) 4–6 weeks of student training and practice in the instructional method (except the Control group), (c) the experimental treatments and posttest, and (d) the follow-up test and study:

(a) Teacher Training: The instructors for the three treatment groups will be regular middle-grade and university instructors. In order to control for Hawthorne and John Henry effects, all instructors will participate in pre-treatment instructional sessions. All groups will receive sessions covering concept development, direct instruction, dictionaries usage, word play, semantic webbing, semantic features, synonym webs, word maps, word origins, Word Wizard, and roots and affixes. The instructional program will be based on works such as Deighton 1978, Gipe 1980, Eeds & Cockrum 1985, McKeown 1985, Schwartz & Raphael 1985, Beck et al. 1987, Nagy 1988, Blachowicz & Fisher 1996, Stahl 1999, and Johnson 2000.

Those teaching the CVA treatment will be given additional sessions on how to help students learn the CVA strategies specified in the CVA curriculum developed in Phase 2, and shown how to use the CVA instructional materials. For those employing the dictionary/glossary method, additional sessions on dictionary techniques for teaching vocabulary will be given based especially on the guidelines of Deighton 1978, Blachowicz & Fisher 1996, and Johnson 2000.

(b) Student Training: After teacher training, the CVA group in grades 5–8 will be given 4–6 weeks of direct instruction of the CVA curriculum developed in Phase 2. As previously described in Phase 2, students will be required to (a) note hard words, (b) identify and (c) evaluate context clues, (d) identify appropriate prior knowledge, (e) attempt to define the word, and (f) evaluate their confidence in the deduced meaning. In addition, they will (g) encounter many of the hard words frequently during training. Beck et al. (1987: 150) recommend that all activities "encourage children to make their thinking explicit". To do this, most instruction will be with large or small groups for sharing of insights and strategies. We plan for very little independent or "seat work" activities during these sessions for teaching the CVA curriculum.

Using the practice texts used in the CVA treatment, students in the dictionary/glossary method will be given 4–6 weeks of instruction on the use of dictionaries, glossaries, and finding correct definitions for the hard words in the texts. Most of the students will have already developed adequate dictionary skills, but it is important to the integrity of the experiment that this group go through the same practice texts as the CVA group (though with no CVA instruction) and be encouraged to focus on the meanings of hard words in SMET texts for the same length of time as the CVA group.

The control group will receive no special directions or instruction during the training period.

For the college students in computer science, it is anticipated that the student training can be condensed to 3–6 hours.

(c) Experimental Treatments and Posttests: Each treatment will run for 8–10 weeks for both the middle-grades and university. The texts used in these settings will be those ordinarily used, though Cassie will have already evaluated them, and they will have been tried out in Phases 1 and 2. A goal of Phases 1 and 2 is to determine how many encounters with a word are typically required for an approximately correct meaning of the word to be determined. In Phase 3, as students progress through the text and encounter hard words the number of times determined in Phases 1 and 2, they will be tested on the meanings of those words on a weekly test (the "immediate retention test"). We anticipate each week's test to evaluate 4–10 words. Each word will be tested only once on the immediate retention tests. A week after the chapter or instructional unit has been completed, a posttest will be administered that tests knowledge of all the hard words in that chapter or unit of the text (the "delayed test"). These two measurements will assess the last two dependent variables (numbers of word definitions approximately correct and of hard words accurately defined; §4.2.2(2a)). In addition to tests of word meanings, tests of content knowledge of the units studied will also be administered during and at the end of the unit of study (the "content test").

(d) Follow-Up: One month after the treatments, as a follow-up to determine if students taught CVA strategies still apply them in their studies of SMET texts, we will conduct structured interviews with a representative sample of students to determine what strategies they use as they encounter a hard word in a text. Ss will be interviewed individually

and assigned a section from their text containing several hard words. After reading the section of text, they will be asked to: (1) list the hard words noted in the text, (2) identify all segments of available context found useful for CVA, (3) judge the helpfulness of the available context, (4) give a meaning for the word, and (5) judge confidence in the accuracy of the derived meaning (the first five dependent variables identified in §4.2.2(2a)). From this, measures of the last two dependent variables can be gained.

3. Data Analysis. Because of the difference between middle-grade and college students in experience with CVA, we will analyze the data separately. For the middle-grade students, our research will have a 2 x 3 x 2 x 2 factorial mixed-design with repeated measures on the last factor. The between-subjects factor are intelligence (above average and average/below average); method of vocabulary instruction (context, dictionary/glossary, and control); and subject (science or technology); the fourth, the within-subjects factor, is time of testing (immediate and delayed). The dependent variables in this quantitative analysis will be content knowledge and number of words learned (§4.2.2(2)). If we find no significant main or interaction effects for subject, we will collapse the analysis to a 3-way mixed ANOVA. Only a 3-way design will be used with the college students, because we are teaching only one subject.

The data from the structured interviews of students after the treatments will be analyzed quantitatively with percentages and qualitatively (Seidman 1991). Teachers will also be interviewed regarding their reactions to the CVA methods as they applied them in their classes.

It is anticipated that the findings of this study will not only inform classroom methods of teaching CVA to students studying SMET texts, but it is also expected that we will be able to derive guidelines for writing SMET texts that will assist authors in presenting their text in a manner that enhances CVA.

5 EFFECT ON EDUCATION AND ON HUMAN RESOURCES. The project team will consist of William J. Rapaport (Associate Professor of Computer Science, Adjunct Professor of Philosophy, and Associate Director, SNeRG) and Michael W. Kibby (Professor, Department of Learning and Instruction; Director, Center for Literacy and Reading Instruction) as co-PIs, and their graduate students in the Department of Computer Science & Engineering, the Department of Learning and Instruction, the Center for Cognitive Science, and the Center for Literacy and Reading Instruction. The overall theory of vocabulary expansion by "syntactic-semantic" techniques was developed by Rapaport, and the details of the theory and its implementation were developed by Karen Ehrlich (now in the Department of Mathematics and Computer Science, SUNY College at Fredonia), who will consult on this project. The topics have been the focus of graduate seminars taught by Rapaport, which will continue, along with joint seminars with Kibby. The development of SNePS is being conducted by a larger research group of M.S. and Ph.D. students, co-directed by Stuart C. Shapiro and Rapaport (with the participation of Jean-Pierre Koenig, Department of Linguistics), of which this project will be an integral component. We plan to involve other faculty and graduate and undergraduate students in various capacities on this project, ranging from fairly straightforward lexicon and grammar development to research on, and implementation of, the proposed extensions to the theory, and testing and training with readers. This will include experts at SUNY Buffalo on SMET education: Douglas H. Clements (computer education), Rodney L. Doran (science education), Carol Hosenfeld (L2 education), Thomas Schroeder (mathematics education). In addition to researchers from the PIs' home departments, we expect to involve other researchers and students from the SUNY Buffalo Center for Cognitive Science in communicative disorders, linguistics, psychology, and/or philosophy.

6 SUMMARY. We have preliminary evidence that such a computational CVA system can be built. We seek funds to explore this further, to develop a CVA curriculum, and to perform experiments (both in the lab and in classroom settings in the school systems) to see if it helps students learn new words and concepts. If it is successful, we hope that it will be acceptable to school systems. We believe that our project is consistent with the goal of improving SMET educational practices in both classroom and informal learning environments, by focusing on concept formation, acquisition, and change in novel domains, in a way that integrates the reader's existing background knowledge with new information (i.e., the unknown term).

Section D. References Cited

1. Aist, Greg (2000), “Helping Children Learn Vocabulary during Computer Assisted Oral Reading”, *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000; Austin, TX)* (Menlo Park, CA, & Cambridge, MA: AAAI Press/MIT Press): 1100–1101.

2. Ali, Syed S., & Shapiro, Stuart C. (1993), “Natural Language Processing Using a Propositional Semantic Network with Structured Variables,” *Minds and Machines* 3: 421–451.

3. Almeida, Michael J. (1995), “Time in Narratives,” in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 159–189.

4. Anderson, R.C., & Freebody, P. (1981), “Vocabulary Knowledge”, in John Guthrie (ed.), *Comprehension and Teaching: Research Reviews* (Newark, DE: International Reading Association): 77–117.

5. Asker, L.; Gamback, B.; & Samuelsson, C. (1992), “EBL^2: An Approach to Automatic Lexical Acquisition” *COLING-92*: 1172–1176.

6. August, D.L.; Flavell, J.H.; & Clift, R. (1984), “Comparison of Comprehension Monitoring of Skilled and Less Skilled Readers”, *Reading Research Quarterly* 20(1): 39–53l.

7. Beck, I.J.; McKeown, M.; & McCaslin, E. (1983), “All Contexts Are Not Equal”, *Elementary School Journal* 83: 177–181.

8. Beck, I.J.; McKeown, M.G.; & Omanson, R.C. (1987), “The Effects and Use of Diverse Vocabulary Instructional Techniques”, in M.G. McKeown & Mary Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 147–163.

9. Berwick, Robert C. (1983), “Learning Word Meanings from Examples”, *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI-83; Karlsruhe, West Germany)* (Los Altos, CA: William Kaufmann): 459–461.

10. Blachowicz, C., & Fisher, P. (1996), *Teaching Vocabulary in All Classrooms* (Englewood Cliffs, NJ: Merrill).

11. Carbonell, Jaime G., & Hayes, Patrick J. (1984), “Coping with Extragrammaticality, ” *COLING-84*: 433–437.

12. Carter, D. (1989), “Lexical Acquisition in the Core Language Engine”, *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics*: 137–144.

13. Cohen, Robin (1987), “Analyzing the Structure of Argumentative Discourse,” *Computational Linguistics* 13: 1–10.

14. Cohen, Robin (1990), “Implementing a Model for Understanding Goal-Oriented Discourse,” *International Journal of Intelligent Systems* 5: 1–31.

15. Cooney, T., et al. (2000), *Scott Foresman Science* (Glenview, IL: Scott Foresman).

16. Copestake, Ann (1999), “The LKB System”, [http://www-csli.stanford.edu/˜aac/lkb.html].

17. Cravo, Maria R., & Martins, João P. (1993), “SNePSwD: A Newcomer to the SNePS Family”, *Journal of Experimental and Theoretical Artificial Intelligence* 5: 135–148.

18. Dale, E., & O’Rourke, J. (1986), *Vocabulary Learning* (Columbus, OH: Zaner-Bloser).

19. Deighton, L.C. (1978), *Vocabulary Development in the Classroom* (New York: Teachers College Press).

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

21. Dunn, L.M., & Dunn, L.M. (1997), *Peabody Picture Vocabulary Test; 3rd Edition* (Circle Pines, MN: American Guidance Service).

22. Durkin, D. (1978/1979), "What Classroom Observations Reveal about Reading Comprehension", *Reading Research Quarterly* 14(4): 481–533.

23. Eeds, M., & Cockrum, W.A. (1985), "Teaching Word Meanings by Extending Schemata vs. Dictionary Work vs. Reading in Context", *Journal of Reading* 28(6): 492–497.

24. Ehrlich, Karen (1995), "Automatic Vocabulary Expansion through Narrative Context", *Technical Report 95-09* (Buffalo: SUNY Buffalo Department of Computer Science).

25. 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.

26. Elshout-Mohr, Marianne, & van Daalen-Kapteijns, Maartje M. (1987), "Cognitive Processes in Learning Word Meanings", in Margaret G. McKeown & Mary E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 53–71.

27. Fisher, P.L. (1986), *An Investigation of Selected Factors Affecting Readers' Understanding of Unfamiliar Words in Text*, doctoral dissertation (SUNY Buffalo Faculty of Educational Studies).

28. Friedland, E.S. (1992), *The Effect of Context Instruction on the Vocabulary Acquisition of College Students*, Ph.D. dissertation (SUNY Buffalo Faculty of Educational Studies).

29. Gates, W.H.; Gates, R.K.; Maria, K.; & Dreyer, L.G. (2000). *Gates-MacGinitie Reading Tests, 4th Edition* (Chicago: Riverside Publishing).

30. Gentner, Dedre (1981), "Some Interesting Differences between Nouns and Verbs", *Cognition and Brain Theory*, 4: 161–178.

31. Gipe, J.P. (1980), "Use of a Relevant Context Helps Kids Learn New Word Meanings", *The Reading Teacher* 33(4): 398–402.

32. Gleitman, Lila (1990), "The Structural Sources of Verb Meanings", *Language Acquisition* 1: 1–55.

33. Gleitman, Lila (1994), "A Picture is Worth a Thousand Words—But That's the Problem" (talk presented at the 1st International Summer Institute in Cognitive Science, SUNY Buffalo, July 1994); abstract in *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum Associates): 965.

34. Granger, Richard H. (1977), "Foul-Up: A Program that Figures Out Meanings of Words from Context", *Proceedings of the 5th International Joint Conference on Artificial Intelligence (IJCAI-77, MIT)* (Los Altos, CA: William Kaufmann): 67–68.

35. Granger, Richard H. (1983), "The NOMAD System: Expectation-Based Detection and Correction of Errors during Understanding of Syntactically and Semantically Ill-Formed Text", *American Journal of Computational Linguistics* 9: 188–196.

36. Haas, Norman, & Hendrix, Gary (1983), "Learning by Being Told: Acquiring Knowledge for Information Management", in Riszard S. Michalski, Jaime G. Carbonell, & Tom M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach* (Palo Alto, CA: Tioga Press): 405–428.

37. Haller, Susan M. (1993), "Planning for Intentions with Rhetorical Relations," *Proceedings of a Workshop on Intentionality and Structure in Discourse Relations (Ohio State University)* (Morristown, NJ: Association for Computational Linguistics): 23–26.

38. 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.

39. Harris, A.J., & Sipay, E.R. (1990), *How to Increase Reading Ability* (White Plains, NY: Longman).

40. Hastings, Peter M. (1994), "Automatic Acquisition of Word Meanings from Natural Language Contexts", Ph.D. disseration (Ann Arbor: University of Michigan Department of Computer Science and Engineering).

41. Hastings, Peter M., & Lytinen, Steven L. (1994a), "The Ups and Downs of Lexical Acquisition", *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94, Seattle)* (Menlo Park, CA: AAAI Press/MIT Press): 754–759.

42. Hastings, Peter M., & Lytinen, Steven L. (1994b), "Objects, Actions, Nouns, and Verbs", *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum Associates): 397–402.

43. Hearst, M.A. (1992), "Automatic Acquisition of Hyponyms from Large Text Corpora," *COLING-92*: 539–545.

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

45. Hunt, Alan, & Koplas, Geoffrey D. (1998), "Definitional Vocabulary Acquisition Using Natural Language Processing and a Dynamic Lexicon", *SNeRG Technical Note 29* (Buffalo: SUNY Buffalo Department of Computer Science, SNePS Research Group).

46. Ismail, Haythem O., & Shapiro, Stuart C. (2000), "Two Problems with Reasoning and Acting in Time", in A. G. Cohn et al. (eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the 7th International Conference* (Morgan Kaufmann): 355–365.

47. Johnson, D.D. (2000), *Vocabulary in the Elementary Classroom* (New York: Allyn & Bacon).

48. Johnson, D.D., & Pearson, P.D. (1984), *Teaching Reading Vocabulary, 2nd ed.)* (Holt, Rinehart, Winston).

49. Johnson, Frances L., & Shapiro, Stuart C. (2000a), "Formalizing a Deductively Open Belief Space", *Technical Report 2000-02* (Buffalo: SUNY Buffalo Department of Computer Science & Engineering).

50. Johnson, Frances L., & Shapiro, Stuart C. (2000b), "Implementing Integrity Constraints in an Existing Belief Revision System", in C. Baral & M. Truszczynski (eds.), *Proceedings of the 8th International Workshop on Non-Monotonic Reasoning*.

51. Johnson-Laird, Philip N. (1987), "The Mental Representation of the Meanings of Words", in Alvin I. Goldman (ed.), *Readings in Philosophy and Cognitive Science* (Cambridge, MA: MIT Press, 1993): 561–583.

52. Kibby, Michael W. (1995), "The Organization and Teaching of Things and the Words that Signify Them", *Journal of Adolescent and Adult Literacy* 39(3): 208–223.

53. Kiersey, David M. (1982), "Word Learning with Hierarchy Guided Inference", *Proceedings of the National Conference on Artificial Intelligence (AAAI-82, Pittsburgh)* (Los Altos, CA: Morgan Kaufmann): 172–178.

54. Kucan, L., & Beck, I. (1997), "Thinking Aloud and Reading Comprehension Research: Inquiry, Instruction, and Social Interaction", *Review of Educational Research* 67(3): 271–299.

55. Lammens, Johan (1994), "A Computational Model of Color Perception and Color Naming", *Technical Report 94-26* (Buffalo: SUNY Buffalo Department of Computer Science).

56. Lipson, M.Y., & Wixon, K.K. (1997), *Assessment and Instruction of Reading and Writing Disability* (New York: Longman).

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

58. Mann, William, & Thompson, S. (1986), "Rhetorical Structure Theory: Description and Construction of Text Structures", in G. Kempen (ed.), *Natural Language Generation* (Boston: Kluwer): 279–300.

59. Martins, João, & Cravo, Maria R. (1991), "How to Change Your Mind", *Noûs* 25: 537–551.

60. Martins, João, & Shapiro, Stuart C. (1988), "A Model for Belief Revision", *Artificial Intelligence* 35: 25–79.

61. Marzano, Robert J.; Paynter, Diane E.; Kendall, John S.; Pickering, Debra; & Marzano, Lorraine (1991), *Literacy Plus: An Integrated Approach to Teaching Reading, Writing, Vocabulary, and Reasoning* (Columbus, OH: Zaner-Bloser).

62. McKeown, M.G. (1985), "The Acquisition of Word Meanings from Context by Children of High and Low Ability", *Reading Research Quarterly* 20: 482–496.

63. McRoy, Susan W.; Ali, Syed S.; & Haller, Susan M. (2000), "Mixed Depth Representations for Dialog Processing", in Łucja 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): 283–299.

64. Mueser, Anne Marie (1984), *Practicing Vocabulary in Context: Teacher's Guide* (Macmillan/McGraw-Hill).

65. Mueser, Anne Marie, & Mueser, John Alan (1984), *Practicing Vocabulary in Context* (SRA/McGraw-Hill).

66. Nagy, William E. (1988), *Teaching Vocabulary to Improve Reading Comprehension* (ERIC Clearinghouse on Reading and Communication Skills, National Council of Teachers of English, & International Reading Association).

67. Nagy, William E. (1997), "On the Role of Context in First- and Second-Language Vocabulary Learning", in Norbert Schmitt & Michael McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy* (Cambridge, UK: Cambridge University Press): 64–83.

68. Nagy, William E. & Herman, Patricia A. (1987), "Breadth and Depth of Vocabulary Knowledge: Implications for Acquisition and Instruction" in Margaret G. McKeown & Mary E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 19–35.

69. Neal, Jeannette G., & Shapiro, Stuart C. (1987), "Knowledge Based Parsing," in L. Bolc (ed.), *Natural Language Parsing Systems* (Berlin: Springer-Verlag): 49–92.

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

71. Peters, Sandra L., & Shapiro, Stuart C. (1987a), "A Representation for Natural Category Systems I," *Proceedings of the 9th Annual Conference of the Cognitive Science Society* : 379–390.

72. Peters, Sandra L., & Shapiro, Stuart C. (1987b), "A Representation for Natural Category Systems II," *IJCAI-87*: 140–146.

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

74. Preissle, J., & LeCompte, M.D. (1993), *Ethnography and Qualitative Design in Educational Research; 2nd Edition* (San Diego: Academic Press).

75. Project 2061 (1999), "Heavy Books: Light on Learning", [www.project2061.org/newsinfo/research/textbook/articles/heavy.htm].

76. Project 2061 (2000), "Algebra for All — Not with Today's Textbooks, Says AAAS", [www.project2061.org//newsinfo/press/r1000426.htm].

77. Pustejovsky, James (1987), "On the Acquisition of Lexical Entries: The Perceptual Origin of Thematic Relations," *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*: 172–178.

78. Quillian, M. Ross (1968), "Semantic Memory," in Marvin Minsky (ed.) *Semantic Information Processing* (Cambridge, MA: MIT Press): 227–270.

79. Quillian, M. Ross (1969), “The Teachable Language Comprehender: A Simulation Program and Theory of Language,” *Communications of the ACM* 12: 459–476.

80. Quine, Willard Van Orman (1951), “Two Dogmas of Empiricism,” in *From a Logical Point of View, 2nd Edition, Revised* (Cambridge, MA: Harvard University Press, 1980): 20–46.

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

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

83. Rapaport, William J. (1988), “Syntactic Semantics: Foundations of Computational Natural-Language Understanding”, in James H. Fetzer (ed.), *Aspects of Artificial Intelligence* (Dordrecht, Holland: Kluwer Academic Publishers): 81–131; reprinted in Eric Dietrich (ed.), *Thinking Computers and Virtual Persons: Essays on the Intentionality of Machines* (San Diego: Academic Press, 1994): 225–273.

84. Rapaport, William J. (1995), “Understanding Understanding: Syntactic Semantics and Computational Cognition”, in James E. Tomberlin (ed.), *AI, Connectionism, and Philosophical Psychology*, Philosophical Perspectives, Vol. 9 (Atascadero, CA: Ridgeview): 49–88; reprinted in Toribio, Josefa, & Clark, Andy (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.

85. Rapaport, William J. (2000a), “How to Study: A Brief Guide”, [www.cse.buffalo.edu/˜rapaport/howtostudy.html#readactively].

86. Rapaport, William J. (2000b), “How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition”, *Journal of Logic, Language, and Information* 9: 467–490.

87. Rapaport, William J. (2000c), “A Bibliography of (Computational) Theories of Contextual Vocabulary Acquisition”, [www.cse.buffalo.edu/˜rapaport/refs.vocab.html].

88. Rapaport, William J., & Ehrlich, Karen (2000), “A Computational Theory of Vocabulary Acquisition”, in Łucja 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.

89. Rapaport, William J.; Shapiro, Stuart C.; & Wiebe, Janyce M. (1997), “Quasi-Indexicals and Knowledge Reports”, *Cognitive Science* 21: 63–107.

90. Reed, J.C. (1953), “Developing Word Recognition and Meaning Vocabulary in College”, in Helen M. Robinson, *Corrective Reading in Classroom and Clinic* (Chicago: University of Chicago Press): 141–144.

91. Rosch, Eleanor (1978), “Principles of Categorization”, in Eleanor Rosch & Barbara B. Lloyd (eds.), *Cognition and Categorization* (Hillsdale, NJ: Lawrence Erlbaum Associates): 27–48.

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

93. Schatz, E.K., & Baldwin, R.S. (1986), ”Context Clues Are Unreliable Predictors of Word Meanings”, *Reading Research Quarterly* (Fall), 21(4): 439–453.

94. Schwartz, R.M., & Raphael, T. (1985), “Concept of Definition: A Key to Improving Students’ Vocabulary”, *The Reading Teacher* 39: 198–203.

95. Segal, Erwin M. (1995), “A Cognitive-Phenomenological Theory of Fictional Narrative”, in Judith F. Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 61–78.

96. Seidman, I.E. (1991), *Interviewing as Qualitative Research* (Teachers College Press).

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

98. Shapiro, Stuart C. (1989), "The CASSIE Projects: An Approach to Natural Language Competence," in João P. Martins & Ernesto M. Morgado (eds.), *Proceedings of the 4th Portuguese Conference on Artificial Intelligence (EPIA-89)*, *Lecture Notes in Artificial Intelligence*, No. 390 (Berlin: Springer-Verlag): 362–380.

99. Shapiro, Stuart C. (2000), "SNePS: A Logic for Natural Language Understanding and Commonsense Reasoning", in Łucja 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.

100. Shapiro, Stuart C. & The SNePS Implementation Group (1999), "SNePS 2.5 User's Manual", [www.cse.buffalo.edu/SNePS/Manuals/manual25.ps].

101. 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; revised version of the shorter version appears as "A Fully Intensional Propositional Semantic Network", in Leslie Burkholder (ed.), *Philosophy and the Computer* (Boulder, CO: Westview Press, 1992): 75–91.

102. 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.

103. Shapiro, Stuart C., & Rapaport, William J. (1992), "The SNePS Family", *Computers and Mathematics with Applications* 23: 243–275; reprinted in Fritz Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 243–275.

104. Shapiro, Stuart C., & Rapaport, William J. (1995), "An Introduction to a Computational Reader of Narratives", in Judith Felson Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 79–105.

105. Siskind, Jeffrey Mark (1992), "Naive Physics, Event Perception, Lexical Semantics, and Language Acquisition," Ph.D. dissertation (Cambridge, MA: MIT Department of Electrical Engineering & Computer Science).

106. Siskind, Jeffrey Mark (1994), "Lexical Acquisition in the Presence of Noise and Homonymy," *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94, Seattle)* (Menlo Park, CA: AAAI Press/MIT Press): 760–766.

107. Siskind, Jeffrey Mark (1996), "A Computational Study of Cross-Situational Techniques for Learning Word-to-Meaning Mappings", *Cognition* 61: 39–91.

108. Stahl, Steven A. (1999), *Vocabulary Development* (Cambridge, MA: Brookline Books).

109. Sternberg, Robert J. (1987), "Most Vocabulary is Learned from Context", in Margaret G. McKeown & Mary E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Hillsdale, NJ: Lawrence Erlbaum Associates): 89–105.

110. Sternberg, Robert J.; Powell, Janet S.; & Kaye, Daniel B. (1983), "Teaching Vocabulary-Building Skills: A Contextual Approach", in Alex Cherry Wilkinson (ed.), *Classroom Computers and Cognitive Science* (New York: Academic Press): 121–143.

111. *Webster's New International Dictionary of the English Language (2nd Edition, Unabridged)* (Merriam, 1937).

112. Wiebe, Janyce M. (1990), "Identifying Subjective Characters in Narrative," *Proceedings of the 16th International Conference on Computational Linguistics; COLING-90*: 401–408.

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

114. Wiebe, Janyce M. (1994), "Tracking Point of View in Narrative," *Computational Linguistics* 20: 233–287.

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

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

117. Wilks, Yorick, & Fass, Dan (1992), "The Preference Semantics Family," *Computers and Mathematics with Applications* 23: 205–221; reprinted in Fritz Lehmann (ed.), *Semantic Networks in Artificial Intelligence* (Oxford: Pergamon Press, 1992): 205–221.

118. Yuhan, Albert Hanyong, & Shapiro, Stuart C. (1995), "Computational Representation of Space," in Judith Felson Duchan, Gail A. Bruder, & Lynne E. Hewitt (eds.), *Deixis in Narrative: A Cognitive Science Perspective* (Hillsdale, NJ: Lawrence Erlbaum Associates): 191–225.

119. Zernik, Uri, & Dyer, Michael G. (1987), "The Self-Extending Phrasal Lexicon," *Computational Linguistics* 13: 308–327.