What Is the "Context" for Contextual Vocabulary Acquisition?

William J. Rapaport (rapaport@cse.buffalo.edu)
Department of Computer Science & Engineering, Department of Philosophy, Center for Cognitive Science
State University of New York at Buffalo, Buffalo, NY 14260-2000, USA
http://www.cse.buffalo.edu/∼rapaport/cva.html

Abstract

"Contextual" vocabulary acquisition is the active, deliberate acquisition of a meaning for a word in a text by reasoning from textual clues and prior knowledge, including language knowledge and hypotheses developed from prior encounters with the word, but without external sources of help such as dictionaries or people. But what is "context"? Is it just the surrounding text? Does it include the reader's background knowledge? I argue that the appropriate context for contextual vocabulary acquisition is the reader's "internalization" of the text "integrated" into the reader's "prior" knowledge via belief revision.

Introduction

"Contextual vocabulary acquisition" (CVA) is the active, deliberate acquisition of a meaning for a word in a text by reasoning from textual clues and prior knowledge, including language knowledge and hypotheses developed from prior encounters with the word, but without external sources of help such as dictionaries or people. It is the task faced by anyone coming upon an unknown word while reading, who has no outside source of help, but who needs to figure out a meaning for the word in order to understand the text being read.

My colleagues and I are engaged in a project with a dual goal: (1) To develop a computational theory of CVA (Ehrlich 1995, Ehrlich & Rapaport 1997, Rapaport & Ehrlich 2000) and (2) to adapt the strategies for doing CVA (embodied in our algorithms) to an educational curriculum for teaching CVA strategies to students in classroom settings (Rapaport & Kibby 2002). CVA has been investigated (though not hitherto in an integrated fashion) in AI, psychology, first- and second-language (L1, L2) acquisition, and reading education (see Rapaport 2003).

CVA is not restricted to fluent readers faced with a new word. Most of our vocabulary is acquired this way: People know the meanings of more words than they are explicitly taught, so they must have learned most of them as a by-product of reading or listening. The *average* number of word families (e.g., 'help', 'helps', 'helped', 'helping', 'helper', 'helpless', 'helpful' are one word family) known by high school graduates is estimated as at least 45,000. Learning this many words by age 18 means learning an average of some 2500 words each year; yet no more than 400 words per year are directly taught by teachers—4800 words in 12 years of school. Therefore, around 90% of the words we know and understand are learned from oral or written context. Learning words from context is not a once-in-a-while thing; it averages almost 8 words learned per day (Nagy & Anderson 1984).

Some of this "incidental" acquisition is the result of conscious, active processes of hypothesizing a meaning for unknown words from context. How do readers do this? The psychology, L1, and L2 literatures suggest various strategies (e.g., Ames 1966, Clarke & Nation 1980, Van Daalen-Kapteijns & Elshout-Mohr 1981, Sternberg et al. 1983, Kibby 1995, Blachowicz & Fisher 1996, Wesche & Paribakht 1999).

But most of these strategies are quite vague. E.g., Clarke & Nation 1980 gives these directions: (1) "look at the word itself and its surroundings to decide on the part of speech"; (2) "look at the immediate grammar context of the word, usually within a clause or sentence" (this presumably gives such information as who does what to whom, etc.); (3) "look at the wider context of the word usually beyond the level of the clause and often over several sentences" (looking for causal, temporal, class-membership information, etc.); (4) "*guess . . . the word* [my italics] and check . . . that the guess is correct". This is hardly a detailed algorithm that could easily be followed by a student: Step 4 is reminiscent of a famous cartoon showing a complicated mathematical formula, in the middle of which occurs the phrase, "then a miracle occurs"! (For another example, see Sternberg et al. 1983: 139–140.)

Although many authors suggest what contextual clues to look for (as in step 3, above—Sternberg et al. 1983 being the most helpful), few, if any, provide specific advice on what to *do* with them, i.e., what reasoning or other cognitive processes and what prior knowledge should be applied to them. Unfortunately, little (if any) of the computational research on the *formal* notion of reasoning within a context is directly relevant to CVA (e.g., Guha 1991, McCarthy 1993, Iwańska & Zadrozny 1997, Lenat 1998, Stalnaker 1999). Knowing more about the nature of context, having a more precise theory of CVA, and knowing how to teach it will allow us to more effectively help students identify context clues and know better how to use them, leading to larger vocabularies and better reading comprehension.

Learning new concepts and their words is not simply "more facts" or memorizing a definition. Concept learning requires making ever more refined discriminations of ideas, actions, feelings, and objects; it necessitates "assimilating" (Piaget 1952) the newly learned concept with prior knowledge, via inference, belief revision, and reorganization of existing cognitive schemata. We are investigating ways to facilitate readers' natural CVA by developing a rigorous computational theory of how context is used and creating a systematic, viable curriculum for teaching CVA strategies, based on our AI algorithms and on analysis of CVA processes used by good readers.

Almost everyone working on this topic believes that it is possible to "figure out" a meaning for a word "from context". Other terms in the literature include "construct", "deduce", "derive", "educe", "guess", "infer", or "predict"; I prefer to say that the reader "computes" a meaning for an unknown word; that is what our software does, and what our algorithm-based curriculum teaches.

But what is "context"? Most researchers in all disciplines have in mind *written*, as opposed to *spoken*, contexts and as opposed to a broader notion of "context" that might include visual or "situative" contexts (speaker, location, time, etc.). Still, there is ambiguity (see Engelbart & Theuerkauf 1999 for a survey): Informally, many researchers say something like this: "The reader can infer/guess/figure out, etc., the meaning of a word from context ..." (e.g., Werner & Kaplan 1952, McKeown 1985, Schatz & Baldwin 1986). Sometimes they *say* that, but *mean* something like this: "... from context *and* the reader's background knowledge" (e.g., Granger 1977, possibly Sternberg et al. 1983, Hastings & Lytinen 1994). Sometimes, instead of talking about "context *and* background knowledge", they talk about "context *including* background knowledge" (e.g., Nation & Coady 1988; see also Graesser & Bower 1990). But whereas 'context' as used in these studies has the connotation of being in the *external* text containing the word, 'background knowledge' has a connotation of being in the reader's *mind*. What exactly is, or should be, meant by the 'context' for contextual vocabulary acquisition?

Interdisciplinary cognitive scientists, especially, face the problem that many terms are used differently by different researchers, without any notice of the differences, often resulting in confusion. One should always try to figure out (from context, if by no other way!) how an author uses such terms. On the other hand, one should never *use* any of these terms without clearly explaining how one is using it. 'Context' is such a term. Here, I propose some definitions related to this term. I think the concepts are more important than the words we use for them, but we need to clearly define our words (so that they *don't* have to be figured out from context!).

A clue to the nature of context is in our CVA software: We use a knowledge-representation-and-reasoning (KRR) system (SNePS; Shapiro & Rapaport 1987, 1992, 1995) to represent, in a *single* semantic-network knowledge base (KB), *both* the information in the text *and* the reader's background knowledge. This strongly suggests that the relevant "context" for CVA of the unknown word is the entire surrounding *network* (or at least a subpart of it; for a defense of such holistic semantic theories, see Rapaport 2002).

What follows is a sequence of proposed terms and their definitions, leading to a proposal for the proper definition of 'context' as, arguably, it should be used in CVA and that is consistent with our computational approach.

Preliminary Definitions

An **unknown word** for a reader is by definition (is$_{def}$) a word (or phrase) that the reader has either never seen before, or is such that he or she has only the vaguest idea about its meaning. (For a discussion of levels of knowing the meaning of a word, see Kibby 1995.) For convenience, let's symbolize this by '*X*'.

A **text** is$_{def}$ a (written) passage. It could be as short as a sentence or as long as several paragraphs, and it will usually contain *X*. It is not essential that the text be written: Presumably the same techniques could be applied to oral CVA (though there would be attentional and memory limitations).[1] In any case, most CVA research concerns texts that are read, rather than heard.

The next definition uses a possibly awkward term of art, but it serves a useful role, and others have used it before (Brown & Yule 1983: 46–50, citing Halliday; Haastrup 1991): The **co-text** of *X* as it occurs in some text is$_{def}$ the entire text (be it one sentence or several paragraphs)[2] "minus" *X* (i.e., the entire text surrounding *X*). So, if *X* = 'brachet', and our text is:

(T1) 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. (Malory 1470: 66.)

then *X*'s co-text in (T1) is:

> There came a white hart running into the hall with a white ___ next to him, and thirty couples of black hounds came running after them.

The underscore marks the location of the missing *X*. Co-texts are often used in "cloze" tests, in which a passage with a missing word is presented to a subject, who must then "fill in the blank", e.g., determine what that word might have been. Note that in CVA, however, the reader

[1] See Gildea et al. 1990, Beals 1997, Aist 2000. We have experimented with MICASE (2002), though it is transcribed, hence written.

[2] *Pace* Schatz & Baldwin 1986, however, the co-text should not be *limited* to a 3-sentence window around *X*.

is not usually trying to *find* a known but missing *word*; rather, the reader is hypothesizing a *meaning* for a visible but unknown word.

The reader's **prior knowledge** is$_{def}$ the knowledge that the reader has when s/he *begins* to read the text and is able to recall as needed while reading. 'Knowledge' is the common, though probably not the best, term, since usually what is known is true. But obviously a reader might have lots of mistaken beliefs, so 'prior *beliefs*' is probably a better term. We can use these interchangeably as long as it's clear that prior "knowledge" can be false. '*Prior* knowledge' usefully suggests that it's what the reader has *before* reading, i.e., the beliefs that the reader brings to the text and has available for use in understanding it.

Similar terms are used by other researchers, each with slightly different connotations: (1) '*Background* knowledge' lacks the temporal connotation, but is otherwise synonymous for our purposes; it might, however, more usefully refer to the information that the text's *author* assumes that the reader should have. We could then distinguish the background knowledge *necessary* (or assumed) for understanding the text from the reader's actual prior knowledge. (2) '*World* knowledge' connotes general factual knowledge about things *other* than what the text is about. (3) Specialized, subject-specific knowledge about the text's topic is often called '*domain* knowledge'. (4) '*Commonsense* knowledge' connotes the sort of knowledge that "everyone" has (e.g., that water is wet, that dogs are animals, maybe that Columbus discovered America in 1492, etc., but no "domain" knowledge); I would include under this rubric both the sort of very basic commonsense information that the CYC KRR system is concerned with (Lenat 1995) and the somewhat more domain-specific information that the "cultural literacy" movement is concerned with (Hirsch 1987). There is much overlap among these different notions of knowledge. For instance, surely the reader's prior knowledge includes much commonsense knowledge, and the author's background knowledge might include much domain knowledge.

The Proper Definition of 'Context'

How, then, might we define the 'context' of X? 'Context' is the tricky word that should never be used without defining it. I begin with the following preliminary definition, with some caveats to be discussed in a moment (including what the plus-sign ('+') represents):

Definition 1 *The* **context** *of X for a reader is$_{def}$ the co-text of X + the reader's prior knowledge.*

I think it's quite clear that *both* co-text *and* prior knowledge are needed: To take a simple example, after reading:

> **(T2)** Then the hart went running about the Round Table; as he went by the sideboard, the white brachet bit him in the buttock .... (Malory 1470: 66.)

most subjects infer that brachets are (probably) animals. But they do not make the inference solely from this textual premise (T2); they must also use an enthymematic, prior-knowledge premise: if *x* bites *y*, then *x* is (probably) an animal.[3]

To refine Def. 1, recall that "text" (and hence "co-text") is something "out there" in the world, while "prior knowledge" is something "inside" our heads, in our minds. But I believe (as do many other cognitive scientists and many, if not most, reading specialists) that, when you read, you "internalize" the text you are reading, i.e., you "bring it into" your mind (cf. Gärdenfors 1997, 1999ab; Jackendoff 2002, §10.4; Rapaport in press).

Moreover, *this "internalized" text is more important than the actual words on paper*. As a simple example, suppose the text says "I'm going to put the cat out", but you misread it as "I'm going to put the car out". Your subsequent interpretation or understanding of the rest of the text will be quite different from that of someone who didn't misread 'cat' as 'car'. So, what matters for your understanding of the text is not what the text actually *is*, but what you *think* the text is.

We need a name for this "internalized text". Some people might call it a "represented text", since it's a representation of the text. I hesitate about that, since 'representation' is one of those words that I warned you about at the start of this. We might call it the reader's "mental model" of the text, but 'mental model'—though otherwise quite nice—is perilously close to being a brand name (Johnson-Laird 1983) and best avoided. For now, I can't think of a better name than ... 'internalized text'. I'll also use 'internalized co-text' for (you guessed it) the internalized co-text. So, perhaps a better definition for the "context" of *X* would be this:

Definition 2 *The* **context** *of X for a reader is$_{def}$ the reader's internalized co-text of X + the reader's prior knowledge.*

But there's another problem: When the internalized text is "added" to the prior knowledge, the result might not be a simple "conjunction" or "union" of the two things. An active reader will typically make some (possibly unconscious) inferences while reading. E.g., from the short text, "John sang. He was awful.", readers automatically infer that 'he' refers to John. (Some say that 'he' and 'John' co-refer; others, that 'he' refers back to 'John'—the differences don't matter for our purposes.) Or, e.g., a

[3] It's a bit more complex: We don't want to infer merely that this particular white brachet is an animal, but that brachets in general are animals.

reader of the phrase 'a white brachet' might infer (from prior, commonsense knowledge that only physical objects have color) that brachets are physical objects. Similarly, a reader might infer that, if person *A* is shorter than person *B*, who is shorter than person *C*, then *A* is shorter than *C*; or that if a knight picks up a brachet and carries it away, then the brachet (whatever 'brachet' might mean) must be small enough to be picked up and carried. In these cases, the whole is greater than the sum of the parts: The integration of the prior knowledge with the internalized text might include some extra beliefs that are not in the text and that were not previously in the prior knowledge, i.e., that were not previously known; i.e., you can learn from reading!

But the whole might also be *less* than the sum of the parts: From reading, you can also learn that one of your prior beliefs was *mistaken*. (It's less likely, though possible—e.g., in the case of a typographical error—that you'd conclude that a sentence *in the text* was in error; Rapaport & Shapiro 1995.) In that case, you'll be revising your beliefs by *eliminating* something.

So, that plus-sign in Defs. 1 and 2 should be taken with a grain of salt. There is a whole branch of AI, KR, and philosophy called "belief revision" that studies this. (See, e.g., Alchourrón et al. 1985, Martins & Shapiro 1988, Martins 1991, Gärdenfors 1992, Hansson 1999.) Here's a sample of some of their terminology applied to reading (but please also take some of this with a grain of salt, since the terminology isn't universally agreed on):

The plus-sign represents an operation that takes as input the reader's prior knowledge and internalized (co-)text, and that outputs an updated mental KB that is a "belief-revised integration" of the inputs. As the reader reads the text, some passages from it will be *added* (i.e., unioned or conjoined) to the reader's prior knowledge, and perhaps new inferences will be drawn; this is called 'expansion' of the prior KB. Other text passages will be added, followed by the *elimination* of prior-KB beliefs that are inconsistent with it (it is limited to prior beliefs, since a reader typically assumes that the text is correct, as just noted); this is called 'revision'. A few text passages (e.g., those involving typos) might be added, then rejected when seen to be inconsistent with the prior KB; this is called 'semi-revision'. Beliefs that are removed are said to be 'retracted'; such '*con*traction' of a KB might also result in the *re*traction of other beliefs that inferentially depended upon the removed one. After the text has been fully read, the reader might consider all (relevant) beliefs in his or her newly expanded mental KB, make new inferences, and eliminate further inconsistencies (such elimination is called 'consolidation'; cf. Hansson 1999 for definitions of these terms). Let's call the end result the '(belief-revised) integration' of the two inputs.

One more detail: '*X*' was the unknown word *in the text*. But we need a mental counterpart for it, an "internalized *X*", because everything else has been internalized. So, the final(?) definition of 'context' for CVA makes it a ternary relation among a reader, a word, and a text:

Definition 3 *Let T be a text. Let R be a reader of T. Let X be a word in T unknown to R. Let T–X be X's co-text in T. Then* **the context that** *R* **should use to hypothesize a meaning for** *R's* **internalization of** *X* **as it occurs in** *T* $=_{def}$ *the belief-revised integration of R's prior knowledge with R's internalization of T–X.*

I.e., the "context" that the reader should use in order to to compute a meaning during CVA is the single mental KB resulting from the belief-revised integration of the reader's prior knowledge with the reader's internalized (co-)text.

This view of what the full context is for CVA not only meshes nicely with most cognitive-science and reading-theoretic views of text understanding (e.g., Schank 1982, Rumelhart 1985), but also with most KRR techniques in AI for processing text, including our own: The reader's mind is modeled by a KB of "prior knowledge" (including commonsense knowledge, world knowledge, perhaps some domain knowledge, etc.) expressed in a semantic-network language (SNePS). As our computational cognitive agent ("Cassie"; cf. Shapiro & Rapaport 1987, 1995; Shapiro 1989) reads the text, she incorporates the information in the text into her KB, making inferences and performing belief revision along the way (using SNeBR; Martins & Shapiro 1988). Finally, when asked to define one of the words she read, she deductively searches this *single*, integrated KB for information that can fill appropriate slots of a definition frame (for details, see Rapaport & Ehrlich 2000; the notion of a definition frame was adapted from van Daalen-Kapteijns & Elshout-Mohr 1981, and the slots were inspired by Sternberg et al. 1983). Her definition is thus determined by relevant portions of the semantic-network KB (this is a version of a conceptual-role semantics that avoids alleged evils of holism; cf. Rapaport 2002). Thus, from our computational point of view, the "context" that she uses to hypothesize a meaning for a word represented in her KB is a single KB consisting of her prior knowledge as modified by, and including, that part of her KB containing the information that she incorporated into it from the text. This matches our definition of 'context' for CVA.

Distinguishing Co-text and Prior Knowledge

Although all relevant information is in this single KB, we may need to distinguish between beliefs that came from the (co-)text, beliefs that were already in the reader's prior knowledge, and beliefs that arose from inferences from

both of these.

First, when eliminating one of two inconsistent beliefs, we need to know their sources, so that we would know whether to retract a prior belief or a belief originating from the text. We do this by marking each proposition with a "knowledge category": 'story', meaning that the proposition came from the text; 'life', meaning that it came from prior knowledge, etc. (for details and examples, see Ehrlich 1995, Rapaport & Ehrlich 2000).

Second, consider the following text containing the (presumably) unknown word 'detritus':

(T3) The birds alert nearby anglers that a massive school of menhaden is under attack by bluefish. The razor-toothed blues tear at the menhaden like piranhas in a killing frenzy, gorging themselves, some killing even when they are too full to eat, some vomiting so they can eat again. Beneath the blues, weak fish begin to circle, snaring the **detritus** of the carnage. (Franklin 2001.)

What prior knowledge might be useful to compute a meaning for 'detritus' from this passage? One possibility is the following defeasible rule:

(R) If fish *x* attacks fish *y*, and fish *z* is weaker than fish *x*, then fish *z* will only get leftovers.

From this rule and the following part of T3:

(T3.1) [W]eak fish begin to circle, snaring the detritus of the carnage.

we can infer that 'detritus' might be 'leftovers':

Let R′—representing the version of R in the KR language—and the items labeled 'PK*i*' be found in the reader's prior knowledge; let 'WF1' be a Skolem constant referring to some item in the reader's prior knowledge that satisfies the conditions in PK3; and let T3.2 be part of T3):[4]

R′. $(\forall x,y,z)[(\text{Fish}(x) \wedge \text{Fish}(y) \wedge \text{Fish}(z) \wedge \text{Attacks}(x,y) \wedge \text{Weaker-than}(z,x)) \rightarrow \exists w[\text{Leftovers}(w) \wedge \text{Gets}(z,w) \wedge \forall v[\text{Gets}(z,v) \rightarrow v = w]]]$

PK1. Fish(bluefish)

PK2. Fish(menhaden)

PK3. Fish(WF1) ∧ Weaker-than(WF1, bluefish)

PK4. $(\forall x,y)[\text{Tears-at}(x,y) \rightarrow \text{Attacks}(x,y)]$

PK5. $(\forall x,y)[\text{Snares}(x,y) \rightarrow \text{Gets}(x,y)]$

T3.2. Begin-to-Circle(WF1) ∧ Snares(WF1, detritus) ∧ Tears-at(bluefish, menhaden)

Using the substitution $\{x := \text{bluefish}, y := \text{menhaden}, z := \text{WF1}\}$, we can instantiate and apply *modus ponens* (MP) to R′, PK1–PK4, and T3.2 to get:

$\exists w[\text{Leftovers}(w) \wedge \text{Gets}(\text{WF1}, w) \wedge \forall v[\text{Gets}(\text{WF1}, v) \rightarrow v = w]]$

Applying MP to PK5 and T3.2 allows us to infer: Gets(WF1, detritus). Now, *if it were the case* that Leftovers(detritus) $\wedge$ $\forall v[\text{Gets}(\text{WF1}, v) \rightarrow v = \text{detritus}]$ (i.e., there is nothing else for the weak fish to get except the detritus of the carnage), then we would be able to *deductively* infer the consequent of R′. So, we can *abductively* infer Leftovers(detritus) (cf. Hobbs et al. 1993). This gives us a partial definition (or meaning hypothesis) for 'detritus'. This hypothesis, of course, is defeasible (i.e., it might be incorrect), yet plausible, and can serve as a first approximation to a full definition. At the very least—but perhaps most importantly—*it enables the reader to understand this passage*.

However, we *don't* want to infer from T3.1, *which is from the text*, and (e.g.) "They (those weak fish) also snared worms.", which let's suppose is *also* in the text, that 'detritus' are worms. One way to block this is to only allow the previous inference to go through when we use prior knowledge together with internalized text information, rather than two pieces of information from the text. And one way to do *that* is to associate each proposition with its source: text or prior knowledge (or an inference from these).

To sum up: When we speak of figuring out a meaning for a word "from context", we should mean: from the belief-revised integration of the reader's prior knowledge with the reader's internalized co-text of the word, with each proposition in this single mental KB marked with its source.[5]

References

Aist, G. (2000), "Helping Children Learn Vocabulary during Computer Assisted Oral Reading", *AAAI-2000*: 1100–1101.

Alchourrón, C., et al. (1985), "On the Logic of Theory Change", *J. Symb. Logic* 50: 510-530.

Ames, W. (1966), "The Development of a Classification Scheme of Contextual Aids", *Reading Res. Qtly.* 2: 57–82.

Beals, D. (1997), "Sources of Support for Learning Words in Conversation", *J. Child Lang.* 24: 673–694.

Blachowicz, C., & Fisher, P. (1996), *Teaching Vocabulary in All Classrooms* (Prentice-Hall/Merrill).

Brown, G., & Yule, G. (1983), *Discourse Analysis* (Cambridge).

Clarke, D., & Nation, P. (1980), "Guessing the Meanings of Words from Context", *System* 8: 211–220.

Ehrlich, K. (1995), "Automatic Vocabulary Expansion through Narrative Context," *Tech. Rep. 95-09* (SUNY Buffalo Comp. Sci. Dept.).

[4] Here, we ignore difficulties in representing generics like 'bluefish', 'menhaden', etc.

[5] The preparation of this paper was supported by grant REC-0106338 from the National Science Foundation. I am grateful to Marc Broklawski, Frances L. Johnson, Michael W. Kibby, Brian S. Morgan, Scott Napieralski, Stuart C. Shapiro, Karen Wieland, other members of the CVA and SNePS Research Groups, and Jan Chomicki for discussion and for comments on earlier drafts.

Ehrlich, K., & Rapaport, W. (1997), "A Computational Theory of Vocabulary Expansion", *Proc. 19th Conf. Cog. Sci. Soc.* (Erlbaum): 205–210.

Engelbart, S., & Theuerkauf, B. (1999), "Defining Context within Vocabulary Acquisition", *Lang. Teaching Res.* 3: 57–69.

Franklin, H. (Sep. 2001), "The Most Important Fish in the Sea", *Discover*: 44–50.

Gärdenfors, P. (ed.) (1992), *Belief Revision* (Cambridge).

Gärdenfors, P. (1997), "Meanings as Conceptual Structures", in M. Carrier & P. Machamer (eds.), *Mindscapes* (Univ. Pittsburgh Press).

Gärdenfors, P. (1999a) "Does Semantics Need Reality?", in A. Riegler et al. (eds.), *Understanding Representation in the Cognitive Sciences* (Kluwer/Plenum).

Gärdenfors, P. (1999b), "Some Tenets of Cognitive Semantics", in J. Allwood & P. Gärdenfors (eds.), *Cognitive Semantics* (Benjamins).

Gildea, P., et al. (1990), "Contextual Enrichment by Videodisc", in D. Nix & R. Spiro (eds.), *Cognition, Education, and Multimedia* (Erlbaum).

Graesser, A., & Bower, G. (eds.) (1990), *Inferences and Text Comprehension* (Academic).

Granger, R. (1977), "Foul-Up", *IJCAI-77*: 172–178.

Guha, R. (1991), "Contexts" (Stanford Univ. Ph.D. Thesis).

Haastrup, K. (1991), *Lexical Inferencing Procedures or Talking about Words* (Gunter Narr).

Hansson, S. (1999), *A Textbook of Belief Dynamics* (Kluwer).

Hastings, P., & Lytinen, S. (1994), "The Ups and Downs of Lexical Acquisition", *AAAI-94*: 754–759.

Hirsch, E. (1987), *Cultural Literacy* (Houghton Mifflin).

Hobbs, J., et al. (1993), "Interpretation as Abduction", *Artif. Intell.* 63: 69–142.

Iwańska, Ł., & Zadrozny, W., (eds.) (1997), Context in Natural Language Processing, *Comp. Intell.* 13(3).

Jackendoff, R. (2002), *Foundations of Language* (Oxford).

Johnson-Laird, P. (1983), *Mental Models* (Cambridge).

Kibby, M. (1995), "The Organization and Teaching of Things and the Words that Signify Them", *J. Adolescent & Adult Literacy* 39: 208–223.

Lenat, D. (1995), "CYC", *CACM* 38(11): 33–38.

Lenat, D. (1998), "The Dimensions of Context-Space" [http://www.cyc.com/].

Malory, T. (1470), *Le Morte Darthur*, ed. R. Lumiansky (Collier, 1982).

Martins, J. (1991), "The Truth, the Whole Truth, and Nothing but the Truth", *AI Mag.* 11(5): 7–25.

Martins, J. & Shapiro, S. (1988), "A Model for Belief Revision," *Artif. Intell.* 35: 25–79.

McCarthy, J. (1993), "Notes on Formalizing Context", *IJCAI*.

McKeown, M. (1985), "The Acquisition of Word Meaning from Context by Children of High and Low Ability", *Reading Res. Qtly.* 20: 482–496.

MICASE: Michigan Corpus of Academic Spoken English (2002), [http://www.hti.umich.edu/m/micase/].

Nagy, W., & Anderson, R. (1984), "How Many Words Are There in Printed School English?", *Reading Res. Qtly.* 19: 304–330.

Nation, P., & Coady, J. (1988), "Vocabulary and Reading", in R. Carter & M. McCarthy (eds.), *Vocabulary and Language Teaching* (Longman).

Piaget, J. (1952), *The Origins of Intelligence in Children*, M. Cook, trans. (Int'l. Univ. Press).

Rapaport, W. (2002), "Holism, Conceptual-Role Semantics, and Syntactic Semantics", *Minds & Machines* 12: 3–59.

Rapaport, W.J. (2003), "Bibliography of Theories of Contextual Vocabulary Acquisition", online at: [http://www.cse.buffalo.edu/∼rapaport/refs.vocab.html].

Rapaport, W. (in press), "What Did You Mean by That?", *Minds & Machines*.

Rapaport, W., & Ehrlich, K. (2000), "A Computational Theory of Vocabulary Acquisition", in Ł. Iwańska & S. Shapiro (eds.), *Natural Language Processing and Knowledge Representation* (AAAI/MIT).

Rapaport, W., & Kibby, M. (2002), "Contextual Vocabulary Acquisition", *Proc. 6th World Multiconf. Systemics, Cybernetics & Informatics*: 261–266.

Rapaport, W., & Shapiro, S. (1995), "Cognition and Fiction", in J. Duchan et al. (eds.), *Deixis in Narrative* (Erlbaum).

Rumelhart, D. (1985), "Toward an Interactive Model of Reading", in H. Singer & R. Ruddell (eds.), *Theoretical Models and Processes of Reading; 3/e* (Int'l. Reading Assoc.).

Schank, R. (1982), *Reading and Understanding* (Erlbaum).

Schatz, E., & Baldwin, R. (1986) "Context Clues Are Unreliable Predictors of Word Meanings", *Reading Res. Qtly* 21: 439–453.

Shapiro, S. (1989), "The CASSIE Projects", *EPIA 89* (Springer *Lect. Notes in AI 390*).

Shapiro, S., & Rapaport, W. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in N. Cercone & G. McCalla (eds.), *The Knowledge Frontier* (Springer-Verlag).

Shapiro, S., & Rapaport, W. (1992), "The SNePS Family," *Comp. & Math. with Applications* 23: 243–275.

Shapiro, S., & Rapaport, W. (1995), "An Introduction to a Computational Reader of Narrative," in J. Duchan et al. (eds.), *Deixis in Narrative* (Erlbaum).

Stalnaker, R. (1999), *Context and Content* (Oxford Univ. Press).

Sternberg, R., et al. (1983), "Teaching Vocabulary-Building Skills", in A. Wilkinson (ed.), *Classroom Computers and Cognitive Science* (Academic).

van Daalen-Kapteijns, M., & Elshout-Mohr, M. (1981), "The Acquisition of Word Meanings as a Cognitive Learning Process", *J. Verbal Learning & Verbal Behavior* 20: 386–399.

Werner, H., & Kaplan, E. (1952), "The Acquisition of Word Meanings", *Monographs, Soc. Res. Child Dev.*.

Wesche, M., & Paribakht, T. (eds.) (1999), "Incidental L2 Vocabulary Acquisition", *Studies in Second Lang. Acquisition* 21(2).