

# In Defense of Contextual Vocabulary Acquisition: How to do Things with Words in Context

William J. Rapaport

Department of Computer Science and Engineering, Department of Philosophy, and Center for Cognitive Science,  
State University of New York at Buffalo, Buffalo, NY 14260-2000, USA.  
+1-716-645-3180x112; fax +1-716-645-3464  
rapaport@cse.buffalo.edu, <http://www.cse.buffalo.edu/~rapaport/CVA/>

**Abstract.** “Context” is notoriously vague, and its uses multifarious. Researchers in “contextual vocabulary acquisition” differ over the kinds of context involved in vocabulary learning, and the methods and benefits thereof. This paper presents a computational theory of contextual vocabulary acquisition, identifies the relevant notion of context, exhibits the assumptions behind some classic objections, and defends our theory against these objections.

## 1 A Computational Theory of Contextual Vocabulary Acquisition

*Contextual vocabulary acquisition* (CVA) is the active, deliberate acquisition of a meaning for a word in a text by reasoning from context, where “context” *includes*: (1) the reader’s “internalization” of the surrounding text, i.e., the reader’s “mental model” of the word’s “*textual* context” (or “co-text”: Brown & Yule 1983: 46–50, citing Halliday; Haastrup 1991) integrated with (2) the reader’s prior knowledge (including (a) the reader’s knowledge of language and (b) meaning hypotheses developed by the reader from prior encounters with the word), but it *excludes* (3) external sources of help such as dictionaries or people.

CVA is a process that can be used by a reader to figure out a meaning for an unfamiliar word as it occurs in a passage being read. It is what you do when you come across such a word in your reading, realize that you don’t know what it means, and decide that you need to know what it means in order to understand the passage, but there is no one around to ask, and it is not in the dictionary (or you are too lazy to look it up). In such a case, you can try to figure out its meaning “from context”, i.e., from clues in the co-text together with your prior knowledge (PK). (Although CVA can presumably be used in non-reading (non-textual) situations—e.g., in ordinary conversation, when watching TV, etc.—here I focus only on CVA during reading. Everything I say here carries over to spoken domains.)<sup>1</sup>

My colleagues and I have developed a computational theory of CVA, which is implemented in a propositional semantic-network knowledge-representation-and-reasoning system (SNePS; Shapiro & Rapaport 1987, 1992, 1995; Shapiro & Group 2004). Our computational system begins with a stored knowledge base containing SNePS representations of relevant PK. It takes as input SNePS representations of a passage containing an unfamiliar word. The processing begins with inferences drawn from these two sources of information. When asked to define the word, it applies definition algorithms (for nouns and for verbs; adjectives are under investigation) that deductively search the resulting network for information of the sort that might be found in a dictionary definition, outputting a definition frame (or schema) whose slots are the kinds of features that a definition might contain (e.g., class membership, properties, actions, spatio-temporal information, etc.) and whose slot-fillers contain information gleaned from the network. (Details of the underlying theory, representations, processing, inferences, belief revision, and definition algorithms are presented in Ehrlich 1995, 2004; Ehrlich & Rapaport 1997, 2004; Rapaport & Ehrlich 2000; Rapaport & Kibby 2002; Rapaport 2003a.) We are investigating ways to make our system more robust, to embed it in a natural-language-processing system, and to incorporate what Sternberg and colleagues (1983ab, 1987) call “internal” context, i.e., morphological (and etymological) analysis.

Since the CVA process of figuring out a meaning for a word from context is computable (as evidenced by the existence of our algorithms), I prefer the phrase “*compute* a meaning” rather than “*figure out* a meaning”. But I prefer ‘figure out’ to any of the other terms that appear in the literature, such as ‘deduce’ (which is too narrow) or ‘infer’: Simon (1996: 171) observes that it is “more accurate to say that” a text “suggests” meanings than that a reader “infers”

<sup>1</sup> There is evidence that “conversation is not a substitute for reading” in terms of the benefits of reading for improving not only vocabulary but general intelligence, as well as evidence that watching TV has *negative* effects! (Cunningham & Stanovich 1998.)

meanings from it. I especially dislike “guess a meaning”, with its connotation of randomness, as well as for its lack of guidance for readers doing CVA (see §2.2, below). However, since many of my non-computer-science colleagues find my use of the phrase “the reader computes a meaning” to be awkward at best, I shall refrain from using it too often, reverting to “figures out”. (But what I mean in such contexts by ‘figures out’ is “computes”!)

Our research group, which includes reading educators, is exploring ways to apply our methods to the development of what we hope will be a better pedagogical curriculum than the current state of the art for teaching CVA to readers. To this end, we have been studying the CVA literature from a variety of disciplines that, generally speaking, seem to ignore each other’s literature (including computational linguistics, reading education, second-language education, and psychology; see Rapaport 2004 for a bibliography).

Two often-cited papers by reading scientists (Beck et al. 1983; Schatz & Baldwin 1986) have claimed that not only are certain contexts less than useful for doing CVA, but that most “natural” contexts (as opposed to artificial, “pedagogical” contexts) are not helpful at all. However, a careful examination of their arguments points up several assumptions behind these claims that are inconsistent with our computational theory of CVA. Thus, their objections do not apply to our theory.

## 2 Are All Contexts Created Equal?

In a paper subtitled “All Contexts Are Not Created Equal”, Beck et al. (1983) present four prototypical kinds of contexts along a “continuum” (178;<sup>2</sup> see §2.5, below). Their main claim is that “it is not true that every context is an appropriate or effective instructional means for vocabulary development” (177). In this section, we examine their paper carefully to discern their assumptions and to see where our approach agrees and disagrees with theirs.

### 2.1 The Role of Prior Knowledge

Beck et al. begin by pointing out that “the context that surrounds a word in text can give *clues* to *the word’s meaning*” (177, my emphases). But a passage is not a clue without some other information to interpret it as a clue (‘clue’ is a relative term).<sup>3</sup> Therefore:

#### Implication A1

The clues in the text need to be supplemented with other information in order for a meaning to be figured out.

*This supplemental information must be supplied by the reader’s PK.* Such PK (better: such prior *beliefs*, since what the reader believes need not be true) might include general “world” or cultural knowledge (Hirsch 1987, 2003), “common-sense” knowledge (Lenat 1995), specialized “domain” knowledge, and it might or might not include the “background” knowledge that the author assumes that the reader will have: The author “is counting on [the] . . . words [in the text] evoking somewhat the same associations in the reader’s mind as are present in his [*sic*] own mind. Either he has some sense of what his readers know or he assumes that their knowledge stores resemble his” (Simon 1996: 171). However, not all of the reader’s PK may be consciously available at the time of reading. More importantly, each reader will bring to bear upon his or her interpretation of the text *idiosyncratic* PK. (See, e.g., Dulin 1969, Garnham & Oakhill 1990; Hobbs 1990 argues that a text’s meaning is a function of both the text *and* the reader’s mind).<sup>4</sup>

The “context” that the reader uses to figure out a word’s meaning is not just the *textual* context, but a *wider* context consisting of the reader’s available PK “integrated” with the reader’s “internalization” (or “mental model”) of the co-text (though not necessarily in Johnson-Laird’s (1983) sense). Its integration with the reader’s PK will involve belief revision: new beliefs added by inference and old beliefs withdrawn in the face of new evidence. The new beliefs would be inferred as conclusions from arguments in which at least one premise comes from the internalized co-text and at least one premise comes from PK. Typically, withdrawn beliefs are PK beliefs that are inconsistent with co-text propositions. (For further discussion of co-text, PK, and belief revision, see Rapaport 2003a.) In what follows, except

<sup>2</sup> All page references in this section will be to Beck et al. 1983, unless otherwise noted.

<sup>3</sup> I owe this observation to my colleague J.-P. Koenig.

<sup>4</sup> This phenomenon seems akin to a similar one in the area of (visual) perception: Whether an object is seen as white is a function not only of the object in its environment but also of the perceiver’s state of adaptation. If the perceiver has been in a red environment and has, accordingly, visually adapted to red, then an object that looks white to that perceiver will look red to someone entering the red environment from a white environment (Webster 2004).

when quoting, I will use ‘co-text’ or ‘textual context’ to refer to the text surrounding an unfamiliar word, reserving ‘context’ or ‘wide context’ to refer to the reader’s “internalized” co-text “integrated” with the reader’s PK.

The reader’s internalization of the text involves some interpretation (e.g., resolving pronoun anaphora) or the immediate and unconscious drawing of an inference (e.g., that the pronoun ‘he’ refers back to a male or that ‘John’ is a proper name typically referring to a male human) (cf. Garnham & Oakhill 1990: 383). Consider the following natural passage (currently being analyzed by our research group):

The archives of the medical department of Lourdes are filled with *dossiers* that detail well-authenticated cases of what are termed miraculous healings. (Murphy 2000: 45; my italics.)

Is this to be understood as saying (a) that the archives are filled with dossiers, and that *these dossiers* detail cases of miraculous healings? Or is it to be understood as saying (b) that the archives are filled with dossiers, and *dossiers in general* are things that detail cases of miraculous healings? The difference in interpretation has to do with whether “detail . . . miraculous healings” is a restrictive relative clause (case (a)) or a non-restrictive relative clause (case (b)). Now, arguably, it should be understood as in (a); otherwise, the author should have written, “The archives are filled with dossiers, *which* detail miraculous healings”. But a reader might not be sensitive to this distinction (preferably indicated by ‘that’ *without* preceding comma vs. ‘which’ *with* preceding comma). The notion of misinterpretation cuts both ways: The author might not be sensitive to it, either, and might have written it one way though intending the other. It makes a difference for CVA. A reader who is unfamiliar with ‘dossier’ might conclude from the restrictive interpretation that a dossier is something found in an archive and that these particular dossiers detail miraculous healings, whereas a reader who internalized the non-restrictive interpretation might conclude that a dossier is something found in an archive that (necessarily) details miraculous healings. (We have, in fact, anecdotal evidence that at least some readers of this passage do interpret it in the latter way.)

Or consider how even a common word can mean different things to different people: In some dialects of Indian English, something that looks like a sofa, even if it sits only one person, is a ‘sofa’, while it is a ‘chair’ or ‘recliner’ in American English. Thus, two fluent English speakers might interpret a passage containing the word ‘sofa’ rather differently: The text is the same in both cases, but the readers’ *internalized* texts will be quite different.<sup>5</sup>

It can also involve *mis*interpretation (cf. Garnham & Oakhill 1990: 387ff), even simple misreading: I read the sign on a truck parked outside one of our university cafeterias, where food-delivery trucks usually park, as “Mills Wedding and Specialty Cakes”. Why had I never heard of this local bakery? Why might they be delivering a cake? So I re-read the truck’s sign more carefully. It actually said, “Mills Welding and Specialty Gases”!

Another modifying influence, perhaps related to misreading, stems from reading difficulties that might circumscribe the amount of co-text that the reader can understand and therefore integrate into his or her mental model. Stanovich 1986 (2000: 467) notes that we must “distinguish the nominal context (what is on the page) from the effective context (what is being used by the reader)”.

## 2.2 Do Words Have Unique Meanings?

The assumption—inconsistent with our theory—underlying Beck et al.’s use of the phrase ‘the word’s meaning’ is:

**Assumption A2** A word has a *unique* meaning.

The definite description ‘*the* word’s meaning’ or ‘*the* meaning of a word’ is ubiquitous in the literature on reading, vocabulary learning, and semantics. It is hard to avoid its use, but worth the effort, for it incorrectly suggests that a word has a unique meaning. To be charitable, we could say that what’s normally intended by this phrase is “the meaning of a word *in the present context*”: As Deighton (1959) observed, “context always *determines* the meaning of a word, it does not always *reveal* it” (cited in Schatz & Baldwin 1986: 451). But it follows from our observations about implication A1 (that textual clues need to be supplemented with other information) that the reader will supplement the co-text with idiosyncratic PK, and, consequently, each reader will interpret the word slightly differently. Of course, on this reading, Deighton is still essentially correct: Wide context—understood as the reader’s internalized co-text integrated via belief revision with that reader’s PK—determines *a* meaning for the word, though it requires further processing to reveal that meaning.

<sup>5</sup> Shakthi Poornima, personal communication.

The need for further processing underlies Johnson-Laird's (1987) observation that we don't store definitions, even of words we understand. It also undercuts pedagogical strategies for CVA that instruct the reader merely to "guess" the meaning (e.g., Clarke & Nation 1980, Nation & Coady 1988). Nation (2001: 257) even boasts that his guessing strategy "does not draw on background content knowledge" since "linguistic clues will be present in every context, background clues will not". But background knowledge (or PK) is essential and is always used; it is, in fact, unavoidable, even in Nation's own strategy: Where he says "Guess" (the entire step 4 in Nation's 5-part strategy!), he must in fact mean "make an educated guess", i.e., an inference, but that inference must rely on more premises than merely what is explicit in the text; such premises come from PK. (For further discussion of the essential role of PK, see, e.g., Anderson 1984. On so-called "missing" premises, see Singer et al. 1990, Suh & Trabasso 1993, Rapaport 2003a.)

### 2.3 Do Words Have Correct Meanings?

There is a closely related assumption that many authors make:

**Assumption A3** There is such a thing as "the *correct* meaning of a word" (in a given context).

For instance, Beck et al. comment that "even the appearance of each target word in a strong, directive context [i.e., a context conducive to figuring out "a correct meaning"] is far from sufficient to develop *full knowledge* of word meaning" (180, my emphasis).

Perhaps what is meant by the "correct" meaning is that there is a certain meaning that the author intended. But if we are concerned with a word's meaning *as determined by* the reader's internalized co-text integrated with the reader's PK, then it might very well be the case that the *author's* intended meaning is *not* thus determined. Our investigations suggest that this is almost always the case. The best that can be hoped for is that a reader will be able to hypothesize or construct *a* meaning for the word, rather than *the* meaning of the word (i.e., the reader *gives* or *assigns* a meaning to the word). "The meaning of things lies not in themselves but in our attitudes toward them" (St. Exupéry).<sup>6</sup>

If the meaning that the reader figures out *is* the intended one, so much the better. If not, has the reader then *misunderstood* the text? Misunderstanding is not necessarily a bad thing: If no one ever misunderstood texts—or, at any rate, if no one ever understood texts differently from other readers or from the author's intended meaning—then there would be little need for reading instruction, literary criticism, legal scholarship, etc. (Cf. Rapaport 2002, 2003b, where it is argued, following Russell (1918), that—because of individual differences in our idiosyncratic conceptual meanings—we always *misunderstand* each other. Russell, by contrast, celebrates this as being precisely the mechanism that makes conversation and the exchange of information possible.) The important question is not whether a reader can figure out the *correct* meaning of a word, but whether the reader can figure out *a* meaning for the word *that is sufficient to enable him or her to understand the text*.

The reader need not understand the text "perfectly", but merely *well enough to continue reading*. As Johnson-Laird (1987) has pointed out, we don't normally have, *nor do we need*, "full knowledge"—full, correct definitions—of the words that we understand: We can understand—well enough for most purposes—the sentence "During the Renaissance, Bernini cast a bronze of a mastiff eating truffles"<sup>7</sup> without being able to define any of the terms in that sentence, as long as we have even a vague idea that, e.g., the Renaissance was some period in history, 'Bernini' is someone's name, "casting a bronze" has something to do with sculpture, bronze is some kind of (perhaps yellowish) metal, a mastiff is some kind of animal (maybe a dog), and truffles are something edible (maybe a kind of mushroom, maybe a kind of chocolate candy).

Consider the following passage that our research group is currently analyzing:<sup>8</sup>

All chances for agreement were now gone, and compromise would now be impossible; in short, an *impasse* had been reached. (Dulin 1970.)

Here is one way a reader might figure out a meaning for 'impasse' from this text: A compromise is an agreement. If all chances for agreement are gone, then agreement is impossible. So both conjuncts of the first clause say more or less the same thing. 'In short' is a clue that what follows means more or less the same as what precedes it. So, to say

<sup>6</sup> Cited by Michael Sims, *Chronicle of Higher Education* (15 August 2003): B4.

<sup>7</sup> Johnson-Laird, personal communication.

<sup>8</sup> This passage comes from an article detailing possible contextual clues for CVA that can be taught in a classroom; consequently, it might be a "pedagogical", not a "natural", passage (as defined in §2.4, below).

that an impasse has been reached is to say that agreement is impossible. And (perhaps with a bit more plausible PK) that means that an impasse is a *disagreement*.<sup>9</sup>

Is an impasse a disagreement? At least one dictionary defines it simply as a “deadlock” (Waite 1998). Suppose that “deadlock” is “the correct meaning” of ‘impasse’. If the reader decides that ‘impasse’ means “disagreement”, not “deadlock”, has the reader misunderstood the passage? Consider the following scenarios:

1. The reader never sees the word ‘impasse’ again. It then hardly matters whether she has not “correctly” understood the word (though, in the case of this particular bit of CVA, she has surely figured out *a* very plausible meaning).
2. The reader sees the word again in a context in which “disagreement” is a plausible meaning. Since her PK now includes a belief that ‘impasse’ means “disagreement”, this surely helps in understanding the new passage.
3. The reader sees the word again in a context in which “deadlock”, not “disagreement”, is the “best” meaning. E.g., she might read a computer science text discussing operating-system deadlocks, in which a particular deadlock is referred to as an “impasse”. Here, it *might* make little sense to consider the situation as a “disagreement”, so:
  - (a) The reader might decide that this occurrence of ‘impasse’ could not possibly mean “disagreement”. Again, there are two possibilities:
    - i. She decides that she must have been wrong about ‘impasse’ meaning “disagreement”, and she now comes to believe (say) that it means “deadlock”.
    - ii. She decides that ‘impasse’ is polysemous, and that “deadlock” is a second meaning. (Cf. the discussion of the polysemy of the verb ‘to dress’, in Rapaport & Ehrlich 2000, in which it is strongly believed that to dress is to put clothes on, but from textual contexts such as “King Claudas dressed his spear before battle”, it is inferred that to dress is *also* to prepare for battle.)
  - (b) Or the reader might try to reconcile the two possible meanings, perhaps by viewing deadlocks as disagreements, if only metaphorically.<sup>10</sup>

## 2.4 Two Kinds of Textual Context

Let us now turn to Beck et al.’s categorization of textual contexts. They are interested in using *textual* context to help *teach* “the meaning” of the word. We, however, are interested in using *wide* context to help figure out a meaning for an unfamiliar word, for the purpose of *understanding the passage* containing it. These two interests don’t always coincide, especially if the former includes as one of its goals the student’s ability to *use* the word. From the fact that a given co-text might not clearly convey a word’s “correct” meaning, it does not follow that a useful meaning cannot be figured out from it (especially since the wider context from which a meaning is figured out includes the reader’s PK and is not therefore restricted to the co-text). Some co-texts certainly provide more clues than others. The question, however, is whether all CVA is to be spurned because of the less-helpful co-texts.

The top level of Beck et al.’s classification divides all co-texts into two kinds: *pedagogical* and *natural*. The former are “specifically designed for teaching designated unknown words” (178). It will be of interest later that the only example they give of a pedagogical co-text is for a *verb* (italicized below):

All the students made very good grades on the tests, so their teacher *commended* them for doing so well. (178)

By contrast, “the author of a natural context does not intend to convey *the meaning of a word*” (178, my emphasis). Note the assumptions about unique, correct meanings at work. In contrast, and following Deighton 1959 (see §2.2, above), I would say that the author of a natural co-text *does*—no doubt, unintentionally—convey *a* meaning for the word in question. And that meaning is the only one that a reader might be expected to figure out. Beck et al. go on to observe that natural “contexts will not necessarily provide *appropriate* cues to the meaning of a particular word” (178, my emphasis). This does not mean that no cues (or clues) are provided.<sup>11</sup> It may well be that clues *are* provided for *a* meaning that helps the reader understand the passage.

<sup>9</sup> Plausibly, if agreement is *impossible*, then *disagreement* is possible. And, plausibly, if reaching a goal (albeit a negative goal, viz., an impasse—whatever that is) is *also* possible (perhaps because it has happened, and whatever happens is possible), then perhaps an impasse is also a *disagreement*. These are defeasible inferences, but they are the sort of inferences our protocols show that readers actually make. Cf. Hobbs et al. 1993 for discussion of the abductive nature of some of these inferences.

<sup>10</sup> On how CVA can be used to understand metaphor, see Badiu & Anderson 2001. On how metaphor can extend meanings, see Lakoff 1987, Rapaport 2000 and forthcoming.

<sup>11</sup> Many authors write of “cues”; others, of “clues”; some (e.g., Beck et al.), of both. ‘Cue’ suggests a textual element that prompts the reader, perhaps unconsciously, to think of something. ‘Clue’ suggests a textual element that a careful reader can use to (perhaps consciously) infer something. The two terms seem to be interchangeable in the literature, but I will try to use them in the way mentioned here, except when quoting.

Finally, note that the pedagogical-natural distinction may ultimately be hard to maintain: A passage produced for pedagogical purposes by one researcher might be taken as “natural” by another (see §2.6, below).

## 2.5 Four Kinds of (Natural) Co-texts

**Misdirective Co-texts.** Natural co-texts are divided into four categories. “At one end of our continuum are misdirective contexts, those that seem to direct the student to an *incorrect* meaning for a target word” (178, my emphasis). I am willing to believe that some co-texts are misdirective. But Beck et al.’s sole example does not inspire confidence:

Sandra had won the dance contest and the audience’s cheers brought her to the stage for an encore. “Every step she takes is so perfect and graceful,” Ginny said *grudgingly*, as she watched Sandra dance. (178.)

Granted, a reader might incorrectly decide from this that ‘grudgingly’ meant something like “admiringly”. But there are three problems with this example: (1) The authors provide no evidence that this is, indeed, a natural co-text. But this is a minor matter; surely, many such allegedly misdirective co-texts could be found in nature, so to speak.

(2) If it is a natural co-text, it would be nice to see a bit more of it. Indeed, another assumption many CVA researchers make is this:

**Assumption A4** (Textual) contexts have a fixed, usually small size.

But, in the present example, there might be other clues, preceding or following this short co-text, that would rule out “admiringly”. Perhaps we know or could infer from earlier or later passages that Ginny is jealous of Sandra, or that she is inclined to ironic comments. Strictly speaking, one could logically infer from this passage a disjunction of possible meanings of ‘grudgingly’ and later rule some of them out as more occurrences of the word are found.

(3) But, most significantly, ‘grudgingly’ is an adverb. Now, another assumption is this:

**Assumption A5** All words are equally easy (or equally difficult) to learn.

But adverbs, adjectives, and other such modifiers are notoriously hard cases for CVA and for first-language learning. (See, e.g., Granger 1977; Gentner 1981, 1982; Gillette et al. 1999. Not nearly as much research has been done on adjective or adverb learning as on noun and verb learning.)

Thus, the evidence provided for the existence of misdirective co-texts is weak, primarily since there should be *no* limit on the size of a co-text (see §3.2, below) and since the only example concerns an adverb, which can be difficult to interpret in any context. There is no “limit” on the size of the *wide* context. Certainly a reader’s PK (which is part of that wide context) might include lots of beliefs that might assist in coming up with a plausible meaning for ‘grudgingly’ in this passage.<sup>12</sup>

Another assumption is also at work. Beck et al. conclude, “Thus, incorrect conclusions about word meaning are likely to be drawn” from misdirective co-texts (178). This assumes

**Assumption A6** Only one co-text can be used to figure out a meaning for a word.

Granted, if a word only occurs once, and in the most grievous of misdirective co-texts, then it is quite likely that a reader would “draw an incorrect conclusion”, if, indeed, the reader drew *any* conclusion. However, in such a case, it does not matter what the reader concludes or whether the reader concludes anything at all, for it is highly unlikely that anything crucial will turn on such a word. More likely, the reader will encounter the word again, and will have a chance to revise his or her initial hypothesis about what it might mean.

In general, the task of CVA is one of hypothesis generation and testing; it is fundamentally a scientific task of developing a theory about a word’s meaning (or possible meanings). It is not mere guessing (although Loui (2000) has said that, “In science, guessing is called ‘hypothesis formation’”). Or, if you prefer, it is like detective work: finding clues to determine, not “who done it”, but “what it means”. And, like all hypotheses, theories, and conclusions drawn from circumstantial evidence, it is susceptible to revision when more evidence is found.

Admittedly, all of this assumes that the reader is consciously aware of the unfamiliar word and notes its unfamiliarity. It also assumes that the reader remembers the word and its hypothesized meaning (if any) between encounters. None of these further assumptions are, unfortunately, necessarily the case.

<sup>12</sup> Might a wider scope make it *harder* for the reader to identify passages that are relevant for CVA? We take a holistic view of meaning; thus, all passages are potentially relevant (Rapaport 2002). The issue that our definition algorithms help resolve is how to filter out a dictionary-like definition from this wealth of data (Rapaport & Ehrlich 2000).

**Nondirective Co-texts.** The next category is “nondirective contexts, which seem to be of *no* assistance in directing the reader toward any particular meaning for a word” (178, my emphasis). Here is Beck et al.’s example:

Dan heard the door open and wondered who had arrived. He couldn’t make out the voices. Then he recognized the *lumbering* footsteps on the stairs and knew it was Aunt Grace. (178.)

Again, the evidence is underwhelming, and for the same reasons: no evidence of the sole example being natural, no mention of any larger co-text that might provide more clues, and the word is a modifier (this time, an adjective).

Above, I suggested that the reader could ignore a single unfamiliar word in a misdirective text. The same is true of a non-directive text. But could an author use a word uniquely in such a way that it *is* crucial to understanding the text? Yes—authors can do pretty much anything they want. But, in such a case, the author would be assuming that the reader’s PK includes the author’s intended meaning for that word (again, see Simon 1996: 171, quoted above in §2.1). As a literary conceit, it might be excusable; in expository writing, it would not be.

**Syntactic Manipulation.** But even misdirective and non-directive co-texts (as well as the other categories yet to be explored) are capable of yielding a clue. The technique for squeezing a clue out of any co-text is to syntactically manipulate the co-text to make the unfamiliar word its focus, much as one syntactically manipulates an equation in one unknown to turn it into an equation with the unknown on one side of the equals sign and its “co-text” on the other (cf. Higginbotham 1985, 1989; Rapaport 1986). For example, from the above “misdirective” text, we could infer that, whatever else ‘grudgingly’ might mean, it could be defined (albeit only vaguely) as “a way of saying something” (and we could list all sorts of such ways, and hypothesize that ‘grudgingly’ is one of them). Moreover, it could be defined (albeit still vaguely) as “a way of (apparently) praising someone’s performance” (and we could list all sorts of such ways, and hypothesize that ‘grudgingly’ is one of them). I put ‘apparently’ in parentheses, because some readers, depending on their PK, might realize that sometimes praise can be given reluctantly or ironically, and such readers might hypothesize that ‘grudgingly’ is that kind of way of praising. Similarly, from the “lumbering” passage, a reader might infer that lumbering is a property of footsteps, or footsteps on stairs, or even a *woman’s* footsteps on stairs.<sup>13</sup>

**General Co-texts.** Not all co-texts containing modifiers are mis- or nondirective. The next category is a case in point: “general contexts . . . provide enough information for the reader to place the word in a general category” (178–179):

Joe and Stan arrived at the party at 7 o’clock. By 9:30 the evening seemed to drag for Stan. But Joe really seemed to be having a good time at the party. “I wish I could be as *regarious* as he is,” thought Stan. (129.)

Note that this adjective is contrasted with Stan’s attitude. From a contrast, much can be inferred. Indeed, in our research, two of the three adjectives that we have figured out meanings for occur in such contrastive co-texts:<sup>14</sup>

Unlike his brothers, who were noisy, outgoing, and very talkative, Fred was quite *taciturn*. (Dulin 1970.)<sup>15</sup>

Transylvania presents visitors with none of the logistical hurdles encountered in the hardscrabble lands to its east. Trains go almost everywhere, and tickets cost roughly two dollars an hour for first-class travel (first-class Romanian-style, that is, with *tatterdemalion* but comfortably upholstered compartments and equally *tatterdemalion* but solicitous attendants). (Tayler 1997.)<sup>16</sup>

<sup>13</sup> Nation 2001: 235f makes similar points about Beck et al. in general and ‘grudgingly’, in particular.

<sup>14</sup> The third adjective occurs in a (natural) co-text containing a parallel construction, which is equally useful: “In *The Pity of War* (1998), Ferguson argued that British involvement in World War I was unnecessary, far too costly in lives and money for any advantage gained, and a *Pyrrhic* victory that in many ways contributed to the end of the Empire” (Harsanyi 2003). For our computational analysis of this word in context, see Anger 2003.

<sup>15</sup> This is probably not a natural co-text, or else it is what Beck et al. call a “directive” co-text; see footnote 8, above, and the next subsection, below. For our computational analysis of this word in context, see Lammert 2002.

<sup>16</sup> For our computational analysis of this word in context, see Schwartzmyer 2004.

**Directive Co-texts.** Beck et al.'s fourth category is "directive contexts, which seem likely to lead the student to a specific, correct meaning for a word" (179). But, here, their example is that of a noun:

When the cat pounced on the dog, he leapt up, yelping, and knocked over a shelf of books. The animals ran past Wendy, tripping her. She cried out and fell to the floor. As the noise and confusion mounted, Mother hollered upstairs, "What's all the *commotion*?" (179.)

Again, it's not clear whether this is a natural co-text. But, more importantly, the fact that it is a noun suggests that it is not so much the co-text that is helpful as it is the fact that it is a noun, which is generally easier to learn than adjectives and adverbs. (Note, too, that this text is longer than the others!)

## 2.6 CVA, Neologisms, and Cloze-Like Tasks

Beck et al. conducted an experiment involving subjects who were given passages from basal readers. The researchers "categorized the contexts surrounding target words according to" their four-part "scheme", and they "then blacked out all parts of the target words, except morphemes that were common prefixes or suffixes . . . . Subjects were instructed to read each story and to try to fill in the blanks with the missing words or reasonable synonyms" (179).

Independent of the results, there are several problems with this set-up: (1) The passages may indeed have been found in the "natural" co-text of a basal reader, but were the stories in these anthologies written especially for use in schools, or were they truly natural? (Remember: One reader's natural co-text might be another researcher's pedagogical one; see §2.4.) (2) How large were the surrounding co-texts? Recall that a small co-text might be nondirective or even misdirective, yet a slightly larger one might very well be directive. (3) It is unclear whether the subjects were given any instruction on how to do CVA before the test. Here we find another assumption:

**Assumption A7** CVA "comes naturally", hence needs no training.

Our project, by contrast, is not focused on incidental CVA, but on deliberate CVA,<sup>17</sup> carefully taught and practiced.

(4) Another problem arises from the next assumption:

**Assumption A8** Cloze-like tasks are a form of CVA.

A "cloze-like" task involves replacing certain words in a passage with blanks to be filled in. But blacking out a word or asking the subject to fill in a blank is not quite the same task as CVA.

A serious methodological difficulty faces all CVA researchers: If you want to find out if a subject can figure out a meaning for an unknown word from context, you don't want to use a word that the subject knows. You could filter out words (or subjects) by giving a pretest to determine whether the subjects know the test words. But then those who don't know them will have seen them at least once before (during the pretest), which risks contaminating the data. Finding obscure words (in natural co-texts, no less) that are highly unlikely to be known by any subjects is difficult; in any case, one might want to test familiar words. Two remaining alternatives—replace the word with a neologism or a blank—introduce complications: In research that we have done with think-aloud protocols of students doing CVA (reported orally, Kibby et al. 2004), we have found that, when students confront what they believe to be a real (but unknown) word, they focus their attention, thoughts, and efforts on meaning (i.e., what could this word mean?), but when obvious neologisms or blank spaces are used, readers focus on "getting" the word, not on expressing its possible meaning. These tasks are probably related, but they are distinctly different, too.<sup>18</sup> Schatz & Baldwin (1986: 450) also claim that "Using context to guess the meaning of a semantically unfamiliar word is essentially the same as supplying the correct meaning in a cloze task." But this is not the case: In cloze-like tasks, the reader is invited to guess (rather than figure out), and there is a unique, correct answer, whereas, in CVA as we see it, the goal is to figure out a meaning that is sufficient for understanding the passage. (Wolfe 2003 discusses the related problem of inventing names in social science.)

Currently, I have no clever solution to this methodological problem. My preferred technique for now is to use a plausible-sounding neologism (with appropriate affixes) and then to inform the subject that it is a word from another language that might or might not have a single-word counterpart in English, but that in any case the subject's job is to figure out what it might mean, not necessarily find an English synonym, exact or inexact.

<sup>17</sup> On this distinction, cf. Hulstijn 2003.

<sup>18</sup> I am grateful to my co-researcher, Michael Kibby, for this insight.



## 2.7 Beck et al.'s Conclusions

Beck et al. claim that their experiment “clearly support[s] the categorization system” and “suggest[s] that it is precarious to believe that naturally occurring contexts are sufficient, or even generally helpful, in providing clues to promote initial acquisition of a word’s meaning” (179). Significantly, however, “Only one subject could identify any word in the misdirective category” (179). This is significant, not because it supports their theory, but for almost the opposite reason: It suggests that CVA *can* be done even with misdirective co-texts, which supports *our* theory, not theirs.

They conclude that “Children most in need of vocabulary development—that is, less skilled readers who are unlikely to add to their vocabularies from outside sources—will receive little benefit from such indirect opportunities to gain information” (180–181). An assumption underlying this conclusion is that:

**Assumption A9** CVA can be of help only in vocabulary acquisition.

But another potential benefit far outweighs this: CVA strategies, if properly taught and practiced, can improve general reading comprehension. This is because the techniques that our computational theory employs and that, we believe, can be taught to readers, are almost exactly the techniques needed for improving reading comprehension: careful, slow reading; careful analysis of the text; a directed search for information useful to computing a meaning; application of relevant PK; application of reasoning for the purpose of extracting information from the text.

## 3 Are Context Clues Unreliable Predictors of Word Meanings?

Schatz & Baldwin 1986 takes the case against context a giant step further, arguing “that context does not usually provide clues to the meanings of low-frequency words, and that context clues actually inhibit the correct prediction of word meanings just as often as they facilitate them” (440).<sup>19</sup>

### 3.1 Schatz & Baldwin’s Argument

In summarizing the then-current state of the art, they ironically note that “almost eight decades after the publication of . . . [a] classic text [on teaching reading] . . . , publishers, teachers, and the authors of reading methods textbooks have essentially the same perception of context as an *efficient* mechanism for inferring word meanings” (440, my emphasis). Given their rhetoric, the underlying assumption here appears to be:

**Assumption A10** CVA is *not* an efficient mechanism for inferring word meanings.

Their argument seems to be roughly the following:

*Textual* context can’t help you figure out “the” correct meaning of an unfamiliar word.  
Therefore, CVA is not “an effective strategy for inferring word meanings”. (440)

In contrast, I am arguing in this paper that:

*Wide* context *can* help you figure out *a* meaning for an unfamiliar word.  
Therefore, CVA *is* an effective strategy for inferring (better: figuring out, or computing) word meanings.

As with Beck et al., note that the issue concerns the *purpose* of CVA. Insofar as its purpose is thought of as getting “the correct meaning”, it is ineffective. But insofar as its purpose is seen to be getting a meaning sufficient for understanding the passage in which the unfamiliar word occurs, it can be quite effective, even with an allegedly “misdirective” co-text.

Perhaps CVA is thought to be too magical, or perhaps too much is expected of it. Schatz & Baldwin claim that, “According to the current research literature, context clues should help readers to infer the meanings of . . . [unfamiliar] words . . . *without the need for readers to interrupt the reading act* with diversions to . . . dictionaries, or other external sources of information” (441, my emphasis). This could only be the case if CVA were completely unconscious and immediate, as if one could read a passage with an unfamiliar word and instantaneously come to know what it means.<sup>20</sup> In contrast, our theory requires interruption—not to access external sources—but for conscious, deliberate analysis of

<sup>19</sup> All page references in this section will be to Schatz & Baldwin 1986, unless otherwise noted.

<sup>20</sup> This *may* hold for “incidental” CVA (Nagy et al. 1985), but not for “deliberate” CVA.

the passage. Computer models that appear to work instantaneously are actually doing quite a lot of active processing, which a human reader would need much more time for.

In any case, stopping to consult a dictionary does not suffice. With the exception of learner's dictionaries designed primarily for ESL audiences, most dictionaries are notoriously difficult to use and their definitions notoriously difficult to interpret (Miller 1985, 1986). More importantly, *CVA needs to be applied to the task of understanding a dictionary definition itself*, which is, after all, merely one more co-text containing the unfamiliar word (Schwartz 1988: 111). Indeed, CVA is the base case of a recursion one of whose recursive clauses is "look it up in a dictionary". An anecdote might illustrate: My 8-year-old son was reading a book and asked me what a certain word meant. I asked him to read the entire sentence (i.e., to give me the word's co-text). It turned out that the sentence was a definition of the word, but he didn't realize that; he needed to do CVA on this defining context.

### 3.2 Schatz & Baldwin's Methodology

Schatz & Baldwin offer the results of several experiments to support their claims. As with Beck et al.'s experiments, there are a number of apparent problems with their methodology.

**Nouns and Verbs vs. Modifiers.** Their first experiment took 25 "natural" passages from novels, selected according to an algorithm that randomly produced passages containing low-frequency words. But consider some of the words they chose (442): 'cogently', 'cozened', 'ignominiously', 'imperious', 'inexorable', 'perambulating', 'recondite', 'salient'. Note that four (or 50%) are adjectives, two (25%) are adverbs, one (12.5%) is a verb ('cozened'), and one ('perambulating') might be either a verb or an adjective, depending on the co-text. These are only "examples"; we are not given a full list of words, nor told whether these statistics are representative of the full sample. But, if they are, then fully 75% of the unfamiliar words are modifiers, known to be among the most difficult of words to learn meanings for. Schatz & Baldwin's example passages consist of an adverb ('ruefully'), three adjectives ('glib', 'pragmatic', 'waning'), and four nouns ('yoke', 'coelum', 'dearth', 'ameliorating'). This brings the statistics to around 67% for adjectives and adverbs, 27% for nouns, and 6% for verbs (not counting 'perambulating'). Of these, two of the nouns ('dearth', 'ameliorating') are presented as examples of words occurring in "facilitative" co-texts (448). Their example of a "confounding" co-text is for an adjective ('waning').

These examples raise more questions than they answer: What were the actual percentages of modifiers vs. nouns and verbs? Which lexical categories were hardest to determine meanings for? How do facilitative and confounding contexts correlate with lexical category? Schatz & Baldwin observe that, among "potential limitations" of their experiments, "a larger sample of words would certainly be desirable" but that their selection of "70 items ... offer[s] a larger and more representative sample than most studies of context clues" (449). But a representative sample of what? Of co-texts? Or of words? The sort of representativeness that is needed should (also) be a function of the variety of lexical category. What would happen with natural co-texts of, say, all four of Beck et al.'s categories with nouns, verbs, adjectives, and adverbs in each such co-text (i.e., 16 possible types of co-text)? Schatz & Baldwin's (and Beck et al.'s) results may say more about the difficulty of learning meanings for modifiers than they do about weaknesses of contexts.

**CVA vs. Word-Sense Disambiguation.** It should be noted, moreover, that, in two of Schatz & Baldwin's experiments, subjects were *not* involved in the task of CVA. Rather, they were doing a related—but distinct—task known as "word-sense disambiguation" (WSD; Ide & Veronis 1998). The CVA task is to figure out a word's meaning "from scratch". The WSD task is to choose a meaning for a word from a list of given meanings; typically, the word is polysemous, and all items on the list are possible meanings for the word in different contexts. Thus, WSD is a multiple-choice test, whereas CVA is an essay question.<sup>21</sup> In Schatz & Baldwin's experiment, all that the subjects had to do was to replace the unfamiliar word with each multiple-choice meaning-candidate (each of which was a proposed one-word synonym) and see which of those five possible meanings fit better; no real CVA was needed.

In the third experiment, real CVA was being tested. However, assumption A3 (about correct meanings) raises its head: "we were interested only in full denotative meanings or accurate synonyms" (446). There is no reason to believe or to expect that CVA will typically be able to deliver on such a challenge. But neither is there any reason to demand such high standards, and, once this constraint is relaxed, CVA can be seen to be a useful tool for vocabulary acquisition and general reading comprehension.

<sup>21</sup> Ellen Prince, personal communication.

**Space and Time Limits.** Another issue concerns the size of the co-text. The smaller the co-text, the less chance there is of figuring out a meaning, for the simple reason that there will be a minimum of textual clues. The larger the co-text, the greater the chance, for the simple reason that a large enough co-text might actually include a definition of the word! (Recall from §3.1, however, that CVA needs to be applied even in the case of an explicit definition!)

What is a reasonable size for a co-text? Our methodology has been to start small and work “outwards” to preceding and succeeding passages, until enough co-text is provided to enable successful CVA. (Here, of course, ‘successful’ only means being able to figure out *a* meaning enabling the reader to understand enough of the passage to continue reading; it does *not* mean figuring out “the correct meaning of” the word.) This models what readers can do when faced with an unfamiliar word in normal reading: They are free to examine the rest of the text for possible clues. In contrast, Schatz & Baldwin limited their co-text size to 3 sentences (typically, the preceding sentence, the sentence containing the unfamiliar word, and the succeeding sentence). This is arbitrary and too small. An inability to do CVA from such a limited co-text shows at most that such co-texts are too small, not that CVA is unhelpful.

Yet another issue concerns time limits. Schatz & Baldwin do not tell us what limits were set, but do observe that “All students finished in the allotted time” (443). But real-life CVA has no time limits (other than self-imposed ones), and CVA might extend over a long period of time, as different texts are read.

**Teaching CVA Techniques.** Finally, there was no prior training in how to use CVA: “we did not control for the subjects’ formal knowledge of how to use context clues” (449). Their finding “that students either could not or chose not to use context to infer the meanings of unknown words” (444) ignores the possibilities that the students did not know that they *could* use context or that they did not know *how* to. Granted, “incidental” (or unconscious) CVA is something that we all do; there appears to be no other explanation for how we learn most of our vocabulary (Nagy and colleagues 1984, 1985, 1987). But “deliberate” (or conscious) CVA is a skill that, while it may come naturally to some, can—and needs—to be taught, modeled, and practiced.

Thus, Schatz & Baldwin’s conclusion that “context is an ineffective or little-used strategy for helping students infer the meanings of low-frequency words” (446) might only be true for untrained readers. It remains an open question whether proper training in CVA can make it effective and can add it to the reader’s arsenal of techniques for improving reading comprehension (though there is some positive evidence; see Fukkink & De Glopper 1998, Kuhn & Stahl 1998, Swanborn & De Glopper 1999, Baumann et al. 2002). Schatz & Baldwin disagree:

It is possible that if the subjects had been given adequate training in using context clues, the context groups in these experiments might have performed better. We think such a result would be unlikely because the subjects were normal, fairly sophisticated senior high school students. If students don’t have contextual skills by this point in time, they probably are not going to get them at all. (449.)

Assumption A7 (that CVA needs no training) is at work again. But students are not going to get “contextual skills” if they are not shown the possibility of getting them. Moreover, the widespread need for, and success of, critical thinking courses—not only at the high school level, but also in post-secondary education—strongly suggests that students need to, and can, be educated on these matters.

### 3.3 Three Questions about CVA

In their general-discussion section, Schatz & Baldwin raise three questions (447; emphasis in original):

(1) “Do traditional context clues occur with sufficient frequency to justify them as a major element of reading instruction?” This is irrelevant *if* CVA can be shown to foster good reading comprehension and critical-thinking skills. For clues need not occur frequently in order for the techniques for using them to be useful general skills. We believe that CVA can foster improved reading comprehension, but much more research is needed. Our answer to Schatz & Baldwin’s first question is this: Traditional context clues do occur and—augmented by the reader’s PK and training in CVA techniques for developing revisable hypotheses about an unfamiliar word’s meaning—are justified as a major element of reading instruction.

(2) “Does context *usually* provide accurate clues to the denotations and connotations of low-frequency words?” This is also irrelevant under our conception of CVA: We are not interested in “accuracy”. Moreover, a “denotation” (in the sense of an external referent of a word) is best provided by demonstration or by a graphic illustration, and a “connotation” (in the sense of an association of the unfamiliar word with other (familiar) words) is not conducive to the

sort of “accuracy” that Schatz & Baldwin (or Beck et al.) seem to have in mind. Our answer to this question is similar to our answer to the first: Context *can* provide clues to revisable hypotheses about an unfamiliar word’s meaning.

(3) Are “difficult words in naturally occurring prose . . . usually amenable to such analysis”? Our theory holds that such words are always amenable to yielding at least some information about their meaning, as discussed in §2.5 (“Syntactic Manipulation”), above.

## 4 Conclusions: A Positive Theory of Computational CVA

Progress is often made by questioning assumptions (Rapaport 1982). This essay has questioned the assumptions underlying Beck et al.’s and Schatz & Baldwin’s arguments and experiments that challenge CVA. Their papers are best read as asserting that, *given those assumptions*, CVA is not as beneficial as some researchers claim it is. We conclude by presenting our theory’s contrasting beliefs. Readers interested in the details of our computational implementation and algorithm-based curriculum project are directed to Rapaport & Ehrlich 2000 and Rapaport & Kibby 2002.

**B1** Every context can give some clue (albeit only minimal) to a word’s meaning.

More precisely: For every textual context  $C$ , and for every word  $w$  in  $C$ ,  $w$  has a meaning in  $C$  (at the very least, its “algebraic” meaning obtained by rephrasing  $C$  to make  $w$  the subject). But  $w$  will also have a meaning that is partly determined by the reader  $R$  (and his or her accessible PK). And, since the reader’s PK may be time-dependent, we need to add “at time  $t$ ” to this formulation. Thus, for every  $w, C, R, t$ , if  $w$  is in  $C$ , and  $R$  reads  $C$  at  $t$ , then  $w$  has a meaning in  $C$  for  $R$  at  $t$ . And, of course, none of the meanings that  $w$  has for  $R$  is necessarily “the” meaning (in either a dictionary sense or that of a reading teacher).

**B2** The surrounding textual context of a word (its co-text) gives clues to a word’s meaning that must be supplemented by the reader’s PK in order for a meaning to be figured out (i.e., computed).

**Corollary B2.1** There is no such thing as a “good” or “bad” co-text *simpliciter* (or a “misdirective”, “non-directive”, “general”, or “directive” co-text, either); the value of a co-text depends on the reader’s PK and ability to use clues and PK together.

**B3** CVA is distinct from cloze-like tasks.

**B4** Co-texts can be as small as a phrase or as large as an entire book;<sup>22</sup> there are no arbitrary limits.

**B5** Many contexts may be required before CVA can “asymptotically” approach a “stable” meaning for a word.

**B6** A word does not have a unique meaning, even in directive and pedagogical contexts.

**B7** A word does not have a (single) correct meaning, not even in directive and pedagogical contexts.

**Alternative B7.1** A word does not *need* a correct meaning (nor does any such correct meaning need to be understood) in order for a reader to be able to understand the word (in context).

**Corollary B7.2** Even a familiar and well-known word can acquire a new meaning in a new context.

In fact, each new  $C$  and each new  $R$  can yield a new meaning, so meanings are continually being extended. (This often happens when words are used metaphorically; cf. Budiu & Anderson 2001, Rapaport forthcoming.)

**B8** Some words are harder to figure out meanings for than others (nouns are easiest, verbs a bit harder, modifiers the hardest).

**B9** CVA is an efficient method for inferring word meanings.

**B10** CVA can improve general reading comprehension.

**B11** CVA can (and should) be taught.<sup>23</sup>

<sup>22</sup> Cf. Quine 1951: “The unit of empirical significance is the whole of science.” See also Rapaport 2002.

<sup>23</sup> I am grateful to Albert Goldfain, Michael W. Kibby, Jean-Pierre Koenig, Shakthi Poornima, Stuart C. Shapiro, and the members of the SNePS Research Group for comments and advice on earlier drafts.

## References

- Anderson, R.C. (1984), "Role of the Reader's Schema in Comprehension, Learning, and Memory", in R.C. Anderson et al. (eds.), *Learning to Read in American Schools* (Erlbaum): 243–257.
- Anger, B. (2003), "Understanding It All: Defining 'Pyrrhic Victory' from Context" [<http://www.cse.buffalo.edu/~rapaport/CVA/Pyrrhic/Pyrrhic-Victory-Report.pdf>].
- Baumann, J.F.; Edwards, E.C.; Front, G.; Tereshinski, C.A.; Kame'enui, E.J.; & Olejnik, S. (2002), "Teaching Morphemic and Contextual Analysis to Fifth-Grade Students", *Reading Res. Q*ly. 37: 150–176.
- Budiu, R. & Anderson, J.R. (2001), "Word Learning in Context: Metaphors and Neologisms", *Tech. Rep. CMU-CS-01-147* (Carnegie Mellon Univ., School of Comp. Sci.).
- Beck, I.L.; McKeown, M.G.; & McCaslin, E.S. (1983), "Vocabulary Development: All Contexts Are Not Created Equal", *Elementary School Journal* 83: 177–181.
- Brown, G., & Yule, G. (1983), *Discourse Analysis* (Cambridge Univ. Press).
- Clarke, D.F., & Nation, I.S.P. (1980), "Guessing the Meanings of Words from Context", *System* 8: 211–220.
- Cunningham, A.E., & Stanovich, K.E. (1998), "What Reading Does for the Mind", *American Educator* (Spring/Summer): 1–8.
- Dulin, K.L. (1969), "New Research on Context Clues", *J. Reading* 13: 33–38, 53.
- Dulin, K.L. (1970), "Using Context Clues in Word Recognition and Comprehension", *Reading Teacher* 23: 440–445, 469.
- Ehrlich, K. (1995), "Automatic Vocabulary Expansion through Narrative Context", *Tech. Rep. 95-09* (SUNY Buffalo Comp. Sci.).
- Ehrlich, K. (2004), "Default Reasoning Using Monotonic Logic", *Proc. 15th Midwest AI & Cog. Sci. Conf.*: 48–54.
- Ehrlich, K., & Rapaport, W.J. (1997), "A Computational Theory of Vocabulary Expansion", *Proc. 19th Annual Conf. Cog. Sci. Soc.* (Erlbaum): 205–210.
- Ehrlich, K., & Rapaport, W.J. (2004), "A Cycle of Learning: Human and Artificial Contextual Vocabulary Acquisition", *Proc. 26th Annual Conf. Cog. Sci. Soc.* (Erlbaum, 2005): 1555.
- Fukkink, R.G., & De Gloppe, K. (1998), "Effects of Instruction in Deriving Word Meaning from Context", *Rev. Ed. Res.* 68: 450–458.
- Garnham, A., & Oakhill, J. (1990), "Mental Models as Contexts for Interpreting Texts", *J. Semantics* 7: 379–393.
- Gentner, D. (1981), "Some Interesting Differences between Verbs and Nouns", *Cogn. & Brain Theory* 4: 161–178.
- Gentner, D. (1982), "Why Nouns Are Learned before Verbs", in S.A. Kuczaj (ed.), *Language Development: Vol. 2* (Erlbaum): 301–334.
- Gillette, J.; Gleitman, H.; Gleitman, L.; & Lederer, A. (1999), "Human Simulations of Vocabulary Learning", *Cogn.* 73: 135–176.
- Granger, R.H. (1977), "Foul-Up: A Program that Figures Out Meanings of Words from Context", *Proc. IJCAI-77*: 67–68.
- Haastруп, K. (1991), *Lexical Inferencing Procedures or Talking about Words* (Gunter Narr Verlag Tübingen).
- Harsanyi, D. (2003), "The Old Order", *Nat'l. Rev.* (May 5): 45–46.
- Higginbotham, J. (1985), "On Semantics", in E. Lepore (ed.), *New Directions in Semantics* (Academic, 1987): 1–54.
- Higginbotham, J. (1989), "Elucidations of Meaning", *Ling. & Phil.* 12: 465–517.
- Hirsch, Jr., E.D. (1987), *Cultural Literacy* (Houghton Mifflin).
- Hirsch, Jr., E.D. (2003), "Reading Comprehension Requires Knowledge—of Words and of the World", *American Educator* 27 (Spring): 10–13, 16–22, 28–29, 48.
- Hobbs, J.R. (1990), *Literature and Cognition* (Stanford, CA: CSLI).
- Hobbs, J.R.; Stickel, M.; Appelt, D.; & Martin, P. (1993), "Interpretation as Abduction", *Artif. Intell.* 63: 69–142.
- Hulstijn, J.H. (2003), "Incidental and Intentional Learning", in C.J. Doughty & M.H. Long (eds.), *Handbook of Second Language Acquisition* (Blackwell): 349–381.
- Ide, N.M., & Veronis, J. (eds.) (1998), Special Issue on Word Sense Disambiguation, *Comp. Ling.* 24(1).
- Johnson-Laird, P.N. (1983), *Mental Models* (Harvard Univ. Press).
- Johnson-Laird, P.N. (1987), "The Mental Representation of the Meanings of Words", *Cogn.* 25: 189–211.
- Kibby, M.W.; Rapaport, W.J.; & Wieland, K.M. (2004), "Helping Students Apply Reasoning to Gaining Word Meanings from Context: Application of AI and Verbal Protocol Data to Curriculum Design", *49th Annual Conv. Int'l. Reading Assoc.*
- Kuhn, M.R., & Stahl, S.A. (1998), "Teaching Children to Learn Word Meanings from Context", *J. Lit. Res.* 30(1): 119–138.
- Lakoff, G. (1987), *Women, Fire, and Dangerous Things* (Univ. Chicago Press).
- Lammert, A. (2002), "Defining Adjectives through Contextual Vocabulary Acquisition", [[http://adamlammert.com/school/cva/cva\\_report.pdf](http://adamlammert.com/school/cva/cva_report.pdf)].
- Lenat, D.B. (1995), "CYC", *CACM* 38(11): 33–38.
- Loui, M.C. (2000), "Teaching Students to Dream", [<http://wocket.csl.uiuc.edu/~loui/gtc00.html>].
- Miller, G.A. (1985), "Dictionaries of the Mind", *Proc. 23rd Annual Meeting Assoc. Comp. Ling.*: 305–314.
- Miller, G.A. (1986), "Dictionaries in the Mind", *Lang. & Cog. Procs.* 1: 171–185.
- Murphy, J. (2000), *The Power of Your Subconscious Mind*, *Rev. Ed.* (Bantam).
- Nagy, W.E., & Anderson, R.C. (1984), "How Many Words Are There in Printed School English?", *Reading Res. Q*ly. 19: 304–330.

- Nagy, W.E. & Herman, P.A. (1987), "Breadth and Depth of Vocabulary Knowledge", in M.G. McKeown & M.E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Erlbaum): 19–35.
- Nagy, W.E.; Herman, P.A.; & Anderson, R.C. (1985), "Learning Words from Context", *Reading Res. Qlty.* 20: 233–253.
- Nation, I.S.P. (2001), *Learning Vocabulary in Another Language* (Cambridge Univ. Press).
- Nation, P., & Coady, J. (1988), "Vocabulary and Reading", in R. Carter & M. McCarthy (eds.), *Vocabulary and Language Teaching* (Longman): 97–110.
- Quine, W.V.O. (1951), "Two Dogmas of Empiricism", in W.V. Quine, *From a Logical Point of View, 2nd ed., rev.* (Harvard Univ. Press, 1980): 20–46.
- Rapaport, W.J. (1982), "Unsolvable Problems and Philosophical Progress", *Amer. Phil. Qlty.* 19: 289–298.
- Rapaport, W.J. (1986), "Searle's Experiments with Thought", *Phil. Sci.* 53: 271–279.
- Rapaport, W.J. (2000), "How to Pass a Turing Test", *J. Logic, Lang., Info.*, 9: 467–490.
- Rapaport, W.J. (2002), "Holism, Conceptual-Role Semantics, and Syntactic Semantics", *Minds & Machines* 12: 3–59.
- Rapaport, W.J. (2003a), "What Is the 'Context' for Contextual Vocabulary Acquisition?", in P.P. Slezak (ed.), *Proc. 4th Int'l. Conf. Cog. Sci./7th Australasian Soc. Cog. Sci. Conf. (ICCS/ASCS-2003)* (Univ. New South Wales), Vol. 2, pp. 547–552.
- Rapaport, W.J. (2003b), "What Did You Mean by That? Misunderstanding, Negotiation, and Syntactic Semantics", *Minds & Machines* 13: 397–427.
- Rapaport, W.J. (compiler) (2004), "A (Partial) Bibliography (in Chronological Order) of (Computational) Theories of Contextual Vocabulary Acquisition", [<http://www.cse.buffalo.edu/~rapaport/refs.vocab.html>].
- Rapaport, W.J. (forthcoming), "The Turing Test", *Encyc. Lang. & Ling., 2nd ed.* (Elsevier).
- Rapaport, W.J., & Ehrlich, K. (2000), "A Computational Theory of Vocabulary Acquisition", in L.M. Iwańska & S.C. Shapiro (eds.), *Natural Language Processing and Knowledge Representation* (AAAI/MIT Press): 347–375.
- Rapaport, W.J., & Kibby, M.W. (2002), "Contextual Vocabulary Acquisition: A Computational Theory and Educational Curriculum", in N. Callaos et al. (eds.), *Proc. 6th World Multiconf. on Systemics, Cybernetics & Informatics (SCI 2002)* (Orlando: Int'l. Inst. Informatics & Systemics), Vol. II, pp. 261–266.
- Russell, B. (1918), "The Philosophy of Logical Atomism", in B. Russell, *Logic and Knowledge* (Capricorn, 1956): 177–281.
- Schatz, E.K., & Baldwin, R.S. (1986), "Context Clues Are Unreliable Predictors of Word Meanings", *Reading Res. Qlty.* 21: 439–453.
- Schwartz, R.M. (1988), "Learning to Learn Vocabulary in Content Area Textbooks", *J. Reading*: 108–118.
- Schwartzmyer, N.P. (2004), "Tatterdemalion: A Case Study in Adjectival Contextual Vocabulary Acquisition" [[http://www.cse.buffalo.edu/~rapaport/CVA/Tatterdemalion/Tatterdemalion\\_final\\_report.pdf](http://www.cse.buffalo.edu/~rapaport/CVA/Tatterdemalion/Tatterdemalion_final_report.pdf)].
- Shapiro, S.C., & Rapaport, W.J. (1987), "SNePS Considered as a Fully Intensional Propositional Semantic Network," in N. Cercone & G. McCalla (eds.), *The Knowledge Frontier* (Springer-Verlag): 262–315.
- Shapiro, S.C., & Rapaport, W.J. (1992), "The SNePS Family," *Comp. & Math. with Applications* 23: 243–275.
- Shapiro, S.C., & Rapaport, W.J. (1995), "An Introduction to a Computational Reader of Narrative," in J.F. Duchan et al. (eds.), *Deixis in Narrative* (Erlbaum): 79–105.
- Shapiro, S.C., & the SNePS Implementation Group (2004), "SNePS 2.6.1 User's Manual", [<http://www.cse.buffalo.edu/sneps/Manuals/manual261.ps>].
- Simon, H.S. (1996), "Computational Theories of Cognition", in W. O'Donohue & R.F. Kitchener (eds.), *Philosophy of Psychology* (Sage): 160–172.
- Singer, M.; Revlin, R.; & Halldorson, M. (1990), "Bridging-Inferences and Enthymemes", in A.C. Graesser & G.H. Bower (eds.), *Inferences and Text Comprehension* (Academic): 35–51.
- Stanovich, K.E. (1986), "Matthew Effects in Reading", *Reading Res. Qlty.* 21: 360–407; reprinted in R.B. Ruddell & N.J. Unrau (eds.), *Theoretical Models and Processes of Reading, 5th Ed.* (Int'l. Reading Assoc., 2004): 454–516.
- Sternberg, R.J. (1987), "Most Vocabulary is Learned from Context", in M.G. McKeown & M.E. Curtis (eds.), *The Nature of Vocabulary Acquisition* (Erlbaum): 89–105.
- Sternberg, R.J., & Powell, J.S. (1983a), "Comprehending Verbal Comprehension", *Amer. Psychologist* 38: 878–893.
- Sternberg, R.J.; Powell, J.S.; & Kaye, D.B. (1983b), "Teaching Vocabulary-Building Skills: A Contextual Approach", in A.C. Wilkinson (ed.), *Classroom Computers and Cognitive Science* (Academic): 121–143.
- Suh, S., & Trabasso, T. (1993), "Inferences during Reading", *J. Memory & Lang.* 32: 279–300.
- Swanborn, M.S.L., & De Glopper, K. (1999), "Incidental Word Learning while Reading", *Rev. Ed. Res.* 69: 261–285.
- Taylor, J. (1997), "Transylvania Today", *Atlantic Monthly* 279(6): 50–54.
- Waite, M. (ed.) (1998), *Little Oxford Dictionary, Rev. 7th Ed.* (Oxford Univ. Press).
- Webster, M.A. (2004), "Adaptation and the Phenomenology of Perception", unpublished talk, SUNY Buffalo Center for Cog. Sci., [<http://wings.buffalo.edu/cogsci/Activities/Colloquium/CLLQf04/2004fall.htm#webster04>]
- Wolfe, A. (2003), "Invented Names, Hidden Distortions in Social Science", *Chronicle Higher Ed.* (May 30): B13–B14.