UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Levels of Semantic Constraint and Learning Novel Words

Permalink

Journal

Authors

Lampinen, James M.
Faries, Jeremiah M .

Publication Date

1994

Peer reviewed

Levels of Semantic Constraint and Learning Novel Words

James M. Lampinen
Department of Psychology
Northwestern University
Evanston, Illinois,60201
jim@swim-two-birds.psych.nwu.edu

Jeremiah M. Faries
Department of Psychology and the Institute for the
Learning Sciences
Northwestern University
Evanston, Illinois,60201
jerry@swim-two-birds.psych.nwu.edu

Abstract

A common method of teaching vocabulary involves presenting students with new words in context and having the students derive the meaning of these words based on contextual cues. Beck, McKeown and McCaslin (1983) have argued that the contexts used to teach new words should be highly constraining. Although highly constraining contexts avoid ambiguity they do not present the learner with the necessity of combining contextual and word specific information and thus practicing skills needed for general comprehension. We suggest that a superior method of teaching is to relax the amount of contextual constraint because to optimize the learning from the presentation of a sentence the student must use both top down and bottom up processes to discover the meaning of the sentence, thus integrating two sources of knowledge about the word. The present research compares knowledge and use of newly learned words between students who learned the new words using three encounters with highly constraining contexts, three encounters with moderately constraining contexts or three progressively less constraining contexts. Students were given definitional and comprehension tests both immediately after study and at a one week delay. The results suggest that repeated encounters with moderately constraining contexts are superior to repeated encounters with highly constraining contexts.

When people learn new words they typically do so by combining information about the way the word is used in the context of the utterance with information from their pre-established knowledge of the word. In fact, most words are learned contextually and incidentally (e.g., Nagy and Herman, 1987). Moreover, comprehension of word meanings almost always occurs within a context and arguably the processes used to understand general contextual meaning contribute greatly to the activation of what we think of as word meaning. Identification of words is easier and faster when preceded by a relevant context (Schuberth and Eimas, 1977) suggesting that contextual information is an important component for processing individual words. Learning words in the absence of these general comprehension processes (e.g., presenting word-definition pairs to study) may produce an inferior record of word meaning insofar as that word can be used and comprehended quickly and naturally (although such learning may facilitate assimilation for words when encountered in context). Researchers, therefore, who are interested in directly teaching word meanings have explored ways to use context as a teaching method (Kolich, 1991).

Debate about the merits of contextual versus definitional methods (see Stahl and Fairbanks, 1986 for an excellent review of this literature), has often amounted to comparing optimal strategies within one broad instructional method with suboptimal strategies within another broad instructional method (see Hall, 1988). Since any successful approach to vocabulary instruction is likely to include both contextual and definitional components (Stahl and Fairbanks, 1986), it may be more constructive to investigate the boundary conditions that determine the effectiveness of both contextual and definitional methods, rather than to pit one against the other.

The goal of our present research is to determine the set of conditions under which contextual methods of instruction are likely to be most effective. One important variable in the construction of study sentences is how tightly the sentence constrains the meaning of a target word. If one were to replace the target word with a blank line (as in a cloze test) and the meaning intended by the blank line were easily inferred, the sentence would be highly constraining. It has been suggested that contexts used in teaching new words should, in fact, be highly constraining (Beck, McKeown and McCaslin, 1983). They suggest that the contexts which students use to learn new words should quite narrowly limit the number of possible definitions of the new word. A supporting argument for the use of highly constraining contexts is that they make word meaning clear and unambiguous and thus easier to derive. On the other hand, Schank (1982) has argued that students who study a word in a highly constraining context do not have to pay any attention at all to the new word since the meaning of the utterance is clear without knowing the word. If they make any association between the derived meaning and the new word it is incidental and not required by the task. In a study of eye fixation patterns, Ehrlich and Rayner (1981) found that target words from highly constrained passages were not fixated on or were fixated on more briefly than the same words in a poorly constrained passage. Further, the situations in which knowing a word's meaning is most crucial are precisely those situations in which the word's meaning is least likely to be deduced from context. This suggests that moderately constraining contexts might be superior to highly constraining contexts because

it is only in moderately constraining contexts that the reader needs to integrate contextual information with prior knowledge about the word.

The activation of word meaning in everyday use of language typically involves combining top down contextual information with bottom up word specific perceptual information. Vocabulary instruction which uses only highly constraining contexts emphasizes only the use of top down contextual information. Taylor and Taylor (1983) argue that good readers are superior to poor readers because they both use context more efficiently and because they use word specific information more efficiently. Interestingly, they also suggest that good readers do not use context by simply guessing what the target word should be (as might be encouraged by high constraint contexts) but, rather, they "narrow... down alternatives" (as might be encouraged by moderate and low constraint contexts).

A model of word learning should include learning of how the word fits into a variety of contexts as well as the learning of word specific associations. Consider for instance, how word meaning is accessed in an interactive reading model like the one proposed by Adams(1990). Adams proposes three separate processing modules all feeding into a central meaning processor. Phonological, orthographic and context processors all combine to activate word meaning.[1] Accessed meaning then becomes part of the the current text representation available in the context processor. Stahl(1991) has proposed that in learning word meanings from context, the meaning produced by this interaction of contextual and word specific information then becomes linked to the orthographic and phonological representations of the word.[2] This account seems plausible as far as it goes and argues for learning words in a context. However, for the association to be made between the contextual and word specific information it is important that such an association be required as part of the learning task. The strength of the link must depend crucially on how strongly the word is attended to and how important the word's meaning is to the overall meaning of the utterance. As we've seen above, high constraint contexts require

---

[1] Adams' model is discussed here as an illustrative example. The crucial point is that many models of reading propose that word meaning activation involves both the orthographic and phonological features of the word and expectations generated by the context.

[2] Word meaning is represented in Adams' model in a semantic network. Thus a word's meaning is to be understood in terms of the word's relationship to other concepts. A particular contextual instantiation of a word is likely to pick out only a few of those semantic relations. Moderately constraining contexts require that the reader combine the semantic information garnered in previous encounters with the word with the new semantic information garnered in the current context. Thus moderately constraining contexts encourage the development of a rich integrated conceptual understanding of the new word.

readers to attend only minimally to the target word because the context itself is heavily redundant with the meaning of the word. When readers encounter words in moderately constraining contexts, the story is different. In these contexts, the word's meaning can only be derived by combining information from the context processor with information from the orthographic and/or phonological processor. Thus for readers in this situation, the word itself is crucial to the understanding of the passage.

Ideally, to provide practice using the skills they will eventually need, students must encounter words in sentences of varying degrees of contextual constraint and the goal of these encounters should be to encourage students to combine contextual information with word specific information to quickly access word meaning and utterance meaning. At the limit students should even be able to access word meaning when the contextual constraint is weak and the word itself contributes greatly to the intent of the utterance.

Pilot research from our lab suggested that readers learned new words most efficiently when they were provided with study sentences in which the amount of constraint provided was progressively decreased. We reasoned that words to be learned in context should be seen in moderate and low constraint contexts in addition to high constraint contexts and that the presentations of these mixed contexts should progress from high constraint sentences to low constraint sentences. Progressively decreasing contextual constraint, encourages students to integrate new contextual information with information garnered from previous encounters with the word, but it also allows students to encounter the new word in ever more difficult contexts without fear of reaching an impasse, and is thus similar to the scaffolding provided by intelligent tutoring systems.

In the pilot study, subjects were presented with spaced presentations with three example sentences for each of 40 words. These example sentences were either highly constraining, or they progressed from highly constraining to minimally constraining. We found that subjects who studied the new words using the progressively less constraining sentences learned the new words better when measured on a host of definitional and comprehension measures. One difference between subjects in the high constraint and decreasing constraint group was that subjects in the decreasing constraint group encountered at least one moderately constraining sentence. If most of the advantage of the mixed constraint condition comes from requiring subjects to combine contextual information with knowledge gained in previous encounters with the word, it is possible that moderate constraint was the source of the effect of variation. The present experiment was designed to test this possibility. One group of subjects was taught 40 low frequency English words by presenting them with spaced encounters with the words in progressively less constraining contexts. Another group of subjects learned the 40 words

relying exclusively on encounters with the words in three different high constraint contexts. A third group of subjects learned the 40 words relying exclusively on encounters with the words in three different moderate constraint contexts. Beck et al's (1983) notion of a pedagogical context would suggest that students using the highly constraining contexts would learn the new words better. The argument developed above would suggest that students who encounter the words in moderate constraint or in progressively less constraining contexts would do better.

Method

Subjects

Subjects were 85 Northwestern undergraduates who received course credit for their participation.

Materials

Word lists: Two lists of twenty low frequency English words were compiled. These words were rated as unfamiliar by 5 independent subjects. The words were divided into two lists of 20 words so that one list could serve as an interference list for the other list.

Study sentences: For each word five highly constraining sentences were generated, five moderately constraining sentences were generated, and one minimally constraining sentence was generated. An example of the sentences used to teach the word *insalubrious* were:

High: John ate a healthful diet, but Karl's was insalubrious.
Moderate: Researchers believe that there is nothing insalubrious about an occasional glass of wine.
Low: It is not clear from what I've read whether it is insalubrious or not.

The average length in words of sentences in each condition was equated (High: 13.30 (3.68); Moderate: 13.35, (3.36); Low: 13.58, (3.32); $F<1$). Five subjects were presented with these sentence lists and were asked to rank order them based on how likely they would be to use the sentences if they wanted someone to guess the meaning of the target word. Based on these subjects' rankings we selected three different high constraint, three different moderate constraint and one low constraint sentences as study sentences for each word. Subjects in the high constraint condition were presented with three different high constraint sentences. Subjects in the moderate constraint condition were presented with three different moderate constraint sentences. Subjects in the mixed constraint condition were presented with a randomly chosen high constraint sentence, a randomly chosen moderate constraint sentence and a low constraint sentence.

Dependent measures: Subjects were tested at immediate and one week retention intervals.

i) Sentence verification: Subjects read 80 sentences, half of which made sense given the meaning of the target word and half of the sentences did not make sense given the meaning of the target word. The 80 sentences presented at the immediate test were different from the 80 sentences presented at one week delay. The sentences were presented in a different randomized order for each subject. Subjects were instructed to press the "z" key on their keyboard for sentences that made sense and the "/" key for sentences that didn't make sense. Subjects were instructed to respond as quickly as they could while still being accurate.

ii) Definition test: The definition test asked subjects to provide definitions for the 40 words studied. The immediate and delayed tests differed only in that the words were presented in different randomized orders.

Procedure

Subjects were run individually or in small groups. Each student provided us with their score on the verbal portion of the Scholastic Aptitude Test which served as our measure of prior verbal ability. They were seated in front of a computer terminal and were told that they would be learning 40 vocabulary terms and that they would do so by learning the words from context. The student read a sentence and was asked to define the target word based on that sentence. After the student typed in his/her response, the correct definition was presented on the screen for 2 seconds and the subject preceded to the next sentence. Immediately after the students were done with the study phase, they took the tests described above. Subjects returned one week later and retook the alternate versions of the tests.

Results

The results for the immediate accuracy data are shown in figures 1 and 2. As can be seen, they support the claim that medium levels of constraint are superior to high levels of constraint, and this appears to be especially true for high ability students. Any advantage of using progressively decreasing levels of constraint appears to follow from the inclusion of medium constraint sentences in the study sentences. We now describe the results according to the kind of test used.

Immediate testing

Sentence Verification: Accuracy in sentence verification reflects the degree of usable knowledge the subjects had acquired from the study sessions. For YES sentences the only significant predictor was verbal SAT, $F(1,79)=9.775$, $p<.002$. For NO sentences, level of constraint, $F(2,79)=3.425$, $p<.04$, and the interaction between self-rated verbal SAT and level of constraint, $F(2,79)=3.628$, $p<.04$, were both significant predictors of accuracy. Analysis of the response time data rules out the possibility that these results are an artifact of a speed/accuracy trade off because there was no effect of condition on response time, $F<1$. Planned comparisons

showed that students who studied the medium constraint sentences, $F(1,54)=6.57$, $p<.02$, and mixed constraint sentences, $F(1,51)=4.11$, $p<.05$, were better at rejecting inappropriate usage than students who studied only high constraint sentences. The advantage of the medium and mixed constraint conditions was greater for higher ability students than it was for lower ability students as demonstrated by the significant interaction between level of constraint and ability for both the high/medium, $F(1,54)=7.04$, $p<.01$, and high/mixed, $F(1,51)=4.11$, $p<.05$, comparisons. There were no significant differences between the medium and mixed constraint conditions. Overall, these results indicate that subjects who studied using either the medium or mixed constraint study sentences were better able to reject inappropriate word usage than were subjects who received only high constraint sentences.

Figure 2: Accuracy on immediate definition test.

Delayed testing

Students returned to be tested again after a one week interval. Again, the results showed an advantage of medium and mixed constraint conditions over the high constraint condition. As before, we present our results by type of test.

Sentence verification: Planned comparisons again revealed a significant advantage of the medium constraint over the high constraint condition on correct NO responses, $F(1,49)=5.51$, $p<.03$ and a significant interaction between constraint level and ability, $F(1,49)=5.36$, $p<.03$. There was also a marginally significant advantage of medium over mixed constraint, $F(1,50)=3.18$, $p<.09$, for correct NO responses on the delayed sentence verification test. The interaction between ability level and level of constraint was also marginally significant, $F(1,50)=3.55$, $p<.07$ and inspection reveals that this is due to poorer performance of low ability students in the mixed constraint condition. There was no difference between subjects in the high constraint and mixed constraint conditions. Thus, while medium levels of constraint improve the ability of students to correctly reject incorrect word usages even at a weeks delay, mixed levels of constraint appear to have a deleterious effect on the ability to reject incorrect usage at a weeks delay.
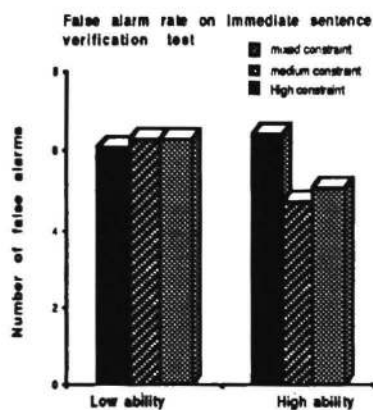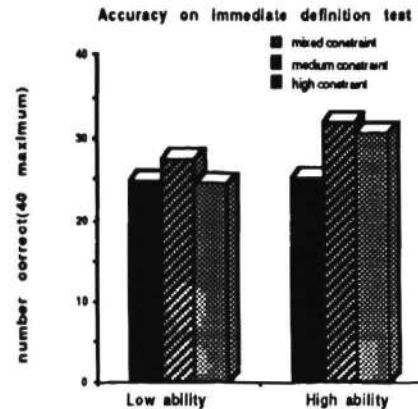

Figure 1: False alarm on immediate sentence verification test

Definitions: This measure indicates how accurately students can provide a definition for words they have learned. Accuracy on the definition test was significantly predicted by level of constraint, $F(2,79)=3.364$, $p<.04$, by verbal SAT, $F(1,79)=6.338$, $p<.02$, and by the interaction between verbal SAT and level of constraint, $F(2,79)=3.855$, $p<.03$. As with the sentence verification test, planned comparisons indicated that subjects in the medium, $F(1,54)=4.505$, $p<.04$, and mixed constraint, $F(1,51)=4.959$, $p<.03$, conditions outperformed those in the high constraint condition and the effect of level of constraint was amplified for high ability level students for both the high/medium, $F(1,54)=5.67$, $p<.02$ and high/mixed, $F(1,51)=5.50$, $p<.03$, comparisons. There were no significant differences between the medium and mixed conditions. So, as with verbal comprehension as measured by sentence verification accuracy, students who learned new words with medium constraint or mixed constraint example sentences were better able to provide definitions than students who learned new words using high constraint sentences.

Definitions: Planned comparisons indicated the same basic pattern of results on the delayed test as on the immediate test. Students who studied using medium constraint sentences outperformed students using high constraint sentences, $F(1,49)=7.22$, $p<.01$, and this advantage was greater for high ability students, $F(1,49)=8.23$, $p<.01$. There was also a marginally significant advantage for the mixed constraint condition over the high constraint condition, $F(1,49)=3.23$, $p<.08$, and a marginally significant interaction between level of constraint and ability, $F(1,49)=3.868$, $p<.06$.

Discussion

Beck et. al.(1983) have argued that students should be taught new words using highly constraining contexts. We suggested above that students would do better when they learn words in contexts which require them to integrate information from the current context with information from previous encounters with the word. To test this hypothesis, we taught college students 40 low frequency English words using either three encounters with the words in high constraint contexts, three encounters with words in moderate constraint contexts or using three encounters with the words in progressively less constraining contexts. We have shown that students who study new words in moderate constraint contexts develop a better ability to make use of these new words as measured by both definitional and sentence verification tasks than do students who study words in highly constraining contexts.

The results of this experiment are consistent with the view that teaching new words with moderate constraint contexts is superior to teaching new words in high constraint contexts. Moderate constraint contexts are superior because students are required to combine top down contextual information with bottom up word specific perceptual information as they do during normal reading. Many models of word recognition and meaning activation (e.g., Adams, 1990) hold that word meaning is activated in precisely this way. Our results suggest that learning words in a fashion that uses both sources of information facilitates subsequent use of the acquired meaning. The word knowledge acquired when both sources of information are attended to is more accessible when there is a need to combine contextual information with word specific information, namely, most reading or listening situations.

Our results also suggest that the optimal level of constraint to use in example sentences is contingent upon the ability level of the students. High ability level students appear to show the largest gains from encountering words in moderately constraining contexts. We can rule out prior knowledge of the particular word's meaning because performance prior knowledge would also have elevated scores in the high constraint condition. Van Daalen-Kapteihns and Elshout-Mohr(1981) suggested that lower ability level students developed holistic models of word meaning based on minimal contextual encounters. Since these models are holistic they are difficult to modify in light of new information. Since learning word meaning from less constraining contexts requires the acceptance of a provisional model of word meaning followed by updating that model as new information becomes available, lower ability level students may be less able to take advantage of moderate constraint contexts since their models are less modifiable. Nevertheless, it should be noted that at all ability levels, moderately constraining sentences are at least as efficacious as highly constraining sentences.

The present study also suggests that our finding of advantages for progressively less constraining contexts over high constraint contexts may have been due primarily to the presence of a moderately constraining sentence in the progressively less constraining study list. This is especially clear in the delayed accuracy results where subjects in the moderate constraint condition continue to outperform subjects in the high constraint condition, but where subjects in the mixed constraint condition no longer significantly outperform those in high constraint condition. It is possible that the decreasing constraint sentences we used made jumps which were too large to be effective. The final example sentence, for instance, may have been too sparse for a third encounter. On the other hand, one potential benefit of decreasing constraint may be in situations where the definition is not provided. We are currently comparing these methods under conditions in which definitions are not provided as part of the study procedure and this may show an advantage for progressively decreasing constraint.

Teaching vocabulary is an important part of learning a language for the first time or as a second language. Word meaning knowledge is an important component of general comprehension. Indeed vocabulary knowledge is highly correlated with general comprehension abilities(Anderson and Freebody, 1981). Further, it is reasonable to assume that incidental acquisition of vocabulary is dependent upon current word knowledge, such that learning new words should increase the ability of students to learn other new words in the future. Beck et. al.(1982) have found such incidental gains in students who have participated in their vocabulary building curriculum.

Beck et. al.(1983) advised educators that "not all contexts are created equal." They argued that the contexts used to teach new words should be contexts which highly constrain the meanings of the new words. While this may be good advice for the elementary school teachers they were addressing, our research suggests that less constraining contexts may be superior for older students. Students need to learn to quickly access and use word meaning when reading or listening to normal paced speech. To do that, they need to strengthen the same kinds of general comprehension processes during study that they will use during reading. Our research suggests that there is some advantage to be gained from providing study sentences which are either moderately constraining or which vary in constraint. Using moderately constraining or progressively less constraining contexts results in a better integration of multiple encounters with novel words and consequently, more usable word knowledge.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge MA: MIT Press.

Anderson, R. and Freebody, P.(1981). Vocabulary knowledge. In J. Guthrie(Ed). *Comprehension and Teaching: Research Reviews.* 77-117, Newark DE: W.W. Norton.

Beck,I.L., McKeown, M.G., and McCaslin, E.S. (1983). Vocabulary development: All contexts are not created equal. Elementary School Journal, 83, 177-181.

Beck, I.L., Perfetti, C.A. and McKeown, M.G.(1982) Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology,* 74, 506-521.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior,* 22, 75-87.

Hall ,J.W. (1988) On the utility of the keyword mnemonic for vocabulary learning. *Journal of Experimental Psychology,* 80, 554-562.

Kolich, E.M.(1991). Effects of computer-assisted vocabulary training on word knowledge. *Journal of Educational Research,* 84, 177-182.

Nagy W. and Herman, P. (1987) Breadth and depth of vocabulary knowledge. In M.G. McKeown and M.E. Curtis(Eds).*The Nature of Vocabulary Acquisition.* Lawrence Erlbaum, Hillsdale NJ pp. 19-35.

Schank, R.C.(1982). *Reading and Understanding: Teaching from the Perspective of Artificial Intelligence.* Lawrence Erlbaum, Hillsdale, NJ.

Schuberth, R. E., & Eimas, P. D.(1977). Effects of context on the classification of words and nonwords.*Journal of Experimental Psychology:Human Perception and Performance,* 3, 27-36.

Stahl, S.A.(1991). Beyond the instrumentatlist hypothesis: Some relationships between word meanings and comprehension. In P.J. Schwanenflugel (Ed). *The Psychology of Word Meanings.* Lawrence Erlbaum, Hillsdale, NJ.

Stahl, S.A. and Fairbanks, M.(1986)The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research,* 56, 72-110.

Taylor, I. and Taylor, M. M., (1983) *The Psychology of Reading.* Academic Press, NY.

Van Daalen-Kapteijns and Elshout-Mohr(1981). The acquisition of word meanings as a cognitive learning process. *Journal of Verbal Learning and Verbal Behavior,* 20, 386-399.