

## Context effects in sentence comprehension: A study of the subjective lexicon\*

CAROL CONRAD†

*University of Oregon, Eugene, Oregon 97403*

These studies explore the role of context in determining what information about the meanings of words is activated in memory at the time a word is encountered in a sentence. Using a color-naming paradigm, it was shown that both meanings of a word that has two distinct meanings are activated in memory at the time the word is heard in a sentence. This activation occurs even when there is sufficient contextual information to indicate which meaning was intended by the speaker. These results support the hypothesis that there exists in memory an isolable subjective lexicon. They suggest that context which is effective in disambiguating lexical ambiguities in the language has its effect only at a relatively late stage in the cognitive processing involved in language comprehension.

In his 1970 presidential address to the American Psychological Association, George Miller posed a problem to psychologists interested in studying human language. The problem was basically this: Can we show that there exists a subjective lexicon which is an isolable subsystem in memory? The subjective lexicon, in the sense that Miller speaks of, refers to our intrinsic knowledge of the meanings of words. It is a set of concepts which have been acquired and refined over a relatively long period of time, a set of dictionary entries which define words, specify their selection restrictions, and perhaps provide information about their possible syntactic roles.

The notion of an isolable subsystem is a complex one: Posner (1973) has presented evidence which suggests that a stimulus may have several different codes and that different classes of mental operations are performed upon these codes. Thus, a written word may have both a physical code and a name code. It also has a meaning, which may be represented in a lexical system in memory. To be isolable, each system or code must have its own properties (see Katz & Fodor, 1963, for an early theoretical attempt at defining the properties of a lexical system) and each must make a separable contribution to a complex task such as language comprehension. One way of establishing experimentally that a coding system is isolable is to show that manipulating the availability of one coding system does not affect the other codes for the same stimulus. An example might be the visual and

auditory systems; brain damage to one of these systems may not affect the other.

To answer Miller's question, it seems necessary to demonstrate at least two things. First, to show that a lexicon exists, it must be demonstrated that, at some level of sentence comprehension, Ss activate in memory the meanings of the individual words in a sentence. Second, to show that the lexicon is an isolable system, it must be demonstrated that the lexical look-up process is not influenced by either the linguistic or nonlinguistic context provided by the sentence or the environment or by the listener's more general knowledge of the world.

The experiments reported here represent an attempt to demonstrate that there does exist a lexical look-up process which is not influenced by linguistic context. The methodology used is a variation on a technique used by Warren (1972). Warren aurally presented Ss with three words which all had the same category name (e.g., robin, canary, sparrow); this was followed by the visual presentation of a single word which was in the list (e.g., robin), the category common to all of the words in the list (e.g., bird), or an unrelated word. The words were printed in colored ink, and Ss were required to name the color of the word as quickly as possible after the word appeared.

Color-naming RTs showed a significant interference effect for the category names as well as for the words in the list. Warren concluded from this that the category names for words are activated in memory along with the representation of the word itself at the time that the words are heard.

Warren's technique provides one way of studying the activation of meanings for the individual words in a sentence. By using sentences instead of word lists, it is possible to manipulate the linguistic context for a word so that the activation of the meanings for that word can be studied under varying contextual conditions. The experiments reported here use contextual constraints on ambiguous words in order to test the notion that there exists an isolable lexicon which plays a role in sentence comprehension.

\*This research was based on a dissertation submitted in partial fulfillment of the requirements for the PhD degree at the University of Oregon and was conducted while the author held an NSF predoctoral traineeship. Funds for Ss and equipment were provided by NSF Grant GB 21020 and by USOE Grant OEG-0-72-0717 from the U.S. Department of Health, Education and Welfare. The author thanks Michael Posner, Ray Hyman, Gerald Reicher, and Merrill Garrett for their helpful comments and criticism on this research.

†Requests for reprints should be sent to Carol Conrad, Department of Psychology, New School for Social Research, 66 West 12th Street, New York, New York 10011.

Recently, a number of people have tried to take advantage of ambiguities in the language to study sentence comprehension. Methodologies have included sentence completion tasks, probe techniques to measure interference, picture-sentence matching tasks which provide reaction-time measures, and dichotic listening tasks. They have been used to study the role of context in disambiguating sentences and to determine how the processing demands of ambiguous sentences vary as a function of context. Several of these studies have shown that the probability of a particular interpretation of an ambiguous sentence can be influenced by the form or meaning of the sentences or words which precede it (Marshall, 1965; Carey, Mehler, & Bever, 1970; Lackner & Garrett, 1973; MacKay, 1973). These studies indicate that the interpretation of an ambiguous sentence may be influenced by linguistic context. In addition, a recent study by Bransford and Johnson (1973) suggests that even the comprehension of unambiguous sentences may depend on having available prior linguistic context. Thus, it is clear that context plays an important role in the comprehension of language and that any theory of comprehension must be able to specify how the processing of contextual information contributes to our understanding of the language we hear or read. The present studies evaluate one hypothesis concerning this processing: that prior linguistic context influences sentence comprehension by suppressing the activation of inappropriate meanings for words in a sentence.

The research on the effects of context on ambiguity arose from the question of whether or not people process both meanings of an ambiguous sentence. This research has given rise to several theories of how ambiguities are processed. The first of these, the garden path theory, was initially suggested by Lashley (1951), who proposed that people process only one meaning or structure of a sentence at a time and that if the original interpretation later becomes inappropriate, we then process the other meaning. Lashley offered the following sentence as an example: "Rapid / rajtin / with his uninjured hand saved from loss the contents of the capsized canoe." Almost everyone tends to interpret / rajtin / as "writing"; however, the context which follows the word forces us to reinterpret it as "righting." The theory fits nicely with the observation that although ambiguities are quite frequent, at least in the English language, we seldom are aware of them.

An alternative to the garden path theory, the exhaustive computation hypothesis, suggests that all possible interpretations of an ambiguous sentence are computed at an early stage of processing and, at some later point, a single reading is selected for attention. A more explicit version of this hypothesis is MacKay's perceptual suppression theory, which proposes that all of the meanings of an ambiguous sentence are processed in parallel at an unconscious level. In order to perceive one of the meanings, the other meaning must be suppressed, and the time for suppression depends on the

a priori likelihood of the suppressed meaning, independent of context. Context does have an effect, however, by strengthening the activation of the appropriate meaning.

At the present time, the data that might allow us to choose between these theories are themselves ambiguous. Several experiments argue strongly for the garden path theory (Foss, Bever, & Silver, 1968; Carey, Mehler, & Bever, 1970), while others suggest that at some point, both meanings of an ambiguity are processed (MacKay, 1966; Fodor, Garrett, & Bever, 1968; Lackner & Garrett, 1973) and that processing demands differ for ambiguous and unambiguous sentences (Foss, 1970; Garrett, 1970). Garrett (1970), in reviewing these studies, has suggested that whether or not effects of ambiguity are found depends on whether testing occurs during the processing involved in comprehending the sentence or after the sentence has already been comprehended. The latter measures tend to support the garden path theory, while the former measures indicate that ambiguous sentences are more difficult to process than are unambiguous ones.

The study of lexical ambiguities may provide information about the subjective lexicon. Both ambiguous and unambiguous sentences can contain a lexical item with two or more different meanings. The two types of sentences differ in that the unambiguous sentences contain contextual information which tells the listener which meaning of the word is appropriate, while the ambiguous sentences do not contain this type of contextual information. If there are differences in interference on the color-naming task for ambiguous and unambiguous sentences which may be attributed to these contextual differences, then we must conclude that the subjective lexicon is not context-independent. A failure to find these differences, however, would suggest that, at some stage, the lexicon is not influenced by contextual information. In addition, a study of this type may provide information to help us choose between the several existing theories of ambiguity.

## EXPERIMENT I

This experiment was designed to test the effects of context on the activation of the category names for ambiguous words which were part of a sentence. Context which served to disambiguate the word was incorporated into the sentence and occurred either prior to or subsequent to the ambiguous word. The paradigm used was a modification of the color-naming task used by Warren (1972).

The goal of the experiment was to determine (1) whether context places constraints on the word meanings which are activated in memory or (2) whether more than one meaning is initially activated and context has its disambiguating effect subsequent to a lexical look-up process. Finding 1 would suggest that there does not exist a lexical look-up process which is independent

**Table 1**  
**Examples of the Sentences and Colored Words Used in Each Condition in Experiment I**

Display Type	Sentence	Display	
	<b>Context After Word</b>		
Ambiguous Word	The nickel is not enough to buy the candy bar.	Nickel	(Red)
Appropriate Category	That nickel is from her coin collection.	Money	(Red)
Inappropriate Category	The nickel was lost from her purse.	Metal	(Red)
	<b>Context Before Word</b>		
Ambiguous Word	The girl found a nickel.	Nickel	(Red)
Appropriate Category	The toy costs a nickel.	Money	(Red)
Inappropriate Category	He bought a comic book for a nickel.	Metal	(Red)
	<b>Controls</b>		
Ambiguous Word	He brought a bottle of wine to dinner.	Nickel	(Red)
Appropriate Category	The craftsman made the chair.	Money	(Red)
Inappropriate Category	The boy next door mows our lawn.	Metal	(Red)

of contextual influence, whereas Finding 2 would provide some support for the notion of an isolable lexical system in memory.

The position of the context relative to the ambiguous word was varied, because if context does influence the lexical look-up process, it would do so only when it is already available at the time the word is heard. Thus, given the assumptions of the color-naming task, context could only have an effect when it preceded the word. In addition, if context has an effect on the lexical look-up process, the manipulation of the position of the context would provide a test of the assumption that the color-naming task reflects activation at the time that a word is heard. A result indicating that context subsequent to the word can influence lexical look-up would strongly suggest that activation does not take place at the time that the word is heard. Category names were used as representations of the meanings of an ambiguous word because previous work (Conrad, 1971) had indicated that the category name for a word is an integral part of the reported meaning of the word.

### Method

#### *Subjects*

Thirteen male University of Oregon students served as Ss in the experiment. They were obtained from the S pool for the Center for Perceptual and Cognitive Studies and were paid \$3 for their participation. All Ss were native English speakers, had good vision and hearing, and, when questioned, all reported that they were not color-blind.

#### *Procedure*

For each trial in the experiment, Ss listened to a single tape-recorded sentence under instructions to try to remember the sentence. Immediately following the sentence, a word, printed in colored ink, appeared on a rear projection screen in front of the Ss. The words were displayed using a Kodak Carousel slide projector equipped with a Lafayette tachistoscopic shutter. They were printed in capital letters and were centered on a pale gray background. Four colors of ink—red, blue, green, and purple—were used. Ss were instructed to name aloud the color of the word as quickly as possible and then to repeat the sentence verbatim. The word remained on the screen until the S had named the color aloud. Naming the color

activated a voice-operated relay which simultaneously terminated the display, stopped a Hunter millisecond timer which had been started at the onset of the display, and positioned the slide tray for the next display. Both color-naming reaction times and Ss' recall were recorded by E. Ss received 25 practice trials followed by 96 experimental trials on each of 2 consecutive days. Each session lasted approximately 40 min.

#### *Words*

Eight ambiguous words having two or more different meanings, e.g., pot, were used. They were selected from a set of 35 words which had been presented to 25 Ss in a pretest. These Ss were asked to name two categories for each word and to indicate which category they had thought of first. The goal was to find a set of words which would fit into two different categories, with each category corresponding to one of the meanings of the word. In addition, the categories had to be given the same name by a majority of the Ss to insure that the structures selected for testing would be common to as many Ss as possible. The eight words selected met the following criteria: at least 40% of the Ss in the pretest mentioned the same category name for a word and at least 40% indicated that they had thought of that category first. In addition, the results of a pilot study, similar to Experiment II, indicated that all of the category names for the eight words showed some interference with the color-naming task.

#### *Sentences*

The eight ambiguous words were incorporated into two experimental sets of 48 unambiguous sentences. In the first set, the context after set, the word was always the second word in the sentence and was made unambiguous by the words which followed it. In the second set, the context before set, the ambiguous word was always the last word in the sentence and was disambiguated by the words which preceded it. Each of the eight words appeared in six sentences in each set so that both meanings of the word were represented three times in each set. In addition, 48 unambiguous sentences that did not contain any of the ambiguous words were used as a control set.

#### *Conditions*

Each of the experimental sentences was followed by a visual display of the ambiguous word used in the sentence (ambiguous word), or by the category name for that word which was appropriate to the meaning expressed by the sentence (appropriate category), or by the category name for the word which was inappropriate to the meaning expressed by the sentence (inappropriate category). Both the context before and context after sentences occurred in each of the three

**Table 2**  
Mean RTs and Their Differences for Correct Color Responses  
in Each Condition in Experiment I

Display Type	Trial Type		
	Experi- mental	Control	Diffe- rence
Context Before Word			
Ambiguous Word	1032	932	100
Appropriate Category	990	884	106
Inappropriate Category	936	872	64
Context After Word			
Ambiguous Word	977	958	19
Appropriate Category	917	915	2
Inappropriate Category	913	878	35

*Note*—RTs are expressed in milliseconds. Difference equals experimental mean minus control mean.

experimental conditions. In the control conditions, the same words, printed in the same color of ink, were paired with the control sentences. Table 1 shows examples of the sentences and colored words used in each condition.

On each day, Ss received half of the experimental trials and their appropriate controls. The sentences were divided in half so that on each day, Ss received six sentences containing each ambiguous word, three for each meaning, with one sentence in each condition for both the context before and context after sets.

### Results

Mean RTs for correct color responses for each S on each day were calculated. The means of these means, collapsed over days and Ss, are shown in Table 2.

A five-way analysis of variance was done on the means. The five factors were Ss, number of days of practice (days), the context position with respect to the ambiguous word (context), the type of visual display (display type), and the type of trial, i.e., experimental vs control (trial type).

As might be expected, there was a significant main effect of days [ $F(1,12) = 5.7, p < .05$ ]. This reflects a decrease in overall RTs from Day 1 to Day 2. In addition, days interacted with both trial type [ $F(1,12) = 16.4, p < .01$ ] and context [ $F(1,12) = 7.9, p < .01$ ]. There was a larger decrease in mean RTs for experimental trials from Day 1 to Day 2 (106 msec) than for control trials (39 msec) and a larger decrease in mean RTs, over days, for trials in which the context preceded the ambiguous word in the sentence (119 msec) than for trials where the context came after the ambiguous word in the sentence (63 msec).

Because of these interactions, the data were reanalyzed separately for Days 1 and 2, using a four-way analysis of variance. The means for Days 1 and 2 are shown in Table 3.

The analysis of the Day 1 data indicates a significant main effect of trial type [ $F(1,12) = 28.0, p < .01$ ]. A look at the data shows that for five of the six comparisons, RTs for the experimental trials were longer than those for the appropriate controls. Only the

condition where the visual display was a category which corresponded to the meaning of the ambiguous word in the sentence, and the context came after that word in the sentence, does not show this interference effect. In addition, there was a significant Trial Type by Context interaction [ $F(1,12) = 23.6, p < .01$ ]. This reflects the fact that there were larger differences between the mean RTs for experimental and control trials when the context preceded the ambiguous word in the sentence than when the context followed the word. Thus, Ss demonstrated larger interference effects when the context came before the ambiguous word.

An a priori test of paired comparisons, using the method suggested by Winer (1962), was done on the means for each pair of experimental and control trials. For the context after sentences, only RTs for the ambiguous word condition were significantly larger than their control group ( $p < .01$ ). For the context before sentences, however, mean RTs for all of the experimental conditions were larger than their appropriate controls ( $p < .025$ ).

Overall mean RTs were faster following sentences in which the context came after the ambiguous word than following sentences in which the context preceded the ambiguous word, producing a significant main effect of context [ $F(1,12) = 10.1, p < .01$ ]. Thus, in the context after condition, Ss were both faster at color naming and showed less interference from the meaning of the ambiguous word.

Finally, there was a significant main effect of display type [ $F(2,24) = 26.0, p < .01$ ]. Mean RTs were fastest for the inappropriate category displays and slowest for the ambiguous word displays. An a priori test of paired comparisons (Winer, 1962), done on the experimental mean RTs for the context before sentences, indicated that there was no significant difference between the

**Table 3**  
Mean RTs and Their Differences for Each Condition  
on Each Day in Experiment I

		Trial Type		
Day	Display Type	Experi- mental	Control	Diffe- rence
Context Before Word				
1	Ambiguous Word	1097	999	98
1	Appropriate Category	1069	938	131
1	Inappropriate Category	982	920	62
2	Ambiguous Word	967	870	97
2	Appropriate Category	910	829	81
2	Inappropriate Category	890	823	67
Context After Word				
1	Ambiguous Word	1052	980	72
1	Appropriate Category	926	935	-9
1	Inappropriate Category	956	897	59
2	Ambiguous Word	902	936	-34
2	Appropriate Category	907	895	12
2	Inappropriate Category	870	859	11

*Note*—RTs are expressed in milliseconds. Difference equals experimental mean minus control mean.

means for the ambiguous word and appropriate category displays [ $F(1,96) = 1.15$ ] but that there was a significant difference between the appropriate category and inappropriate category means [ $F(1,96) = 11.11$ ,  $p < .01$ ]. However, a test of the difference scores for these means and their controls indicated that there was no significant difference between the scores for appropriate and inappropriate categories, suggesting that the control means were also different for the two conditions. Thus, the evidence with respect to whether or not appropriate category names produced more interference than inappropriate category names is inconclusive. The same *a priori* test, done on the experimental means for the context after conditions, indicated a significant difference between the means for the ambiguous word and appropriate category displays [ $F(1,96) = 27.1$ ,  $p < .01$ ] but no significant difference between the means for the appropriate category and inappropriate category displays [ $F(1,96) = 1.3$ ]. Thus, for the sentences in which the context followed the ambiguous word, Ss were slower at color naming when the visual display was a word used in the sentence than when the display was either an appropriate or inappropriate category for that word.

An analysis of variance done on the Day 2 data showed a significant main effect of trial type [ $F(1,12) = 15.0$ ,  $p < .01$ ] and of display type [ $F(2,24) = 5.0$ ,  $p < .05$ ]. Again, for five of the six comparisons, RTs were larger for experimental trials than for control trials, with only the ambiguous word display following sentences in which the context came after the ambiguous word, showing no interference effects. Means for the ambiguous word displays were the slowest and those for the inappropriate category displays were fastest. In addition, there was a significant Trial Type by Context interaction [ $F(1,12) = 5.4$ ,  $p < .05$ ], reflecting the larger differences in RTs between the experimental trials and their appropriate controls for the sentences in which the context preceded the ambiguous word.

An *a priori* test of paired comparisons (Winer, 1962), done on the mean RTs for each experimental condition and its control, indicated that, for the context after sentences, none of the mean RTs for the experimental conditions was significantly different from the mean RTs for the control groups. For the context before sentences, both the ambiguous word and appropriate category conditions had mean RTs significantly larger than those for their control groups.

The overall error rate for color naming in the experiment was 1.24%. The types of errors and their frequency of occurrence were as follows: (1) Ss named the wrong color (.16%), (2) Ss named the word which was displayed instead of the color (.12%), (3) Ss recalled the sentence without first naming the color (.16%), and (4) equipment or E errors (.80%). Verbatim sentence recall was quite good, with only a .16% error rate.

### Discussion

One rather unexpected finding was that there was a

marked decrease in interference effects on the second day of practice. Unlike Warren's findings for category names, only the mean RT for appropriate categories in the context before condition showed an interference effect on the second day. This may be due to the decrease in RTs for Day 2, since we know the Ss having very fast RTs often do not show even the usual Stroop effect.

The lack of interference effects for the context after conditions, even on Day 1, may be attributed to the longer delay between aural presentation of the ambiguous word and the visual display. Since the ambiguous word appeared at the beginning of the sentence in these conditions and the visual display was not presented until Ss had heard the entire sentence, there was a variable delay interval during which Ss were listening to and processing the remainder of the sentence. Thus, the lack of interference may be due either to loss of activation over the interval, attributable to a fast decay rate, or to interference from the processing which occurred during the interval. A third explanation might be that there was no activation of the meanings of ambiguous words in the context after condition. However, this seems unlikely, since there is evidence for activation in the context before condition, the condition where it might least be expected. In addition, an experiment by Foss and Jenkins (1972) also suggests that activation does take place. Foss found that there was significant interference in RTs to a phoneme-monitoring task immediately following the presentation of an ambiguous word and that this interference occurs both when the word has been disambiguated by prior context and when it has not yet been disambiguated. Thus, it seems probable that word meanings were activated in the context after condition. However, it is not clear whether the activation was terminated because of a fast decay rate or as a result of further processing.

The most interesting results are those for the context before conditions on Day 1. To insure that the results were not due to the repetitions of the ambiguous words in the sentences, the data for only the first occurrence of each ambiguous word in a sentence and its appropriate control were analyzed. It was found that, even on its first occurrence, an ambiguous word activated both appropriate and inappropriate category names. The fact that all the appropriate and inappropriate categories show interference effects indicates that both meanings of the word are activated in memory, even when the preceding context indicates which meaning is intended. Context may be having some effect on this activation, since the experimental means for appropriate categories are significantly longer than the means for the inappropriate ones. However, this seems unlikely, since a more appropriate test, that of the difference scores, indicates that there is no significant difference in the amount of interference for the two category conditions. Thus, the results do not unequivocally support or refute the notion of an isolable subjective lexicon. They do,

however, support the existence of a subjective lexicon by showing that, during sentence comprehension, Ss do activate in memory the meanings of individual words in the sentence.

## EXPERIMENT II

The previous experiment suggested that linguistic context might have some minimal effect on the strength of activation of category names for ambiguous words which occur in a sentence. It would seem desirable to replicate that finding and to compare the results with those for ambiguous sentences containing the same ambiguous words. This would provide a condition in which there is no context to indicate which meaning of the ambiguous word is intended. If context is effective in reducing the strength of activation of inappropriate category information, then one might expect a larger interference effect of both categories for ambiguous sentences than for the inappropriate category for unambiguous sentences. In addition, the data for ambiguous sentences may provide further evidence with which to choose among the current theories of ambiguity. The garden path theory would predict that, for the ambiguous sentences, only a single meaning of an ambiguous word is activated in memory. Both the exhaustive computation hypothesis and the perceptual suppression hypothesis, on the other hand, predict that all meanings of the word will be activated. The latter also predicts an inequality of strength of activation for the various meanings of a word.

The present experiment, then, attempts to replicate the findings of Experiment I for sentences containing ambiguous lexical items which are disambiguated by prior context, to extend these findings to sentences which contain no disambiguating context, and to decide between the various theories of ambiguity.

## Method

### Subjects

Fourteen University of Oregon students served as Ss in the experiment. Ten were obtained from the S pool and were paid \$3 for their participation. The remaining four were volunteers who received extra credit for a psychology class. All had good vision and hearing; none reported being color-blind or had participated in Experiment I or the pretest for Experiment II. Ten of the Ss were male and four were female.

### Procedure

The procedure and equipment were the same as in Experiment I. On each trial, Ss listened to a single sentence which was followed immediately by the visual presentation of a word printed in colored ink. Ss named the color of the word as quickly as possible and then recalled the sentence aloud. Color-naming RTs and recall errors were recorded by E. Four colors—red, blue, green, and purple—were used. Ss received 25 practice trials, followed by 96 experimental trials on each of 2 consecutive days. Each session lasted approximately 30 min.

### Sentences

The eight ambiguous words used in Experiment I were included in each of two types of sentences, and in all cases, the ambiguous word was the last word in the sentence. Forty-eight of the sentences, 6 for each word, contained a lexical ambiguity (e.g., "She has lost her bat."). A pretest had indicated that all of these sentences were behaviorally as well as logically ambiguous. That is, both meanings of each sentence were plausible enough so that Ss could find both meanings when asked to do so, and not all Ss gave it the same meaning initially. The remaining 48 were unambiguous (e.g., "She fed the bat."). Within the unambiguous set, there were three sentences for each meaning of each of the ambiguous words. In addition, there was a control set of 48 unambiguous sentences which did not contain any of the ambiguous words.

### Conditions

Each of the sentences was followed by a visual display of the ambiguous word used in the sentence (ambiguous word), by the category name for that word which was appropriate to the meaning expressed by the sentence (appropriate category), or by the category name for the ambiguous word which was inappropriate to the meaning expressed by the sentence (inappropriate category). In reality, the appropriate-inappropriate category distinction could not be made for the ambiguous sentences, since it was impossible with this experimental design to know what interpretation each S gave to the sentences. For these sentences, then, there was no difference between the appropriate category and inappropriate category conditions. However, the distinction was used to maintain a balanced design for purposes of data analysis.

In the control conditions (control), each of the words used in each experimental condition was paired with a control sentence. Experimental and control words were printed in the same color of ink.

Table 4 shows examples of the experimental sentences and colored words for each condition.

On each day, Ss received half of the experimental trials and their appropriate controls. The sentences were divided so that each day, Ss heard three ambiguous and three unambiguous sentences containing each ambiguous word, with one sentence of each type in each condition.

## Results

Mean RTs for correct color responses for each S on each day were calculated. The means of these means, collapsed over days and Ss, are shown in Table 5.

A five-way analysis of variance was done on the means. The five factors were Ss, number of days of practice (days), type of sentence, i.e., ambiguous vs unambiguous (sentence type), the type of visual display (display type), and the type of trial, i.e., experimental vs control (trial type).

As expected, there was a significant main effect of days [ $F(1,13) = 25.0, p < .01$ ], reflecting the decrease in RTs from Day 1 to Day 2. Since there were no interactions with days as a factor, means for both days combined were calculated for each condition for each S, and these were analyzed using a four-way analysis of variance.

The analysis showed a significant main effect of both trial type [ $F(1,13) = 35.8, p < .01$ ] and display type [ $F(2,26) = 15.8, p < .01$ ], but no main effect of

**Table 4**  
**Examples of the Sentences and Colored Words Used in Each Condition in Experiment II**

Display Type	Sentences	Display	
Ambiguous Sentences			
Ambiguous Word	The cupboard contains the pot.	Pot	(Red)
	I traded my desk for the pot.	Pot	(Blue)
Category Word	John bought the pot.	Utensil	(Red)
	He finished the pot.	Marijuana	(Blue)
Unambiguous Sentences			
Ambiguous Word	Beans are cooked in the pot.	Pot	(Red)
	He was arrested for possessing pot.	Pot	(Blue)
Appropriate Category	We boiled water in the pot.	Utensil	(Red)
	It is illegal to smoke pot.	Marijuana	(Blue)
Inappropriate Category	We made tea in the pot	Marijuana	(Red)
	The field was full of pot.	Utensil	(Blue)
Control Sentences			
Ambiguous Word	The truck collided with the wall.	Pot	(Red)
	She made a pie for dinner.	Pot	(Blue)
Appropriate Category	The race track was thoroughly prepared.	Utensil	(Red)
	The man owns a sailboat.	Marijuana	(Blue)
Inappropriate Category	The television station covered the political campaign.	Marijuana	(Red)
	He bought a flowered shirt.	Utensil	(Blue)

sentence type ( $F = 1.0$ ). In addition, there were no significant interactions.

An a priori test of paired comparisons, using the method suggested by Winer (1962), was done on each of the experimental means and their appropriate control to determine whether or not there were significant interference effects for each experimental condition. All of the comparisons were significant (with  $p < .01$ ).

Although there was a significant main effect of display type, a comparison of the difference scores (experimental means minus control mean) for the ambiguous word and category word conditions for both the ambiguous and unambiguous sentences suggests that there are no differences in degree of interference for different types of displays. In addition, the lack of a main effect for sentence type indicates that there were no differences in amount of interference for ambiguous and unambiguous sentences, although, in all cases, the

difference scores are longer for ambiguous sentences than for unambiguous ones.

Since information about the bias of each interpretation of the ambiguous sentences was available from a pretest, it was possible to rescore the category responses and their controls for the ambiguous sentences separately for categories which were appropriate to that meaning of the sentence which had a bias of  $> 50\%$  (likely category) and for the categories which corresponded to the meaning which had a bias of  $\leq 50\%$  (unlikely category). The mean RT, over days, for the likely category was 1,014 msec, with a control of 939 msec, and the mean for the unlikely categories was 1,017 msec, with a control of 943 msec. Both experimental means were significantly larger than their controls ( $p < .001$ ) but were not different from one another.

The error rate for color naming was 1.02%; naming the wrong color constituted .18% of the errors; repeating the word name, .03%; color omissions, .18%; and equipment or E errors, .63%. Verbatim sentence recall was quite good, with an error rate of .18%.

**Table 5**  
**Mean RTs and Their Differences for Correct Color Responses in Each Condition in Experiment II**

Display Type	Trial Type		
	Experimental	Control	Difference
<b>Ambiguous Sentences</b>			
Ambiguous Word	1044	971	73
Appropriate Category	1015*	938*	77
Inappropriate Category			
<b>Unambiguous Sentences</b>			
Ambiguous Word	1059	1001	58
Appropriate Category	997	936	61
Inappropriate Category	993	943	50

*Note*—RTs are expressed in milliseconds. Difference equals experimental mean minus control mean.

\*Since there was no difference between appropriate and inappropriate category conditions for ambiguous sentences, these were combined into a single mean.

## Discussion

The results for unambiguous sentences in which the ambiguous word is preceded by disambiguating context again show significant interference effects for both appropriate and inappropriate categories, but no difference in the size of the interference effects for the two types of categories. In addition, there were no significant differences in the strength of activation of the category names for ambiguous and unambiguous sentences. These findings, then, provide further support for the idea that there may be a context-independent lexical look-up stage which occurs early in sentence processing and that context has its effect on sentence

comprehension at some later stage of processing.

The results also seem to provide further support for the exhaustive computation hypothesis, at least to the extent that it was intended to explain lexical ambiguities. This hypothesis, like the perceptual suppression theory, suggests that at some level of processing, all meanings of an ambiguous sentence are computed. Clearly, the results suggest that at least two interpretations of a lexically ambiguous sentence are computed at an early stage of processing.

The results do not seem to support MacKay's perceptual suppression theory. The theory proposes that in order to perceive one meaning of an ambiguity, the other meaning must be suppressed. This would seem to predict a difference in the strength of activation for the likely and unlikely categories in the present experiment. However, the results indicate that there is not a difference in the strength of activation for the two category types. Although the present study uses a limited number of words to test this hypothesis, the probabilities reflecting the likelihood of a meaning did range from .8 to .2, with a fairly equal distribution of scores throughout the range.

### CONCLUSIONS

The results of these experiments suggest that there does exist a subjective lexicon as Miller suggested and that linguistic context does not have an effect on the lexical look-up process. However, since there were some significant differences in the strength of activation of appropriate and inappropriate category names in Experiment I, it may be possible that small differences do exist. If so, perhaps they could be measured with a more sensitive experimental technique.

It is possible to argue that the color-naming paradigm used in these experiments does not measure the activation of meaning which occurs when a word is encountered in a sentence. Instead, one might argue that the category names are activated when they are presented as the visual display and only then are they associated with the ambiguous word which occurred in the sentence. However, there are several reasons for believing that a backwards association theory of this sort cannot account for the data. First, a theory of this type would seem to predict that at least some of the color-naming errors would be intrusions of the ambiguous word. The fact that there were no errors of this type in either of the present experiments or in Warren's (1972) study suggests that this is not a reasonable explanation of the effects found in these experiments. Second, Ss' reports following the experiments indicated that they were not aware of the relationship between the last word in the sentence and the word presented visually. This suggests that if Ss were associating the two words only at the time the second word was presented, this was not done consciously. Third, a recent experiment by Warren (1973)

provides a direct test of the backwards association theory by using words which are highly associated in only one direction in the color-naming paradigm. For example, we know that given the word "bitter" in a free association test, "sweet" is a highly probable first response. On the other hand, "bitter" is a very unlikely response to "sweet" (Palermo & Jenkins, 1964). Warren aurally presented Ss with words like "bitter" followed by a visual presentation of words like "sweet" and compared the amount of color-naming interference to a condition in which "sweet" preceded "bitter." He found significant interference for the condition where "bitter" preceded "sweet" but no interference when "sweet" preceded "bitter." These results do not support a backwards association theory; rather, they strongly argue that the visually presented word was activated in memory prior to its actual presentation. Finally, Cairns (1973), in an experiment where Ss were forced to reprocess phonological ambiguities (e.g., weight vs wait), has found that the time for reprocessing is on the order of 350 msec. This time to find the other meaning of an aurally presented word is considerably longer than the difference scores found in the present experiments. The Cairns experiment, then, suggests that if a backwards association explanation of the present data were correct, one might expect the differences in RT for experimental and control trials to be three times as large as they actually were. Thus, the best account of the color-naming interference effect found in the present studies would seem to necessitate the assumption that hearing the ambiguous word automatically activates two or more of its meanings in memory and that this activation occurs prior to the visual presentation of a word.

The present experiments suggest that, although linguistic context can be effective in resolving ambiguities in the language, this disambiguating effect does not occur during the lexical look-up process. Rather, there seems to be a process during sentence comprehension in which the meanings of words in a sentence are activated in memory without regard to the available linguistic context. This, then, implies that context has its disambiguating effect at some time subsequent to the lexical look-up process. Thus, we are still left with the problem of how contextual information and lexical information are integrated to produce a meaning for an ambiguous sentence. MacKay's perceptual suppression hypothesis attempts to account for this type of integration, but the results of Experiment II do not support his hypothesis. The simplest alternative might be a theory in which contextual information (and perhaps other types of information) increases the strength of one or more of the activated meanings of a word to the point that it reaches threshold and becomes conscious. Certainly, other alternative theories are also plausible. A next step, then, in trying to gain an understanding of sentence



comprehension might be to formulate and test these integration theories. The color-naming methodology used in the present studies may prove to be a useful tool for these investigations.

## REFERENCES

- Bransford, J. D., & Johnson, M. K. Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- Cairns, H. S. Effects of bias on processing and reprocessing of lexically ambiguous sentences. *Journal of Experimental Psychology*, 1973, 97, 337-343.
- Carey, P. W., Mehler, J., & Bever, T. G. When do we compute all the interpretations of an ambiguous sentence? In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. New York: American Elsevier, 1970.
- Conrad, C. An analysis of a hierarchical model of semantic memory organization. Unpublished master's thesis, University of Oregon, 1971.
- Conrad, C. Studies of the subjective lexicon. Unpublished PhD dissertation, University of Oregon, 1972.
- Fodor, J., Garrett, M., & Bever, T. Some syntactic determinants of sentential complexity, II: Verb structure. *Perception & Psychophysics*, 1968, 3, 453-461.
- Foss, D. J. Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning & Verbal Behavior*, 1970, 9, 699-706.
- Foss, D. J., Bever, T. G., & Silver, M. The comprehension and verification of ambiguous sentences. *Perception & Psychophysics*, 1968, 4, 304-306.
- Foss, D. J., & Jenkins, C. M. Some effects of context on the comprehension of ambiguous sentences. Cited by D. T. Hakes, On understanding sentences: In search of a theory of sentence comprehension. Final report, October 1972, U.S. Office of Education, Washington, D.C.
- Garrett, M. F. Does ambiguity complicate the perception of sentences? In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. New York: American Elsevier, 1970.
- Katz, J. J., & Fodor, J. A. The structure of a semantic theory. *Language*, 1963, 39, 170-210.
- Lackner, J. R., & Garrett, M. Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition*, 1973, in press.
- Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*. New York: Wiley, 1951.
- MacKay, D. G. To end ambiguous sentences. *Perception & Psychophysics*, 1966, 1, 426-436.
- MacKay, D. G. Mental diplopia: Towards a model of speech perception at the semantic level. In G. B. Flores d'Arcais and W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. New York: American Elsevier, 1970.
- MacKay, D. G. Aspects of the theory of comprehension, memory and attention. *Quarterly Journal of Experimental Psychology*, 1973, 25, 22-40.
- Marshall, J. C. Syntactic analysis as a part of understanding. *Bulletin of the British Psychological Society*, 1965, 18, 28.
- Miller, G. The subjective lexicon. Presidential address, 78th Annual Convention, American Psychological Association, September 1970.
- Palermo, D. S., & Jenkins, J. J. *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press, 1964.
- Posner, M. I. Coordination of internal codes. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.
- Warren, R. E. Association, directionality and stimulus encoding. *Journal of Experimental Psychology*, 1973, in press.
- Warren, R. E. Stimulus encoding and memory. *Journal of Experimental Psychology*, 1972, 94, 90-100.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received for publication May 25, 1973;  
revision received June 25, 1973.)