

# Spilling the beans on understanding and memory for idioms in conversation

RAYMOND W. GIBBS, JR.

*University of California at San Diego, La Jolla, California 92093*

Three experiments examine people's understanding and memory for idioms. Experiment 1 indicates that in a conversational context, subjects take less time to comprehend conventional uses of idiomatic expression than unconventional, literal uses. Paraphrase judgment errors show that there is a strong bias to interpret idiomatic expressions conventionally when there is no preceding context; however, subjects interpret literal uses of these expressions correctly when there is appropriate context. Experiment 2 showed that in a free recall task, literal uses of idioms are remembered better than conventional uses of these utterances. Experiment 3 indicated that in conversation, literal and idiomatic recall prompts facilitate memory for literal uses of idioms equally well. The results from these experiments suggest that memory for conventional utterances is not as good as for unconventional uses of the same utterances and that subjects understanding unconventional uses of idioms tend to analyze the idiomatic meaning of these expressions before deriving the literal, unconventional interpretation. It is argued that the traditional distinction between literal and metaphoric language is better characterized as a continuum between conventional and unconventional utterances.

Ordinary language is full of interesting phrases and expressions. Many of these often have intended meanings that differ from their literal interpretations. For example, colloquial statements like "You can let the cat out of the bag," will in many contexts be used to mean "You can reveal the secret." When a listener hears an idiomatic utterance in conversation, the problem is to interpret the intended meaning of the speaker, not just the literal meaning of the sentence.

In the present experiments, I examined how people interpret and remember conventional and unconventional uses of idioms. Other research on nonliteral language processing has focused on the role that literal meaning plays in understanding figurative sentences. Clark and Lucy (1975; see also Clark & Clark, 1977) presented results on understanding conversationally conveyed requests that support a serial process model. This model says that a person understanding a conveyed request like "Can you pass the salt?" must (1) determine the literal meaning of the utterance, (2) compare the literal meaning with the context, (3) if the literal meaning is inappropriate, derive the conveyed meaning via a cooperative principle or some other social convention

(Grice, 1975; see also Gordon & Lakoff, 1971), and, finally, (4) use the utterance on the basis of its conveyed meaning.

In earlier work (Gibbs, 1979), I questioned whether people must compute the literal meaning of an utterance before they can derive the intended or indirect meaning. My hypothesis was that hearers can use context to understand indirect utterances directly, without first analyzing the literal form of the sentence. In those experiments, I had subjects read and then verify paraphrases of indirect requests embedded within a story context. The story context provides a more realistic test of normal sentence understanding than the isolated sentences studied by Clark and Lucy (1975). Using these more natural stimuli, I found that subjects took longer to understand literal sentences than conveyed ones. Without story context, as in the Clark and Lucy experiment, I found that indirect sentences took longer to comprehend than literal sentences. A second experiment found that, in a story context, subjects took longer to understand a direct request, "Do not open the window," than to verify an indirect request, "Must you open the window?" The data from these experiments suggest that a person understanding an indirect request need not construct the literal interpretation before deriving the conveyed request. The reaction time (RT) differences found in the Clark and Lucy (1975) experiment, then, may not reflect any additional stages of processing in which the literal meaning of an utterance is first determined. Rather, the differences may indicate problems of understanding requests of different linguistic styles without appropriate context.

Under most situations of language use, then, the presence of appropriate contextual information should

This research was supported by National Institute of Education Grant G-77-0015 to David E. Rumelhart and Public Health Service Grant MH 15828-09 to the Center for Human Information Processing. I would like to thank Robert Glushko and David Rumelhart for their helpful comments on an earlier draft of this paper. Additional thanks are due to the LNR Research Group for their advice during the course of this work. Reprint requests may be sent to Raymond W. Gibbs, Jr., Department of Psychology C-009, University of California at San Diego, La Jolla, California 92093.

affect ease of comprehension. If the context in which an utterance is spoken is sufficient, it will provide the listener a semantic framework in which to understand the sentence. Without context, the listener has no predictive power and, consequently, must process the utterance in more of a bottom-up manner (see Ortony, Schallert, Reynolds, & Antos, 1978; Rumelhart, 1977, for further details of such a model). While this model of linguistic processing seems adequate for most types of utterances, does it still apply to highly conventionalized language like idioms? Because idioms have strong conventional meanings associated with them, it is possible that context plays much less of a role in helping the listener construct an appropriate interpretation. At the same time, context should play a crucial role in getting the listener to understand an unconventional use of an idiomatic expression. What, then, are the processes by which conventional and unconventional uses of idioms are understood in conversation?

Experiment 1 investigates subjects' comprehension of idioms that can have either a literal or an idiomatic meaning. My hypothesis is that the literal meaning of idioms is not important in understanding these expressions. Rather, it is the conventionality of the utterance that effects its ease of comprehension. Therefore, subjects need not interpret the literal meaning of a conventional use of an idiom in conversation before deriving the conveyed interpretation. On the other hand, when subjects hear an unconventional, literal use of an idiom, they will automatically analyze the conventional, idiomatic interpretation before deciding that the literal meaning is appropriate.

In Experiment 1, subjects read stories one line at a time; the last line was an idiomatic expression. After reading the target sentence, subjects made a paraphrase judgment. The amount of time it took subjects to read each line and to make the paraphrase judgment was measured. I predicted that in a story context, it would take subjects less time to understand and make paraphrase judgments for idioms with conventional meaning than for those same sentences with unconventional interpretations. In order to examine the effect of con-

text on understanding of conventional and unconventional uses of idioms, I also had subjects read and make paraphrase judgments for the same sentences, without any preceding story context. Here, I predicted that subjects would also be faster at comprehending conventional uses of idioms than unconventional uses. However, subjects will make many more paraphrase judgment errors for the literal uses of idioms, because there is no context to induce the appropriate unconventional meaning of these utterances.

### EXPERIMENT 1

#### Method

**Subjects.** Forty-eight undergraduates attending the University of California at San Diego served as subjects. They received either pay or credit for a course requirement. All subjects were native speakers of English.

**Stimuli.** Four types of experimental stimuli were used. These were produced by factorially crossing context (with context or no context) with sentence content (literal or idiom).

Table 1 presents examples of the four types of stimuli. Within each context condition, subjects were randomly assigned to one of two sets. In each condition one set was the complement of the other.

In the with-context condition, each paragraph ended in a target sentence that was followed by a possible paraphrase. The paraphrase was either literal (literal condition) or idiomatic (idiom condition). Note that in each set the same target means something different. In Set A, the target sentence has a literal interpretation; in Set B, the same sentence has an idiomatic meaning.

Each subject in the with-context condition was presented eight stories with a "literal" target sentence and eight stories with an "idiomatic" target sentence. Both of these sentence contents had true paraphrases. In addition to these sentences, there were 12 stories that had target sentences whose paraphrases were false. These filler stories were not included in data analysis. Each subject therefore saw 28 stories in the session, along with 4 practice stories: 1 literal and 1 idiomatic story, along with 2 filler stories. The stories averaged six lines in length, and the target sentences averaged six words in length. The paraphrases averaged five words in length. For a given target sentence, both types of its paraphrases, the literal and the idiomatic, had exactly the same number of words. All of the idioms were taken from Boatner, Gates, and Makkai (1975).

Stimuli in the no-context condition were simply the final sentences and paraphrases used in the with-context condition.

Table 1  
Stimuli

Set A (Literal Context)	Set B (Idiomatic Context)
With Context	With Context
Nick and Sue were listening to Jackson Browne on the radio. "All Jackson Browne songs sound alike." Sue said. "Now isn't that the same song we heard him do on TV recently." "No." Nick replied; "He's singing a different tune."	On TV there was a program discussing Carter's first year in office. One reporter talked about the military budget. "In the campaign Carter promised to cut that budget." "But now that he's President," "He's singing a different tune."
Paraphrase: "He's not singing the same song."	Paraphrase: "He has now changed his mind."
No Context	No Context
"He's singing a different tune."	"He's singing a different tune."
Paraphrase: "He's not singing the same song."	Paraphrase: "He has now changed his mind."

In one set the paraphrase was the literal interpretation of the sentence. The other set had a paraphrase for the same target sentence that was the idiomatic interpretation of the sentence. As in the with-context condition, the fillers for the no-context condition were the same for both sets. The fillers here included the target sentences and their false paraphrases, used as fillers in the with-context condition. However, to insure that subjects did not simply interpret every sentence they saw as an idiom, 16 additional sentences were included that were not idiomatic expressions. Eight of these had true paraphrases, and eight had false paraphrases.<sup>1</sup>

**Procedure.** The stimuli were presented in a different random order for each subject on a computer terminal screen that was under the control of a PDP-11/45 computer. The computer controlled the response keyboard and recorded the latencies of the comprehension and paraphrase responses. In the with-context condition, subjects were seated in front of the terminal screen. They pushed any button on the keyboard, and the first sentence appeared. Subjects read the sentence and pushed a button as soon as they understood what the sentence meant. When they gave their responses, the first sentence was replaced by the second sentence of the story. Subjects read and gave a comprehension response for each sentence of the story. After he gave his comprehension responses for the last sentence, which was the target, the subject heard a tone, and a possible paraphrase of the previous sentence appeared. The tone signified to the subjects that they were to make a true-false paraphrase judgment on a designated part of the keyboard. After they had done this as quickly as possible, the trial was over. To get the first sentence for the next trial, the subject could push the comprehension response button. For both the comprehension response and the true-false paraphrase judgment, subjects were timed from the onset of the display to the instant the response button was pushed. The instructions emphasized that the subject should try to understand exactly what each sentence said before he pushed the comprehension response button or the true-false paraphrase buttons.

Each session in the with-context condition lasted about 20 min, and each session in the no-context condition lasted about 10 min.

## Results and Discussion

Only those target sentences that were followed by a correct paraphrase judgment were included in the data analysis. Means were calculated by averaging across subjects.

**Paraphrase judgments and error rates.** Table 2 presents the response times to make the paraphrase judgments along with the proportion of errors that subject made in making these judgments.

With a story context, subjects took 696 msec longer to make the paraphrase judgments for the literal sentences than they did for the idiomatic sentences. Statistical analyses revealed a significant main effect of sentence type (i.e., literal or idiomatic) [ $\min F'(1,47) =$

15.72,  $p < .001$ ]. There was no significant effect of context (i.e., with context or no context), nor was there a significant interaction between sentence type and context ( $F_1$  and  $F_2 < 1$ ). Newman-Keuls tests on the individual means revealed significant differences ( $p < .001$ ) between the conveyed and literal sentences in both the with-context and no-context conditions. This result confirms the Ortony et al. (1978) finding that conventional uses of idioms do not take longer to comprehend than literal uses of the same expressions.<sup>2</sup> Moreover, the RT differences found in the present experiment indicate that idioms take significantly less time to process than literal interpretations of these expressions.

When there was no accompanying story context, subjects took 679 msec longer to make paraphrase judgments for the literal sentences than for the conveyed. This result is quite similar to the with-context condition. However, examination of the error rates confirms the importance of context in understanding the target sentences. In the no-context condition, subjects made far more errors for the literal paraphrases than for the idiomatic ones [ $\min F'(1,27) = 41.05$ ,  $p < .0001$ ]. That is, more than half of the time, subjects responded that the literal paraphrase was not a correct paraphrase of that sentence out of context. Without context, there are not enough pragmatic and semantic clues for the subject to recognize the potential unconventional interpretation. A sentence like "You can let the cat out of the bag," heard with no supporting contextual information, appears to mean something like "You can reveal the secret." Therefore, subjects usually give a false response to a literal paraphrase, "You can let the cat out."

**Comprehension times for the target sentences.** While the RTs for the paraphrase judgments are an excellent indicator of comprehension, it is also important to examine the comprehension times for the target sentences. To do this, the mean response times for the correct paraphrase judgments and the reading times for their preceding target sentences were added together to form a conservative measure of comprehension time. This measure is useful because a subject may push the comprehension button for the target without fully understanding its meaning. If the subject does understand the target sentences when he pushes the comprehension button, then the paraphrase judgment should be made quickly. Adding the paraphrase judgment RT to the target sentence RT insures a more accurate measure of the subjects' comprehension of the target sentence. This technique has been used in similar research paradigms by Clark and his associates (cf. Springston, cited in Clark & Clark, 1977) and Gibbs (1979).

In the with-context condition, subjects took less time to comprehend and make paraphrase judgments for the idiomatic sentences (3,424 msec) than for the literal sentences (4,328 msec) [ $\min F'(1,24) = 11.45$ ,

Table 2  
Mean Response Times (RT) and Error Rates (ER)  
for Paraphrase Judgments

	Literal Sentences		Idiomatic Sentences	
	RT	ER	RT	ER
With Context	2504	.08	1808	.05
No Context	2544	.51	1865	.15

$p < .01$ ]. This provides further evidence for the idea that people do not understand the literal interpretation of sentences before they derive the idiomatic or conveyed meaning, and this supports the results from Gibbs (1979) on understanding indirect requests and Ortony et al. (1978) on comprehension of metaphor.

These results demonstrate that subjects spend significantly more time processing idioms with literal meanings than those with idiomatic interpretations. This suggests that ease of comprehension for nonliteral language may be more a matter of how conventional a sentence is than how literal or metaphoric it may be. The more conventional an utterance, the easier it will be for a person to find an appropriate interpretation in the right context. Unconventional utterances, like the literal use of idiomatic expressions, will require additional processing in order to find and verify some schemata in memory to account for the sentence (cf. Rumelhart, 1979). What does this suggest about how people remember idiomatic expressions in conversation?

Because conventional uses of idioms are very familiar, one could predict that people will remember them very well in a recall test. However, when an idiom is used in a literal context, people might automatically interpret it according to its conventional meaning first, before deciding that its literal, unconventional meaning is really intended. This "double-take" reaction could lead to excellent recall for these unconventional uses of idioms because of the extra processing the listener has to do to understand these sentences. Memory for conventional utterances will not be as distinctive, despite the fact that they are easier to understand, because they do not require elaborative processing in order to be integrated into the context of the conversation. Therefore, I predicted that subjects would recall idiomatic expressions with literal, unconventional meaning better than sentences with conventional interpretations.

In Experiment 2, subjects heard two repetitions of the series of stories used in Experiment 1. After an interval of 24 h, subjects were asked to recall the stories and write down the last line of each story. In addition, different subjects heard the same stories but with the paraphrases from Experiment 1 as the last sentences. This provided a control for context to insure that each context type (literal and idiomatic) equally induced their appropriate interpretation.

## EXPERIMENT 2

### Method

**Subjects.** Forty-eight undergraduates from the University of California at San Diego served as subjects. They received course credit or were paid for their participation. All were native speakers of English.

**Stimuli and Procedure.** The stimuli for this experiment were the same used in Experiment 1. All of the stories were recorded on a tape cassette using three speakers: one male and one female, who engaged in a dialogue, and one male narrator. The stories were read in a normal intonation. In each story, the male and

female spoke their appropriate parts, and the narrator said those parts of each story that were not dialogue.

There were two parts to the procedure: acquisition and recall. During acquisition, subjects entered an experimental room and were told to listen carefully to the series of stories. The tape was played once. While the tape was being rewound for the second presentation, subjects were told to listen carefully again and that they were going to be asked some questions when they returned the next day. Twenty-four hours later, subjects came back and were asked to remember the stories they had heard the previous day and to write down, as exactly as possible, the last line from each story.

### Results and Discussion

A sentence was scored as being correct if it contained all the important content words and had the same syntactic structure as the acquisition sentence. The mean proportion of each type of sentence correctly recalled by subjects is shown in Table 3. Recall for the literal target sentences was better than for the idiomatic expressions. A planned-comparison test revealed this difference to be highly significant ( $p < .01$ ). However, subjects recalled just as many idiomatic paraphrases as literal ones. Statistical analysis revealed a significant main effect of sentence type (i.e., idiom expression or paraphrase) [ $\text{min } F'(1,72) = 8.05, p < .01$ ]. The effect of sentence content (i.e., literal or idiom) was not significant [ $\text{min } F'(1,74) = 1.35$ ]. The interaction of sentence type and sentence content, however, was reliable [ $\text{min } F'(1,75) = 6.47, p < .05$ ]. These data indicate that, while idiomatic expressions are more familiar, literal interpretations of these expressions are better recalled. In other words, under normal circumstances of conversation, people remember unconventional uses of idioms better than conventional uses.

The fact that subjects recalled the same proportion of idiomatic and literal paraphrases demonstrates that there is no biasing in the story context that would facilitate understanding or recall of the target sentences in one story type over another. This is congruent with the Ortony et al. (1978) finding that literal equivalents take the same amount of time to understand as do idiomatic paraphrases. Since subjects process the paraphrases to the same extent, they are remembered equally well in the recall task.

These results are consistent with the descriptions framework for memory retrieval suggested by Norman and Bobrow (1979). A description is the specification of a target item in memory used by the retrieval mechanism in recalling that item. It can be part of the content of the target item or any specification of the operations

Table 3  
Mean Proportion of Sentences Correctly Recalled

Sentence Type	Sentence Content	
	Literal	Idiomatic
Idiom	.61	.43
Paraphrase	.40	.39

that will lead to successful retrieval. A memory record of an utterance will contain a description of it based on the interpretation of the item when it is first heard. The recall of an item, therefore, is dependent upon recreating the description made at the time of input and using it to recall the utterance. The more descriptions associated with an utterance in memory, the more unique its encoding, and the more discriminable it will be from other utterances in memory.

Experiment 1 suggests that subjects elaborate upon their initial interpretation of the utterance until they understand the utterances appropriately as having literal meaning. In doing this additional processing, the number of descriptions for the utterance increases, thus raising the discriminability of the utterance in memory. Given this greater discriminability, the retrieval process is more efficient, and the utterances are better remembered. As Norman and Bobrow (1979) note, this idea is similar to the "levels of processing" view proposed by Craik and Lockhart (1972). However, the emphasis here is on the "levels of descriptions" used at the time of retrieval.

Along this same line, Jacoby and Craik (1979) have recently emphasized the importance of the interaction between distinctive encoding and adequate retrieval information for recall. Specifically, if the original encoding of an item is difficult to accomplish, later memory of the event will usually be good. The results of Experiment 1, in which unconventional uses of idioms were more difficult to understand than conventional uses, along with the findings of the present experiment, in which unconventional sentences were remembered better than conventional ones, support this notion. Both Jacoby and Craik (1979) and Norman and Bobrow (1979) suggest that initial difficulty of encoding is associated with the formation of a more complete "description" of the stimulus item and thus with a more unique trace of the sentence in memory.

Of the sentences recalled, 90% were scored as correct. The remaining 10% either had key content words missing, as in "You can let it out of the bag," or had a syntactic structure different from the original. While the actual numbers of these recall errors are small, some of them do indicate that subjects make use of memory descriptions in retrieval. Since subjects formulate descriptions to aid in recall of the sentences, the descriptions they use will sometimes incorrectly retrieve utterances that are quite similar in some way to the original target item. For example, some of the recall errors were more literal paraphrases of the idioms. If subjects heard "He's singing a different tune," they sometimes recalled it as "He's singing a different song." Although this type of error was more frequent when subjects heard a literal context ( $n = 7$ ), it also occurred when subjects heard the idiomatic conversation ( $n = 3$ ). Occasionally, subjects substituted one idiom of the same meaning for another ( $n = 3$ ). One subject, for example, recalled, "Please don't make waves," when he heard "Please don't rock the boat," meaning "Please don't cause trouble."

The results of the present experiment, along with those of Experiment 1, suggest that subjects do some sort of double-take when they hear an unconventional use of an idiom. This double-take, in which the conventional interpretation of the utterance is first analyzed and rejected, results in subjects' taking longer to process unconventional uses of idioms than conventional uses, as Experiment 1 has demonstrated, and in greater memory for unconventional idioms, as the present study has shown. Nonetheless, the extra amount of time to comprehend unconventional uses of idioms does not necessarily support the notion that subjects process these utterances to a deeper level of analysis (Craik & Tulving, 1975). To further examine what exactly is going on during this extra processing of literal uses of idioms, Experiment 3 was conducted.

In Experiment 3, a cued recall task was used to examine the memory representation for the unconventional and conventional interpretations of idioms. As shown by other researchers (Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974; Honeck, Reichmann, & Hoffman, 1975; Tulving & Thompson, 1973; Verbrugge & McCarrell, 1977), prompted recall provides a sensitive measure of any inferential activity that may occur during comprehension, and it examines the representation in memory for linguistic material.

Experiment 3 tested whether a literal paraphrase of an idiomatic expression can be an effective prompt for its recall. If subjects analyze the conventional, idiomatic meaning of these expressions at some point during comprehension, then a verbal statement of this should be an effective recall prompt. However, if comprehension of an unconventional use of an idiomatic expression does not necessitate some analysis of its conventional meaning, then a literal prompt taken from the expression (like "cat" from "you can let the cat out of the bag") would facilitate recall better than would the idiomatic paraphrase.

### EXPERIMENT 3

#### Method

**Stimuli and Design.** The same two sets of stories (Sets A and B) used in Experiments 1 and 2 were played to different groups of subjects. The recall prompts were assembled into two sets of cues, literal (corresponding to the unconventional interpretation) and idiomatic (corresponding to the conventional meaning of the expression). All of the literal prompts were judged to be key words in the expression (e.g., "cat" from "You can let the cat out of the bag" or "ice" from "Aren't you skating on thin ice?"); the idiomatic prompts were two-word paraphrases of the meaning of each expression (e.g., "reveal secret" for "You can let the cat out of the bag" and "dangerous risk" for "Aren't you skating on thin ice?").

For both Sets A and B, subjects heard stories with literal and idiomatic target sentences. During recall, however, subjects received a booklet that contained either literal or idiomatic prompts. Therefore, each subject saw some prompts that were not relevant to the target sentences. For example, a subject would hear "You can let the cat out of the bag" with a literal

meaning, but during recall, he would be presented with an idiomatic prompt ("reveal secret").

In addition, there was a no-context condition. Here, subjects heard a list of 40 sentences, 16 of which were idiomatic expressions. As in the with-context conditions, subjects heard the list of sentences twice, and the next day they were given a list of prompts and asked to recall the appropriate sentences.

**Subjects.** Subjects were 48 undergraduates at the University of California at San Diego. All received credit for a course requirement. Subjects were randomly assigned to one of three acquisition conditions: Set A, Set B, or no context. During recall, half of the subjects in each condition received the literal prompts; the other half were presented the idiomatic cues.

**Procedure.** The acquisition instructions were the same as in Experiment 2. For the recall phase, subjects were given the list of prompts and were asked to write down, as verbatim as possible, the last line of each story, given each prompt. It was emphasized that the prompts need not have anything to do with the actual words in the target sentences but that they may be related to their meaning. Following the recall instructions, subjects were given approximately 16 min to recall the sentences.

### Results and Discussion

Sentences were scored correct according to the same criteria used in Experiment 2. The mean proportion of target sentences correctly recalled by subjects is given in Table 4. The results for the with-context condition indicate that subjects better recalled both the literal and idiomatic interpretations of the expressions when given literal cues than when they were given idiomatic prompts [ $\min F'(1,58) = 4.14, p < .05$ ]. Moreover, for both literal and idiomatic prompts, there were only very slight differences in the proportion of literal and idiomatic target sentences recalled ( $F1$  and  $F2 < 1$ ). The near-equal proportion of literal sentences recalled, given either literal or idiomatic prompts, suggests that subjects analyzed these sentences idiomatically at some point during comprehension. Thus, in conjunction with the results of Experiment 2, these results appear to indicate that the double-take reaction that occurs when subjects hear an unconventional use of an idiom is really the subjects' automatically analyzing the conventional, idiomatic meaning of these expressions first, before deciding that the literal meaning is appropriate.

For the no-context condition, subjects recalled significantly more target sentences when given literal prompts than when given idiomatic prompts [ $\min F'(1,11) = 4.91, p < .05$ ]. An analysis of the individual proportion of sentences correctly recalled between the with-context and no-context conditions revealed significant differences for

both literal cues [ $\min F'(1,62) = 13.20, p < .001$ ] and idiomatic cues [ $\min F'(1,60) = 8.49, p < .01$ ]. This indicates that having an appropriate context in which to understand these expressions will later facilitate their recall, given literal and idiomatic prompts.

The results in the no-context condition, however, point out one potential problem in interpreting the results of the present experiment. The superiority of the literal cues over idiomatic cues may not be a difference between prompting the literal vs. the idiomatic meaning of the utterances as it is a comparison between lexical and idiomatic prompting. Thus, the literal sense of the expressions are not being prompted in the literal cue condition, but key words in the sentences are being prompted. The ideal way to control for this problem would be to use literal cues that do not contain any of the same words in the target sentence. For example, the phrase "release animal" could be used as a literal prompt for the expression, "You can let the cat out of the bag." However, it is not possible to create adequate literal prompts for all of the sentences without using some of the key words in the expressions. Good literal prompts for expressions like "Aren't you skating on thin ice?" or "That's for the birds" are impossible to create without using the lexical items "ice" or "birds." Therefore, it appears to be quite difficult to do such a control experiment.

Nonetheless, the advantage of the literal over the idiomatic prompts in the no-context condition, particularly in comparison with the nearly equal facilitation of recall between the literal and idiomatic prompts in the with-context condition, highlights the importance of appropriate context in the interpretation of idiomatic expressions. Thus, in the context of a conversation subjects will tend to analyze the conventional, idiomatic form of the utterance at some point during comprehension of unconventional uses of idioms, but it is not clear that subjects necessarily do this when there is no adequate context. The results from the no-context condition of Experiment 1 support this idea. There, subjects' error rates reveal that, almost half the time, subjects interpreted the unconventional, literal meaning of these expressions as being appropriate. It appears that while highly conventionalized utterances like idioms are somewhat context-independent, without appropriate context subjects will occasionally interpret these utterances according to their putative, literal meaning.

Within Norman and Bobrow's (1979) memory framework, the present data suggest that the description of the information being sought (the recall prompts) matches the descriptions created by the subject when first listening to these expressions. Recall prompts provided the subjects with descriptions that guide the memory search process and help verify whatever memory record is retrieved. Since the description of the information sought (an idiomatic prompt) matches the descriptions created by subjects, when first listening to

Table 4  
Mean Proportion of Sentences Correctly Recalled

Sentence Prompts	Sentences Recalled		
	With Context		No Context
	Literal	Idiomatic	
Literal	.73	.71	.53
Idiomatic	.57	.54	.31

the sentences subjects recall literal uses of idioms just as well when given idiomatic prompts as when given literal recall cues.

The recall errors were very similar to those described in Experiment 2. In the with-context condition, 94% of the sentences recalled were scored as correct, and 85% were correct in the no-context condition. While there are too few recall errors to justify any firm conclusions, it appears that the conversational context provides additional information to enable subjects to formulate more specific descriptions of the target item. This supports the idea that there is better and more accurate memory for idioms heard in the context of a conversation than for idioms with no accompanying context.

### GENERAL DISCUSSION

The results of the present experiments suggest that what is important in language processing is not whether a particular utterance is stated literally or metaphorically. Rather, it is the conventionality of a sentence that makes it easier or more difficult to comprehend and remember. The results of Experiment 1 indicate that conventional utterances are easier to integrate into the context of the conversation than are unconventional uses of the same sentences. Moreover, as the results of Experiment 3 demonstrate, when an unconventional use of an idiom is encountered, people tend to automatically analyze the conventional, idiomatic meaning of the utterance before deciding that the literal meaning is appropriate. These data argue against a serial process model of nonliteral language processing, in which the literal meaning of an utterance is first determined, before the conveyed interpretation is derived. Literal meaning is only important to the extent that it relates to the speakers' intentions in realistic linguistic and social contexts.

Just as the conventionality of an idiom is important in understanding it, so is it important in remembering it. Experiment 2 indicated that in conversation people remember the literal, unconventional uses of idioms better than the conventional uses. This is so despite the fact that people are very familiar with idiomatic uses of idioms that have become hackneyed through overuse. In conjunction with the results of Experiment 3, in which idiomatic cues and literal cues led to equally good recall for literal uses of idioms, it appears that when people hear an unconventional use of an idiom, it produces a sort of double-take reaction. When this happens, people tend to automatically analyze the conventional interpretation of the utterance, before deciding that the unconventional meaning is appropriate. Thus, people formulate additional descriptions for these utterances that make them more distinctive in memory and facilitate better retrieval at a later time. Memory for the conventional meaning of utterances will not be as distinctive, because these sentences do not require elaborative processing in

order to be integrated into the context of the conversation.

Future research on language processing should focus on the relation between the conventionality of an utterance and the specific context in which it is spoken. What effect does the interaction of convention and context have on understanding and memory for various types of speech? By concentrating on these aspects of language use, we can begin to formulate models of linguistic processing based on how speakers conventionally use words to convey meaning instead of on the putative, literal interpretation of sentences outside of everyday conversation.

### REFERENCES

- ANDERSON, R. C., & ORTONY, A. On putting apples into bottles—A problem of polysemy. *Cognitive Psychology*, 1975, 7, 167-180.
- BARCLAY, J. R., BRANSFORD, J. D., FRANKS, J. J., MCCARRELL, N. S., & NITSCH, K. E. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 471-481.
- BOATNER, M. T., GATES, J. E., & MAKKAJ, A. *A dictionary of American idioms*. New York: Baron's Educational Series, 1975.
- CARPENTER, P., & JUST, M. Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 1975, 82, 45-73.
- CLARK, H. Linguistic processes in deductive reasoning. *Psychological Review*, 1969, 74, 387-404.
- CLARK, H., & CHASE, W. On the processes of comparing sentences against picture. *Cognitive Psychology*, 1972, 3, 472-517.
- CLARK, H. H., & CLARK, E. V. *Psychology and language*. New York: Harcourt, Brace & Jovanovich, 1977.
- CLARK, H., & LUCY, P. Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 56-72.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, 104, 268-294.
- GIBBS, R. W. Contextual effects in understanding indirect requests. *Discourse Processes*, 1979, 2, 1-10.
- GLUSHKO, R., & COOPER, L. Spatial comprehension and comparison processes in verification tasks. *Cognitive Psychology*, 1978, 10, 391-421.
- GORDON, D., & LAKOFF, G. Conversational postulates. *Papers from the Seventh Regional Meeting, Chicago Linguistic Society*, 1971, 63-84.
- GRICE, P. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3). New York: Academic Press, 1975.
- HONECK, P., REICHMANN, P., & HOFFMAN, R. Semantic memory for metaphor: The conceptual base hypothesis. *Memory & Cognition*, 1975, 3, 409-415.
- JACOBY, L. L., & CRAIK, F. I. M. Effects of elaboration of processing on encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, N.J.: Erlbaum, 1979.
- NORMAN, D. A., & BOBROW, D. G. Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 1979, 11, 107-123.
- ORTONY, A., SCHALLERT, D. L., REYNOLDS, R. E., & ANTOS, S. J. Interpreting metaphors and idioms: Some effects of context

- on comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 465-467.
- RUMELHART, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Erlbaum, 1977.
- RUMELHART, D. E. Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and thought*. London: Oxford University Press, 1979.
- TULVING, E., & THOMPSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- VERBRUGGE, R., & MCCARRELL, N. Metaphoric comprehension: Studies in reminding and resembling. *Cognitive Psychology*, 1977, 9, 494-533.

#### NOTES

1. A rating study of all the stories was done to rule out the possibility of any systematic differences of the contexts in inducing the appropriate literal and idiomatic meanings of the target and paraphrase sentences. Twenty-four subjects participated in the rating study. Each subject saw a target sentence embedded in either a literal or an idiomatic story context. Half of the subjects received the literal story for a particular target sentence. The other subject saw the idiomatic story for that same sentence. Subjects were asked to indicate whether a particular target sentence should have a literal or an idiomatic meaning and to make a rating on a 5-point scale of the degree to which a target sentence related to its context. In addition,

subjects were also asked to rate how well they understood the target expressions and how familiar they were with these expressions on a 5-point scale.

From this rating study, 16 out of the 24 stories were selected. For these chosen stories, results showed that 97% of the time subjects agreed that the literal stories induced the literal interpretation of the target sentences; there was 96% agreement for the idiomatic stories. An analysis of the ratings showed that there was no significant difference between the overall effectiveness of the literal and idiomatic stories and that no individual story or target expression in either condition varied more than 1 SD from the mean rating. These results suggest that there are no differences in the story contexts that would bias comprehension of the target sentences.

2. It should be noted that Ortony et al. (1978) did not have subjects make paraphrase judgments for their target sentences. Previous research on sentence comprehension and use (Carpenter & Just, 1975; Clark, 1969; Clark & Chase, 1972; Glushko & Cooper, 1978) has demonstrated the need to use some secondary task, like a sentence verification paradigm, to measure the comprehension criteria used by subjects in understanding sentences. In the present experiment, subjects' errors are important in determining if subjects actually understood the target sentences.

(Received for publication August 15, 1979;  
revision accepted November 30, 1979.)