

# Comprehending ambiguity in the sentence-verification paradigm: Basic process or problem solving?

RICHARD REARDON  
*University of Oklahoma, Norman, Oklahoma*

and

STUART KATZ  
*University of Georgia, Athens, Georgia*

**A linguistic ambiguity experiment using a sentence-verification procedure was performed to determine whether and how subject awareness of ambiguity affects response latency and error. Results showed a decline first in errors and then in latencies for the one sentence pair type most likely to fool subjects (the "ambiguous-unexpected" type), thus supporting the hypothesis that subjects who are aware of ambiguity attempt to jointly minimize latency and error (Pachella, 1974). The findings suggest that research in this area, instead of revealing "basic" information-processing mechanisms, may instead reflect conscious subject strategies adapted to an idiosyncratic context.**

Investigators interested in the mechanisms of natural-language performance have often turned to the "ambiguous" sentence as an experimental device. Ambiguous sentences are those having more than one meaning, either because of a particular syntactic arrangement or because they contain words with more than one meaning. One issue in this research has been whether or not the language user processes just one interpretation of an ambiguous sentence or chooses among several possible interpretations (the "one-meaning" and "many-meaning" hypotheses, respectively; Cairns, Note 1).

Consider, for example, the popular research paradigm in this area called "sentence verification" (Cairns, 1973; Foss, Bever, & Silver, 1968; Mistler-Lachman, 1972). Subjects read or listen to an ambiguous or unambiguous sentence, after which they are presented with a second sentence (or sometimes a picture). Their task is to judge whether the second sentence "follows from" the first. The second sentence may represent either the "expected" or the "unexpected" meaning of the ambiguous sentence, expectancy being a function of the probability that one or another of the meanings of an ambiguous sentence is the first to come to mind. Typically, longer latencies and a larger number of false-negative errors for these judgments have been found for ambiguous sentences followed by a second sentence with an unexpected meaning (the "ambiguous-unexpected" pair type). Such findings have been interpreted as supportive of the one-meaning hypothesis. Cairns (1973), for example,

stated that the longer latencies reflect a "reprocessing" stage; the language-processing mechanism, first committed to the expected meaning of an ambiguous sentence, must switch to the unexpected meaning so that the second sentence does follow from the first. The larger number of errors is explained as a commitment to the expected meaning before reprocessing can take place; the subject indicates that the second sentence does not follow from the first, although, of course, it does.

One wonders, however, whether the findings in these experiments reveal unconscious language-processing mechanisms or whether they reflect behaviors that are a function of the unique characteristics of particular experimental arrangements. Consider again the sentence-verification paradigm. If subjects respond correctly to an ambiguous-unexpected sentence pair, it can be argued that they have become aware of the ambiguity of the first sentence of the pair and are "reprocessing" it only because they realize that the second sentence could follow from the less likely of the two possible meanings of the first. Awareness implies that subjects now know what the experimenter knows, namely, that ambiguity is a special property of the stimulus materials and that its purpose is to aid in determining whether or not the subjects will commit themselves (sometimes incorrectly) to the more likely of two meanings of an ambiguous sentence. This is not the sort of knowledge an experimenter would like subjects to have, and for a very simple reason. If subjects are not aware of ambiguity, then whether or not they assign a single meaning to an ambiguous sentence probably reflects, in some sense, a basic, universal, or invariant method of sentence processing uncompli-

Requests for reprints should be sent to: Richard Reardon, Department of Psychology, University of Oklahoma, Norman, Oklahoma 73019.

cated or unobscured by special strategies. If subjects are aware of ambiguity, then basic processes may be distorted by such strategies. From the subjects' point of view, the experiment is now a peculiar game in which they must match wits with the experimenter to avoid commitment to one meaning of an ambiguous sentence when the other is correct.

Reardon and Katz (1978) demonstrated that the sentence-verification arrangement makes discovery of ambiguity almost certain. After simulating a typical sentence-verification experiment, the authors had subjects fill out a postexperimental questionnaire of the "funnel" type recommended by Page (1973). All subjects admitted being aware that some sentences were ambiguous. Almost 50% reported that their discovery occurred during the first quarter of the experiment, and 85%, during the first half.

It would be helpful to have corroborative evidence from the latency and error data that subjects have in fact used knowledge of ambiguity to behave in a situation-specific way. Let us focus on how we might perform additional analyses of latency and error data, given that subjects may be aware.

If subjects do become aware of ambiguity, it is only after the experiment has begun, usually during the earlier stages of the experiment. Once aware, they must now learn how to make decisions differently, although they must continue to respond as quickly and as accurately as possible. As Pachella (1974) pointed out, there is a delicate relationship between speed and accuracy in reaction time experiments of this sort. On the one hand, too many errors may lead to an increase in latency, which results in subjects' gaining time to reduce the number of errors. On the other hand, once the number of errors has been reduced, latency decreases, because long latencies are as unacceptable as high error rates. Given this logic, consider the latency data. When, early in this experiment, subjects discover ambiguity, their error rates should be high for ambiguous-unexpected sentence pairs, because it is only on this type that they are likely to be incorrect. High error rates are soon accompanied by long latencies; thus, we would also expect longer latencies for ambiguous-unexpected pairs. Because other pairs do not have high error rates, their latencies ought to be significantly shorter. As the experiment progresses, however, subjects should develop effective strategies for avoiding incorrect responses to ambiguous-unexpected pairs. This, in turn, should permit a reduction in latency for this type of sentence pair. Again, all other pairs ought to have low, stable latencies throughout the experiment.

Because of the close relationship between latency and errors, we have, in describing the pattern of the latency data over time, described the pattern of the error data as well. Briefly, there should be a high number of errors initially for the ambiguous-unexpected sentence pairs, with a reduction over the course of the experiment. Other sentence-pair types should produce low, stable error rates throughout.

What, in brief, is the difference between a basic-process hypothesis and a conscious-strategy (i.e., aware-subject) hypothesis? Simply, the former predicts stability in the pattern of latencies and errors for ambiguous-unexpected sentence pairs over time; the latter predicts changes in these patterns. The present experiment was conducted to test the two predictions.

## METHOD

### Materials

The determination of the unexpected versus the expected meaning of an ambiguous sentence can be problematic. Obviously, subjects cannot be exposed to the stimulus sentences before the experiment. Pretesting the sentences with a different population is one possibility, but there is no guarantee that what is unexpected for one subject is unexpected for another. Foss et al. (1968) suggested asking subjects after the experiment to indicate the expectancy of each ambiguous sentence. In the present study, both methods were used: the pretest method to ensure a reasonable a priori distribution of pair types, and the posttest method to fix the expectancy of each sentence, for each subject, for analyses. Stimulus lists were constructed as follows.

Twenty lexically ambiguous sentences were selected from a set of 150. These 20 sentences had been found to have expectancy values of .65 or better. Each was paired with a second sentence that disambiguated it. For 10 of these, the second sentence represented an unexpected meaning. In addition, there were 20 unambiguous sentences, each paired with a second sentence that did or did not "follow from" it. Thus, there were 40 sentence pairs in all: 10 ambiguous-expected (AE) pairs, 10 ambiguous-unexpected (AU) pairs, 15 unambiguous-expected (UE) pairs, and 5 unambiguous-unexpected (UU) pairs. The relatively large number of UE pairs, according to Cairns (1973), makes more comparable the number of compatible ("follows from") and incompatible ("does not follow from") pairs. The second sentence of each pair in all cases agreed with the first in person, gender, number, and other relevant grammatical features.

The various pair types were distributed randomly throughout the list. To minimize possible order or sequence effects, partial counterbalancing was performed as follows. There were two list orders, a forward and a backward. Within each order, the expectancy for each ambiguous sentence was reversed, so that those ambiguous sentences that had been followed by an expected second sentence were now followed by an unexpected sentence, and vice versa. Thus, there were four presentation lists in all. Seven subjects were assigned at random to each list.

### Subjects and Procedure

The subjects were 28 undergraduates enrolled in introductory psychology courses.

Each subject sat at a table in a darkened room and faced a movie screen 1.40 m away. In front of him/her was a small panel with two switches, one marked "yes" and the other "no." A carousel slide projector was located above and behind the subject.

The subjects were informed that the experiment involved "sentence recognition." They were told that they would be reading a number of pairs of sentences. They were instructed to read the first sentence of each pair; when the second sentence appeared, they were to decide if it did or did not "follow from" the first. If it did follow, they were to pull the switch marked "yes"; if it did not, they were to pull the "no" switch.

The first sentence of each pair was projected for approximately 3.5 sec, in capital letters 1.5 cm tall. Exposure time was controlled with a mechanical shutter (Lafayette 43016) placed over the projector lens. The second sentence appeared 1.5 sec after the first was shown. With the onset of the second sentence,

a digital timer (Lafayette 54016) was automatically started. As soon as the subject pulled the "yes" or "no" switch, the timer stopped and the shutter closed. The latency and "yes" or "no" responses were recorded, after which the experimenter went to the next pair. This procedure was repeated for each of the 40 sentence pairs.

After completing the sentence pairs, the subjects were asked to fill out the Reardon and Katz (1978) postexperimental funnel-type questionnaire. Following this, in the manner recommended by Foss et al. (1968), the expectancies for the ambiguous sentences were determined. The subjects were given a list of all of the ambiguous sentences; under each sentence were two paraphrases, one expressing the meaning of one version of the sentence and the other expressing the meaning of the other version. The subjects were asked to circle the paraphrase that represented the first meaning that came to mind when they read the ambiguous sentence; the circled paraphrase was then taken as the expression of the expected meaning and the uncircled paraphrase as the unexpected meaning.

**RESULTS AND DISCUSSION**

The mean latencies, collapsed over trial blocks, for each pair type gave the following results: AE = 2.377 sec; AU = 2.605 sec; UE = 2.333 sec; and UU = 2.358 sec. The error rate for AU pairs was extremely high, approximately 50%, whereas the error rates for the other pair types were between 4% and 11%. These are typical results in sentence-verification experiments (e.g., Cairns, 1973) and indicate that the present study was a successful replication. Every sentence pair was assigned to one of three trial blocks according to where it appeared in the presentation list: first third, second third, or last third. The first three sentence pairs were eliminated from the analysis to minimize possible warm-up effects (this included one UU, one UE, and one ambiguous pair). An ambiguous sentence pair was assigned to the AE or AU conditions on the basis of the response to the relevant item on the expectancy questionnaire given at the conclusion of the experiment (as described in the Method section). The critical comparisons involve these two types. Preliminary inspection and analyses of the latencies and errors for all sentence-pair types, excluding AU pairs, indicated no differences among AU, UU, and UE pairs. The pattern of mean latencies for AU and AE pairs across trial blocks is given in Table 1.

Typically, a 2 x 3 analysis of variance (type x trial block), with repeated measures on both independent variables, would be carried out. However, McCall and Appelbaum (1973) cautioned that repeated-measures designs in which one or more of the independent variables have more than two levels are "at risk" for violation of the homogeneity-of-covariance assumption. Violation of this assumption can produce artificially large F values. Therefore, our data were analyzed using the conservative multivariate-analysis-of-variance approach to repeated-measures designs (see McCall & Appelbaum, 1973, for details of this technique). There was a main effect for pair type: AU latencies were significantly longer than AE latencies [ $F(1,27) = 10.83, p < .003$ ]. The main effect for trial block failed to reach significance. The pair type x trial block interaction

**Table 1**  
Mean Latencies (in Seconds) to Judge Compatibility of Sentence Pairs: Ambiguous-Unexpected and -Expected x Trial Block

	Trial Block		
	Block 1	Block 2	Block 3
Ambiguous Unexpected	2.722 (.53)	2.731 (.45)	2.450 (.37)
Ambiguous Expected	2.432	2.316	2.362

*Note*—Numbers in parentheses indicate error rates for AU pairs across trial blocks.

was marginally significant [ $F(2,26) = 3.20, p < .06$ ]. As Table 1 shows, latencies for the AE sentence-pair types were uniformly low across trial blocks, whereas the latencies were higher for the AU type in Trial Blocks 1 and 2, but not in Trial Block 3. Planned comparisons revealed the following: There were significant differences between the AU type and AE type for Trial Block 1 [ $t(81) = 6.75, p < .01$ ] and Trial Block 2 [ $t(81) = 6.85, p < .01$ ] (Tukey critical value:  $q = 4.99$  for  $p < .01$ ). There was, however, no significant difference at Trial Block 3. These analyses point to one crucial result: The latencies for the AU sentence-pair type dropped substantially in Trial Block 3.

The error data are also summarized in Table 1. There was a significant decrease in the error rate for the AU sentence-pair type [ $\chi^2(1) = 4.061, p < .05$ ]. Moreover, and perhaps of greater importance, the decline in error rate occurred before the decline in latencies. As we have already reported, the number of errors to other sentence-pair types was small (between 4% and 11%), and these errors were not distributed in any systematic fashion over time.

The results of the postexperimental questionnaire on awareness replicate the findings of Reardon and Katz (1978). As in the earlier study, nearly all the subjects (27 of 28) admitted to being aware of ambiguity, and these aware subjects tended to report early discovery.

The data of the present study support a conscious-strategy hypothesis. Latencies for the crucial AU sentence pairs decreased on Trial Block 3 to a level statistically indistinguishable from the latencies for AE pairs, the latter remaining stable throughout the experiment. Errors for the AU sentence pairs also decreased, and began to do so even sooner than the latencies. A basic-process hypothesis cannot explain these findings, for it predicts neither a decrease in latencies nor a decrease in errors for any sentence-pair type. On the other hand, a strategy hypothesis makes just such a prediction. Subjects are aware of ambiguity, and therefore of the possibility and actuality of error. They make errors at the outset associated specifically with the AU sentence-pair type, but because such errors are unacceptable, they must act cautiously—hence, the high latencies. As the experiment proceeds, subjects, by virtue of practice, make fewer errors—hence, the decline in errors during

Trial Block 2. According to the tradeoff law, once subjects have reduced the number of errors, they begin to respond more quickly, for the experiment is, after all, a speed test—hence, the reduction in latencies during Trial Block 3.

The awareness hypothesis works well in this situation, but it is not without difficulties. The subjects were still making a relatively high number of errors with the AU sentence type at the end of the experiment, even though the latencies were statistically comparable to the latencies for other sentence-pair types. If the tradeoff principle had been working perfectly, the errors should have been comparable among pair types if the latencies were comparable. A reasonable explanation is that the simple *perception* of increased accuracy by subjects is itself sufficient for them to feel free to respond more quickly.

Overall, the results of the present study remind us that human subjects are not passive responders, but are active, creative, learning individuals.

#### REFERENCE NOTE

1. Cairns, H. S. *Ambiguous sentence processing*. Paper presented

at the meeting of the Midwestern Psychological Association, Detroit, May, 1971.

#### REFERENCES

- CAIRNS, H. S. Effects of bias on processing and reprocessing of lexically ambiguous sentences. *Journal of Experimental Psychology*, 1973, **97**, 337-343.
- FOSS, D. J., BEVER, T. G., & SILVER, M. The comprehension and verification of ambiguous sentences. *Perception & Psychophysics*, 1968, **4**, 304-306.
- MCCALL, R. B., & APPELBAUM, M. I. Bias in the analysis of repeated-measures designs: Some alternative approaches. *Child Development*, 1973, **44**, 401-415.
- MISTLER-LACHMAN, J. L. Levels of comprehension in processing of normal and ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 614-623.
- PACHELLA, R. C. The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N. J: Erlbaum, 1974.
- PAGE, M. M. On detecting demand awareness by postexperimental questionnaire. *Journal of Social Psychology*, 1973, **91**, 305-323.
- REARDON, R., & KATZ, S. Assessing awareness of ambiguity in the sentence-verification paradigm. *American Journal of Psychology*, 1978, **91**, 501-508.

(Manuscript received for publication May 21, 1983.)