Journal of Memory and Language **41,** 457–468 (1999)
Article ID jmla.1999.2656, available online at http://www.idealibrary.com on IDEAL®

Sense Creation in and out of Discourse Contexts

Richard J. Gerrig and Heather Bortfeld

State University of New York, Stony Brook

What does *Evian boy* mean? Researchers have offered a variety of theories to predict readers' interpretations of such noun–noun combinations outside discourse contexts. But in what ways do these out of discourse context interpretations contribute to the determination of discourse meanings? We define two views: the *interdependence view* suggests that the time course of meaning recovery in discourse contexts is affected by out of discourse context interpretations; the *independence view* does not posit such a relationship. Experiments 1 and 2 established pairs of noun–noun combinations (such as *doll smile* and *baseball smile*) that differed reliably with respect to accessibility of meanings out of discourse contexts. Experiments 3 and 4 demonstrated that noun–noun combinations with highly accessible out of discourse context meanings did not experience interference when they were endowed with different, innovative meanings in discourse contexts. We suggest that the pattern of data is inconsistent with the interdependence view. © 1999 Academic Press

Key Words: comprehension; discourse context; meaning accessibility; noun–noun combinations; sense creation.

Consider this sentence from a magazine article (Callahan, 1996, p. 25): "The Evian boy raised his hand and leaned forward, mouth agape, searching for the right words." What does *Evian boy* mean? In the absence of discourse context, it is possible to imagine a range of meanings: The boy who is drinking Evian, the boy who works for the Evian company, the boy who is wearing a t-shirt adorned with an Evian bottle, and so on. However, in the original context, the meaning of *Evian boy* was quite specific. The first paragraph of the article, which described a talk by Gena Rowlands at the American Academy of Dramatic Arts, included the critical information (pp. 20, 25, 26):

> The auditorium on Lower Madison Avenue was packed with acting students. Onstage, Gena Rowlands was struggling to break the seal on a bottle of Evian when a scrappy, eager young man bounded from his front-row seat and proudly cracked it for her. Without uttering a word, she held his gaze for a moment and smiled serenely. He lurched back to his seat, clearly flushed.

Five paragraphs of text intervened followed by the sentence including the use of *Evian boy*, which then took on (approximately) the meaning "the boy who had helped Rowlands with her Evian bottle."

The purpose of the current article is to contrast circumstances in which readers attempt to determine the meaning of noun–noun combinations, like *Evian boy*, with and without support from surrounding discourse context. Although languages provide a wide range of expressions that allow for *sense creation*—instances in which language producers encapsulate new meanings (Clark, 1983)—a great degree of attention has been focused on noun–noun combinations (e.g., Coolen, Van Jaarsveld, & Schreuder, 1991, 1993; Gagné & Shoben, 1997; Gerrig & Murphy, 1992; Murphy, 1990; Wisniewski, 1996; Wisniewski & Love, 1998). This disproportionate attention stems, in large part, from the fact that noun–noun combinations afford researchers the opportunity to study *conceptual combination*—the processes through which the independent meanings of different concepts interact to yield innovative combinations. Many of the theories of conceptual com-

We thank Talia Ben-Zeev and Arthur Samuel for insightful comments on earlier versions of the manuscript. We also thank Norbert Gershkovich for assistance in data collection.

Please address all correspondence and reprint requests to Richard J. Gerrig, Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794-2500. E-mail: rgerrig@psych1.psy.sunysb.edu.

0749-596X/99 $30.00
Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

bination have been tested with methodologies akin to the challenge we posed earlier with the question "What does *Evian boy* mean?" That is, experimental participants have been asked to provide interpretations of noun–noun combinations outside of discourse contexts.

These tests leave open the question of whether the outcomes they reveal will generalize to circumstances in which speakers or writers use such combinations in richer discourses. Some theorists have argued explicitly that an understanding of the processes that give rise to interpretations outside of discourse contexts must logically precede, and will subsequently contribute to, accounts of the interpretations of noun–noun combinations in discourse: "It makes sense to first identify how the meanings of the constituents (i.e., prior knowledge) affect interpretation. Then the role of discourse setting may be more meaningfully understood in light of these prior knowledge effects" (Wisniewski, 1996, pp. 450, 451). In this article, we challenge the generality of the claim that the meanings readers generate for noun–noun combinations outside discourse contexts affect the recovery of meanings within discourse contexts.

Consider the pair of combinations *doll smile* and *baseball smile.* In the absence of discourse context, most informants find it considerably easier to provide an interpretation for *doll smile* (as confirmed in Experiments 1 and 2). That is, the accessibility of a meaning for *doll smile* in the absence of discourse context is relatively high compared to the accessibility of a meaning for *baseball smile.* Consider, however, the use of the combinations in the concluding sentence of this story:

> Aunt Bev had just come back from visiting her niece and nephew. She was telling her friend Dora about the visit. Bev explained, "I brought Sarah a doll and Paul a baseball."
>
> Dora asked, "Did the children like their gifts?"
>
> Bev replied, "Their mother coached them to give me their biggest smiles." She continued, "I saw a doll smile (*or* baseball smile) that was very impressive."

This story seemingly equates the accessibility of the meanings for the two noun–noun combinations. Our experiments followed the same pattern as this example. That is, we identified pairs of noun–noun combinations that differed reliably in the accessibility of meanings outside of discourse contexts. We then created discourse contexts that we believed would vitiate those differences. Our goal was to demonstrate concrete instances in which accessible out of discourse context meanings for noun–noun combinations do not figure in the recovery of discourse meanings.

Our critical comparisons involved the time course with which readers can recover meanings. Suppose, for example, that individuals could think of a meaning more quickly for the phrase *doll smile* than for *baseball smile* outside discourse contexts. What potential impact could this difference have for processing of the compounds within discourse contexts? One possibility is that highly accessible out of discourse context meanings would impair readers' ability to recover innovative meanings in discourse contexts. Consider experiments that examined readers' ability to recover new meanings based on noun–noun combinations, such as *door man* and *cave man,* that already existed in the mental lexicon (Gerrig, 1989). Reading times were compared for discourse contexts in which the compounds were used with their conventional meanings versus contexts in which they were endowed with innovative meanings (e.g., the *door man* and *cave man* were types of researchers being interviewed for faculty positions by a Fine Arts department). The experiments revealed that those noun–noun combinations that were easier to understand in their conventional uses were more difficult to understand in their innovative uses. That is, the ready accessibility of the conventional meaning appeared to interfere with readers' ability to construct new meanings.

We can extend the logic of this earlier finding to novel uses of noun–noun combinations for which out of discourse context meanings are not prestored as conventional meanings but nonetheless are relatively high in accessibility. In the same way the conventional meaning of a compound such as *door man* can interfere with readers' ability to construct an innovative meaning in discourse context, a highly accessible out of discourse meaning for *doll smile*

could interfere with readers' ability to construct a novel meaning. We call this prediction the *interdependence view* because it arises from an interdependence of meaning recovery in and out of discourse contexts. This interdependence view, as we have defined it, includes the assumption that out of discourse context meanings become sufficiently represented in the course of processing to bring about interference. By contrast, we suggest that the two types of meaning recovery do not necessarily depend on one another. We call this the *independence view* and predict from it no differential interference when noun–noun combinations differing in out of discourse context meaning accessibility are used in discourse contexts. In four experiments, we evaluated the interdependence and independence views.

EXPERIMENT 1

To begin our program of research, we needed pairs of noun–noun combinations that differed reliably in the accessibility of meanings out of discourse contexts. However, we also needed pairs that, for later experiments, we could embed in discourse contexts that we believed would render innovative meanings equally accessible. We generated 30 pairs that, by our intuitions, met these criteria. The goal of Experiment 1 was to identify 20 pairs, via participants' ratings, to pass along to later experiments.

Methods

Participants. Sixty-six undergraduates from the State University of New York at Stony Brook participated. All the students were native speakers of English.

Materials and design. We generated 30 pairs of noun–noun combinations that we intended to differ in the accessibility of meanings outside discourse contexts. (We were also able to anticipate how the members of each pair could function in an equivalent fashion in brief discourses.) The Appendix provides 20 of the 30 pairs (those pairs that we used for subsequent experiments) as well as a gloss of the stories in which they later functioned.[1] (The complete set of stories is available from the first author.)

One noun–noun combination from each pair was distributed to two questionnaires. Each questionnaire included 15 noun–noun combinations that we believed, a priori, would be relatively high on out of discourse context meaning accessibility and 15 that we believed would be relatively low. The combinations (i.e., the tokens from each pair) were presented in different random orders on the two questionnaires and the three pages of the questionnaire were randomized for each participant.

Procedure. Participants were asked to read each noun phrase and rate how difficult it was to come up with a meaning on a scale ranging from 1 ("very easy") to 7 ("very hard").

Results and Discussion

The purpose of this experiment was to confirm our intuitions that one member of each pair of noun–noun combinations had a meaning that was relatively more accessible outside discourse contexts. On average, our intuitions were confirmed. The mean difficulty rating for those noun–noun combinations we believed to have relatively more accessible meanings was 2.62; for those noun–noun combinations we expected to be relatively less accessible the mean rating was 3.58. This difference was reliable with both participants (*F1*) and items (*F2*) treated as random variables ($F1(1,64) = 106.933$, $MSe = 0.263$, $p < .001$; $F2(1, 28) = 33.147$, $MSe = 0.412$, $p < .001$). Despite this overall support for our intuitions, we were correct on only 25 of 30 pairs (i.e., for 5 pairs, participants gave "easier" ratings to the token we anticipated would be relatively low on meaning accessibility). Accordingly, we chose the best 20 pairs—those that had shown the largest differences in the expected direction—for Experiments 2 to 4. The smallest difference among the 20 pairs we retained was 0.75 ratings points. The mean rat-

[1] We had intended that each of the pairs have the same head noun with differing modifier nouns (e.g., *tax lunch* vs. *family lunch*). However, we made an editing error and rendered one pair as *martini lemon* and *martini cantaloupe.* Because the error did not affect the logic of our predictions, it was preserved for subsequent experiments.

ings for this sample of 20 pairs was 2.24 for highly accessible tokens and 3.70 for less accessible tokens.

EXPERIMENTS 2A AND 2B

Because the interdependence and independence make different predictions about moment-by-moment processing, we next needed to confirm that readers could recover meanings more swiftly for the highly accessible token than for the less accessible token from each pair. In Experiment 2a, we asked participants to push a response key to indicate when they had thought of a meaning for each noun–noun combination. In Experiment 2b, we used a task more closely patterned on previous research. Participants were asked to indicate as quickly as possible whether they found a phrase meaningful (Murphy, 1990).

Methods

Participants. Twenty students participated in each version of the experiment. All were native speakers of English.

Apparatus. The experiment was run on two computers that recorded responses and response times. Participants were seated in front of a MicroScan color monitor with their hands resting on the keyboard. They used buttons on the keyboard to respond. The phrases were displayed in the center of the screen in standard upper- and lowercase type.

Materials and design. The 20 pairs from Experiment 1 were randomly distributed to two lists so that each list had 10 highly accessible and 10 less accessible tokens. Five noun–noun combinations eliminated after Experiment 1 were used as practice stimuli.

Experiment 2b included 20 filler noun–noun combinations. We intended 10 of these phrases to be clearly meaningful (e.g., *funeral committee, hamster snack*) and 10 to be clearly nonsensical (e.g., *error oven, hammer ghetto*). We included these fillers so that participants would have easy opportunities to judge combinations as meaningful or not.

Procedure. In Experiment 2a, we asked participants to read the noun–noun combinations one at a time on the computer screen and indicate, by pressing the *L* key on the keyboard, when they had thought of a meaning for the phrase. They used the spacebar to advance to each new phrase.

In Experiment 2b, we asked participants to push a key labeled "yes" to indicate "yes, I can think of a meaning" and to push a key labeled "no" to indicate "no, I don't think the phrase is meaningful." The *A* and *L* keys were labeled "yes" and "no" in a fashion counterbalanced across participants. As in Experiment 2a, participants advanced to each new noun–noun combination by pressing the spacebar.

Results and Discussion

For each of these experiments, as well as for the subsequent experiments, we considered any response times more than 3 *SD* beyond each individual's mean response time to be outliers and eliminated those observations from further analyses. For Experiment 2a, this process yielded a loss of 1.75% of the data. As we had predicted, participants responded considerably more quickly when asked to think of a meaning for the highly accessible tokens from each pair (mean response time = 5.611 s) versus the less accessible tokens (mean response time = 8.541 s). This difference was reliable ($F1(1,18) = 23.068$, $MSe = 3619732$, $p < .001$; $F2(1,18) = 24.210$, $MSe = 2942574$, $p < .001$).

Although the results from Experiment 2a confirmed our prediction, we wished to replicate the high accessibility versus relatively low accessibility difference with a task that had been used previously in the literature, which is why we turned to the "is it meaningful?" task. For Experiment 2b, we again removed 1.75% of the data as outliers. Again, participants made responses more swiftly for highly accessible tokens (mean = 1.918 s) than for less accessible tokens (mean = 2.437 s), a difference that is reliable ($F1(1,18) = 12.042$, $MSe = 212880$, $p = .003$; $F2(1,18) = 9.328$, $MSe = 2824050$, $p = .007$). We also compared the rates at which participants judged the phrases to be meaningful. Participants responded "yes" for 81.2% of the highly accessible tokens versus only 43.4% of the less accessible tokens ($F1(1,18) = 88.676$, $MSe = 165.422$, $p < .001$; $F2(1,18) =$

80.806, $MSe = 175.004$, $p < .001$). Note that the response times in Experiment 2b were considerably shorter than those in Experiment 2a. This suggests, as we might expect, that it takes less time to understand a phrase than to prepare (even mentally) a paraphrase. In any case, the response times from both Experiments 2a and 2b as well as the pattern of "yes" and "no" responses for Experiment 2b strongly suggest that meanings were more available for the highly accessible tokens. In Experiments 3 and 4, we demonstrate that the less accessible noun–noun combinations—which readers often judged to be meaningless in Experiment 2b—become quite easy to understand in discourse contexts.

EXPERIMENT 3

The purpose of the previous experiments was to define two groups of noun–noun combinations—one for which meanings were highly accessible out of discourse contexts and one for which meanings were relatively inaccessible. In Experiment 3, we use these two groups of noun–noun combinations to test the predictions of the *interdependence* and *independence* views of meaning recovery. Specifically, we embedded the matched pairs of highly and less accessible compounds (e.g., *doll smile* and *baseball smile*) in discourse contexts that allowed the compounds to function in a comparable fashion (e.g., smiles on the face of children receiving gifts). The interdependence view suggests that the accessibility of out of discourse contexts meanings should interfere with discourse meanings; the independence view does not anticipate such interference.

In this experiment, participants read stories that ended with either the highly or the less accessible compounds. Half of the time the compounds appeared in discourse contexts that we expected to impart innovative meanings. For the other half of the items, participants read the noun–noun combinations in neutral contexts—contexts that we did not expect to establish unambiguous innovative meanings. In the neutral contexts, we anticipated that the out of context difference between the highly and less accessible compounds would continue to emerge. Thus, the neutral contexts provide a discourse replication of the earlier experiments as well as a baseline difference for comparison to the innovative contexts.

Methods

Participants. Twenty-eight students participated in this experiment. All were native speakers of English.

Apparatus. The same equipment was used as in Experiment 2.

Materials and design. For each pair of noun–noun combinations, we wrote a story that created comparable innovative meanings. The stories mentioned both the head nouns and modifier nouns explicitly. For half of the items, information relevant to the highly accessible compound appeared closer to the end of the story; for the other half the less accessible compound information was closer. The stories ranged from six to eight lines in length. The noun–noun combinations always appeared in the final line of the story. Table 1 presents sample stories. (The Appendix provides the full list of noun–noun combination pairs and glosses for their innovative uses.) We changed each story to create neutral versions. The neutral versions replaced the lines of text that provided the basis for the innovative relationship. For each story we wrote a true/false comprehension question that was the same across innovative and neutral versions. We also wrote five practice stories.

Using a Latin Square design, we constructed four lists of stories so that each story appeared in a different version on each list. Participants each read one list of 20 stories in random order.

Procedure. Participants read the stories on a computer screen, one line at a time. They pressed the *A* key (labeled "NEXT") after they had read and understood each sentence. At the end of the story, a beep sounded from the computer. The phrase "TRUE or FALSE?" appeared on the screen two lines above the comprehension question. Participants used the *J* and *K* keys (labeled "TRUE" and "FALSE") to agree or disagree with a statement about the story.

Results and Discussion

We used the same procedure as in Experiment 2 to remove outliers, which resulted in a loss of 0.71% of the data. Table 2 presents mean response times. Recall that the interdependence view suggests that accessible out of context meanings should interfere with innovative meanings. Thus, this view would predict a reversal of what is hard and what is easy from the neutral to the innovative uses of the compounds. By contrast, the independence view suggests that the comprehension times for the innovative uses will not show a pattern of interference. As revealed in Table 2, the data are more consistent with the independence view. The difference in comprehension time for highly and less accessible compounds is 597 ms in neutral contexts (this difference was 519 ms in Experiment 2b). The difference in the innovative contexts was 96 ms in the same direction, which is inconsistent with the interdependence view's prediction of interference (both F's < 1.5 for the 96-ms difference). This interaction of context (neutral vs. innovative) and out of context accessibility (highly vs. less) was reliable ($F1(1,24) = 8.016$, $MSe = 228344$, $p = .009$; $F2(1,16) = 14.721$, $MSe = 87932$, $p = .001$). The main effects of both context ($F1(1,24) = 23.557$, $MSe = 178553$, $p < .001$; $F2(1,16) = 14.358$, $MSe = 213123$, $p = .002$) and accessibility ($F1(1,24) = 10.663$, $MSe = 325669$, $p = .003$; $F2(1,16) = 6.586$, $MSe = 373871$, $p = .021$) were also reliable. Table 2 also provides error rates for the comprehension questions. There were no reliable differences in error rates as a function of the preceding story (all F's < 1.0).

We interpret our results to be consistent with the independence view. However, our conclu-

TABLE 1

Sample Stories

1. *ghost doctor* vs *artichoke doctor*
   - Neutral context (Experiment 3)
     Rahul and Betty were getting ready to go to a party at their hospital. They were both doctors in the pediatric unit. Rahul and Betty arrived at the party a little late, but they still had plenty of time to mingle. They were both up for an award from the hospital as "best staff member." At the end of the evening, the hospital administrator announced the prize. He said, "We all agree that the winner is the ghost/artichoke doctor."
   - Innovative context (Experiments 3 and 4)
     Rahul and Betty were getting ready for their hospital Halloween Party. They were both doctors in the pediatric unit. Betty had decided to go as a ghost and Rahul was dressed as an artichoke. Both costumes were fairly elaborate. They were both in the running for the award as "best costume." At the end of the party, the hospital administrator announced the prize. He said, "We all agree that the winner is the ghost/artichoke doctor."
   - Comprehension question (Experiment 3): Rahul and Betty were going to a party at a bowling alley.
   - Paraphrase judgment (Experiment 4): The administrator said the doctor in the ghost/artichoke costume won.
2. *family lunch* vs *tax lunch*
   - Neutral context (Experiment 3)
     Each day Maggie ate lunch with her friends. They all sat together and talked about concerns in their lives. Generally, they agreed on the topic ahead of time. Sometimes Maggie wasn't always sure the conversations were worthwhile. She thought back over the topics they'd covered over the last few days. Maggie found the family/tax lunch to be very helpful.
   - Innovative context (Experiments 3 and 4)
     Each day Maggie ate lunch with her friends. They all sat together and talked about concerns in their lives. Generally, they agreed on the topic ahead of time. One day, they discussed how to avoid paying extra income tax. Another day, they discussed how to deal with family problems. Maggie found the family/tax lunch to be very helpful.
   - Comprehension question (Experiment 3): Maggie discussed problems at lunchtime.
   - Paraphrase judgment (Experiment 4): Maggie was helped by the lunchtime discussion of family/tax issues.

TABLE 2

Results of Experiment 3: Reading Times for Noun–Noun Combination Sentences

	Discourse context	
	Neutral	Innovative
Highly accessible	3.745 (10)	3.612 (12)
Less accessible	4.342 (12)	3.708 (10)
Difference	.597 (2)	.096 (−2)

Note. The figures in parentheses are percentages of errors for the comprehension questions.

sion relies on the assumption that the innovative meanings our stories provide for the noun–noun combinations are different from those meanings at which participants would arrive out of discourse contexts. For example, if it were the case that readers understood *doll smile* to mean "a smile on the face of a child receiving a doll as a gift" in the absence of discourse context, then we would predict facilitation, rather than interference—invalidating our test of the interdependence view.[2] To test for this possibility, we asked 20 participants, drawn from the same population as the other studies, to provide out of discourse context interpretations of the noun–noun combinations. Participants received one of two versions of a questionnaire with highly and less accessible tokens assigned to the two questionnaires in a counterbalanced fashion. The participants read these instructions:

> Below, you will read some noun phrases. Noun phrases that you have probably heard before include dog sled, road sign, and car window. Many of the phrases you will see in this experiment are probably novel—you may have never heard them before. Some examples of novel noun phrases might be factory fish, bottle frog, earthquake school, and so on. Your task is to write down a description of the most likely meaning of each novel phrase. Pretend that you have just heard the phrase during a conversation. What would be the meaning of the phrase that seems most natural to you?

One of the authors (R.J.G.) and a research assistant blind to the experimental hypotheses read through the participants' definitions. There were no disagreements about which meanings were consistent with the discourse meanings. Overall, of the participants' 400 attempts to provide meanings out of discourse contexts, 2% (eight instances) of the highly accessible meanings and 1.75% (seven instances) of the less accessible meanings were consistent with the discourse meanings established by our experimental stories. However, one pair was responsible for seven of these cases—five instances for the highly accessible token *television vacation* (e.g., "not watching TV for a while") and two instances for the less accessible token *computer vacation* (e.g., "turn off the computer"). We examined the data for this item to see if this relatively high rate of correspondence between in and out of discourse context meanings provided a pattern of facilitation as predicted by the interdependence view. In fact, reading times for the use of *television vacation* in the innovative context (mean = 3.418) were somewhat longer than those for *computer vacation* (mean = 3.062). If we omit this item from the overall analysis the mean difference for highly and less accessible tokens (3.622 vs. 3.742) grows even a bit further from the prediction of interference (cf. Table 2). Ideally, none of our noun–noun combinations would have elicited interpretations out of discourse context that were equivalent to our innovative meanings. Nonetheless, we believe that the rate of equivalence is sufficiently low and similar across the highly and less accessible tokens that our experiment still provides a meaningful contrast between the interdependence and independence views.

The results of this experiment are consistent with the independence view and at odds with the interdependence view. However, our only measure in this experiment was comprehension time. We did not, for example, ask our participants to indicate their agreement with paraphrase judgments as has been the practice in some previous research (e.g., Gerrig, 1989). The reason for this omission is that we weren't entirely sure what paraphrases to give for the neutral stories, particularly for the noun–noun combinations with relatively less accessible meanings out of context. We didn't want our readers to adopt a strategy of reading the noun–noun combination sentences cursorily in anticipation of a paraphrase. As a consequence, for Experiment 4 we opted to retain only the innovative stories and add paraphrase judgments to those stories.

EXPERIMENT 4

The goal of Experiment 4 was to ensure that readers had understood the innovative uses of the noun–noun combinations to have the meanings we intended. Accordingly, at the end of each story participants were asked to agree or disagree with a paraphrase of the sentence con-

[2] We thank an anonymous reviewer for suggesting this control.

taining the noun–noun combination. A priori, the interdependence view and independence views do not make contrasting predictions for paraphrase judgments. That is, if readers have fully understood the noun–noun combinations before continuing on to the paraphrase, we would expect paraphrase judgment times to be more or less equivalent however they had arrived at their initial "full" understanding.

Method

Participants. Twenty native speakers of English participated in this experiment.

Apparatus and materials. The same equipment was used as in earlier experiments. We retained the innovative stories from Experiment 3. Each story ended with a paraphrase sentence that was meant to capture the intended meanings of the noun–noun combinations (see Table 1 for examples). We wrote 20 filler stories, each of which ended with a sentence containing an innovative use of a noun–noun combination. We intended all the paraphrase sentences to be true for the experimental items and false for the filler items.

Design and procedure. There were two versions of each story (highly vs. less accessible). Using a Latin Square design, we constructed two lists so that each story appeared in a different version on each list. The 20 filler items were identical on the two lists. Participants each read one list of 40 stories in random order.

Participants read the stories on a computer screen, one line at a time. They pressed the *A* key (labeled "NEXT") after they had read and understood each sentence. At the end of the story, a beep sounded from the computer. The phrase "TRUE or FALSE?" appeared on the screen two lines above the comprehension question. Participants used the *J* and *K* keys (labeled "TRUE" and "FALSE") to agree or disagree with the paraphrase of the story's final sentence.

Results and Discussion

We used the same procedure as in earlier experiments to remove outliers, which resulted in a loss of 2.75% of the data. Table 3 presents mean reading times, paraphrase judgment times, and error rates for the paraphrase judgments. Let us begin with the reading times.

TABLE 3

Results of Experiment 4

	Reading Time	Paraphrase judgment time (error rate)
Highly accessible	2.802	3.878 (7.5)
Less accessible	3.035	4.050 (11.0)

Recall that the interdependence view suggests that, by virtue of interference, the reading times for the highly accessible tokens should be longer than the reading times for the less accessible tokens. As seen in the table, this is not the case. In fact, the sentences with the highly accessible tokens were read 233 ms more quickly ($F1(1,18) = 2.965$, $MSe = 210494$, $p = .102$; $F2(1,18) = 5.920$, $MSe = 82333$, $p = .026$). Although this result is inconsistent with the interdependence view as we have defined it, it does raise the unpleasant possibility that readers may not have been recovering our intended innovative meanings. That is, if readers were defaulting to the out of discourse context meanings, the sentences with highly accessible tokens should have been easier to understand than those with less accessible tokens.

Fortunately, the paraphrase judgments allow us to assess that possibility. If readers were recovering the *out* of context meanings quickly for the highly accessible tokens, we would expect paraphrase judgment times to be lengthy and the error rate to be high (i.e., we would predict interference). As seen in the table, the data do not fit that pattern. Instead, participants found it somewhat easier to verify the paraphrases for the highly accessible tokens (a difference of 172 ms) and were somewhat less error prone (a difference of 3.5% errors). Neither of these differences, however, was reliable (paraphrase judgment times, $F1(1,18) = 1.058$, $MSe = 350153$, $p > .10$; $F2(1,18) = 0.338$, $MSe = 274030$, $p > .10$; error rates, $F1(1,18) = 1.785$, $MSe = 101.4$, $p > .10$; $F2(1,18) = 1.638$, $MSe = 111.4$, $p > .10$).[3]

[3] As for Experiment 3, we examined the data separately for the *television vacation* (highly accessible token)–*com-*

Although we believe the data argue against the interdependence view, we are still left with the uncomfortable result that readers found it somewhat easier to understand the innovative meanings based on the highly accessible noun–noun combinations in the discourse settings. This result raises the possibility that the stories themselves may have been biased in favor of the highly accessible combinations—the stories may have provided more contextual support for those readings. Such circumstances would invalidate this experiment as a fair test of the interdependence view. To rule out this possibility, we had 20 students from the same participant population perform a rating task. Each student read one version of each of the 20 experimental stories and rated how well each final sentence (i.e., the sentence containing the noun–noun combination) fit in the context of the preceding sentences on a scale ranging from 1 (*fits poorly in context*) to 7 (*fits excellently with context*) (cf. Gerrig, 1989). The mean ratings on this task were 5.10 for the highly accessible combinations and 5.16 for the less accessible combinations (both t's < 1). These data argue against an inherent bias in the story contexts.

To examine further the reading time advantage for the highly accessible combinations, we turned our attention to the words that made up the combinations. Specifically, we looked at the word frequency for the contrasting member of each combination (e.g., *doll* vs *baseball*) (Kučera & Francis, 1967). The mean frequencies for the unique words were 67.7 for highly accessible combinations and 73.4 for less accessible combinations. This difference was not reliable ($t < 1$). This analysis rules out another easy justification for the reading time difference.

We are left with no ready explanation for why readers found it somewhat easier to understand the highly accessible noun–noun combinations in the discourse settings. It is possible that some of the same attributes (e.g., aspects of their conceptual structure) that allow the highly accessible noun–noun combinations to assume meanings out of discourse contexts may serve them well in discourse contexts. We can call this a *fluidity view* because properties of the component nouns of noun–noun combinations allow fluid recovery of meanings in a variety of contexts. This view is quite different from the interdependence view, as we have defined it. Consider *spaghetti night.* In the Experiment 3 control study, 90% of the out of discourse context interpretations were some variation on "a night when you eat spaghetti" or "a night when you eat pasta." An interdependence view would predict that readers would arrive at this interpretation of *spaghetti night* on their way to understanding this noun–noun combination in discourse context. A fluidity view, by contrast, would suggest that properties of *spaghetti* and *night* make it easier for readers to recover meanings, whatever the context. Thus, the fluidity view, like the independence view, makes no reference to the *content* of out of discourse context interpretations of noun–noun combinations.

Whatever the merits of the fluidity view, the data strongly argue against the prediction that out of discourse context meanings interfere with readers' ability to recover innovative meanings. We defined the interdependence view with respect to this interference prediction because that seemed most strongly supported by conceptual and empirical precedents. That is, if out of discourse context meanings come forcefully to readers' minds, they ought to interfere with readers' ability to construct different meanings. With respect to the independence view, the best result for reading times would have been no difference between the innovative uses of the highly and less accessible noun–noun combinations. However, as we have noted, because we are not sure how recovery of out of discourse

puter vacation (less accessible token) pair. The mean reading times for the sentence containing the combinations were 2.532 (television vacation) and 2.699 (computer vacation). This difference is somewhat smaller than that reported in Table 3. The difference for paraphrase judgments was quite large: 3.403 (television vacation) and 4.615 (computer vacation). If we exclude this item the difference in paraphrase times decreases from 172 ms to 118 ms. Thus, when we exclude the item that was most biased in favor of the interdependence view our results actually become more consistent with the independence view.

context meanings could lead to facilitation, we believe these data are considerably more supportive of independence than interdependence.

GENERAL DISCUSSION

The purpose of these experiments was to contrast two views of the relationship between the meanings readers recover for noun–noun combinations in and out of discourse contexts. One view, which we have called the interdependence view, suggests that the out of discourse context meanings should affect the ease with which readers can recover meanings in discourse contexts. The second view, which we have called the independence view, posits no particular relationship between the two types of meaning. Experiments 1 and 2 provided us with pairs of noun–noun combinations that differed reliably with respect to accessibility of meanings out of discourse contexts. Experiments 3 and 4 demonstrated that noun–noun combinations with highly accessible out of discourse context meanings did not suffer from interference when they were endowed with different, innovative meanings in discourse contexts. We suggest that the pattern of data are inconsistent with the interdependence view.

Our experiments do not, of course, demonstrate that out of discourse context meanings *never* have an impact on discourse context meanings. In fact, we developed our prediction for the interdependence model from a series of experiments that demonstrated that conventional meanings can and do interfere with the recovery of innovative meanings (Gerrig, 1989). Recall that the member of pairs such as *door man* and *cave man* that was easier to understand when used conventionally proved harder to understand when used innovatively. This result arises, presumably, because the conventional meanings of noun–noun combinations such as *door man* and *cave man* are accessed automatically from the mental lexicon at the same time the innovative meanings are being computed. The present results suggest that the processes that give rise to out of discourse context meanings—even *highly* accessible out of discourse context meanings—are not automatically engaged in discourse contexts. Or, if the processes are automatically engaged, the products of these processes are too weakly represented in the course of processing to bring about interference.

Our experiments were designed to provide as fair a test as we could devise for the interdependence view. Out of discourse contexts, our pairs of noun–noun combinations differed quite consistently. In that light, it seems noteworthy that we found no evidence for interference in the final pair of experiments. We believe, therefore, that our experiments set a possible agenda for those researchers who wish to pursue the interdependence view: They could attempt to define the range of circumstances in which out of discourse context meanings affect the process of meaning recovery within discourse contexts and they could attempt to verify those predictions experimentally.

Consider research by Gagné and Shoben (1997) that demonstrated the importance of the relations typically encoded by the modifier noun in readers' understanding of noun–noun combinations. For example, participants found it easier to understand combinations like *chocolate bird* than those like *chocolate plant* because the first relation (a bird *made of* chocolate) functions more often for *chocolate* than does the second relation (a plant *makes* chocolate). We could imagine a series of experiments that would assess the extent to which the typicality of the relation interfered with innovative meanings in discourse contexts. The interdependence view would predict that readers would find it relatively more difficult to understand *chocolate bird* in an innovative context that used a low typicality relation rather than a high typicality relation. Experiments along these lines could further define circumstances in which interdependence versus independence hold.

However, for many ordinary cases of noun–noun combinations, we speculate that the independence of meanings will hold. To motivate this speculation, we return to the example of *Evian boy* in its discourse context. That example shares an important feature with our experimental stimuli: The "definition" of *Evian boy* is provided in the first paragraph; when *Evian boy*

is used much later, it points back to that original episode. Similarly, our experimental stimuli all provided explicit mention of the relationships that the noun–noun combinations would subsequently encode. In all these cases, the noun–noun combinations are functioning, at first, like deictic expressions: They point to entities and relationships that have already been introduced into the discourse representation. We suggest that the process of finding the referent for a deictic expression is a rather different, and probably less computationally complex, process than the one people carry out when they are asked to provide meanings for noun–noun combinations out of discourse contexts.

Another way to make this point is with respect to a contrast made by Gerrig and Murphy (1992) between *exemplar matching* and *concept formation* views of conceptual combination. The exemplar matching view makes very much the same claim we are making here, that readers link the noun–noun combination with an earlier discourse referent. The concept formation view, which Gerrig and Murphy accept, goes beyond that view to suggest that readers actively encode a new relation when they understand a combination. What we wish to add here is that, in discourse contexts, concept formation is parasitic on exemplar matching: Readers' understanding of noun–noun combinations may typically result in more than mere referential linking, but exemplar matching is the way in which the process starts. By contrast, the interpretation of noun–noun combinations out of discourse context does not involve exemplar matching—so there is no head start for concept formation. Thus, we speculate that the independence view is likely to hold widely because we believe that the processes required to recover meanings for noun–noun combinations in many or most discourse contexts are not the same processes required to recover meanings for the same noun–noun combinations out of discourse contexts.

A corollary to this claim is that the meanings people recover will often be different in and out of discourse contexts. In some contexts, readers might be satisfied with exemplar matching: noun–noun combinations might do little more than mark referential continuity between a character who carried out activities early in a discourse and his or her subsequent activities (or, more generally, between a concept introduced into a discourse—such as "the smiles that children produce when they receive gifts"—and a later reference to that concept). In other contexts, readers will produce more elaborate meanings by establishing a new conceptual relation that transcends the original discourse context (Gerrig & Murphy, 1992). The relation encoded in *doll smile,* for example, seems less ephemeral than the one encoded in *Evian boy;* we might expect that readers would construct a more elaborate conceptual representation for *doll smile.* Out of discourse contexts—at least in experiments—readers are *required* to give full conceptual representations. An important empirical question, therefore, is under what circumstances the meanings readers represent in discourse contexts approach the level of concreteness and specificity of those they provide out of discourse contexts. Our program of research does not answer that question, but our data suggest why the question is compelling.

APPENDIX:

NOUN–NOUN COMBINATION PAIRS AND GLOSSES FOR INNOVATIVE USES

Each pair is given with the highly accessible noun–noun combination first.

1. *doll smile* vs. *baseball smile*

Aunt Bev's gifts to her niece and nephew elicited smiles. Bev says, "I saw a doll/baseball smile that was very impressive."

2. *ghost doctor* vs. *artichoke doctor*

Rahul and Betty, both doctors in a pediatric unit, are competing in a "best costume" contest at a Halloween party. The hospital administrator says, "We all agree that the winner is the ghost/artichoke doctor."

3. *family lunch* vs. *tax lunch*

Each day Maggie and her friends have lunchtime conversations about various topics. Maggie found the family/tax lunch to be very helpful.

4. *zebra horse* vs. *ostrich horse*

The horses at the circus have saddles on them with pictures of other animals. Susie says, "I think the zebra/ostrich horse is the most beautiful."

5. *martini lemons* vs. *martini cantaloupes*

A "party planning" magazine suggests serving martinis in scooped out fruit. Maggie says, "I want to try to find some good martini lemons/cantaloupes."

6. *igloo man* vs. *candle man*
The anthropology department is debating who they should hire as a new assistant professor. The chair says, "I really think we ought to hire the igloo/candle man."
7. *radio concerts* vs. *magazine concerts*
Pat and Dale advise Bonnie on how she can stave off boredom at symphony concerts. Bonnie is asked whether she must attend many more concerts and she replies, "Yes, but maybe now I'll make them radio/magazine concerts."
8. *spaghetti night* vs. *fork night*
Jenny is teaching her son Carlos new skills. She asks her husband, "When do you want me to plan the spaghetti/fork night?"
9. *sun canary* vs. *fog canary*
Tim is contemplating buying canaries whose singing depends on the weather. Tim finally makes up his mind to buy a sun/fog canary.
10. *mosquito screen* vs. *toaster screen*
Alex is installing a new screensaver on his computer. Alex decides he likes the mosquito/toaster screen best.
11. *baby couch* vs. *aunt couch*
Jim's baby brother and aunt each nap on a different couch in the living room. Jim advises his friend, "The baby/aunt couch is more comfortable."
12. *clown bike* vs. *kitten bike*
Lisa and Marie are decorating their bikes with stickers. Their friend Michelle comes over and says, "I wish I could own a clown/kitten bike."
13. *water sandals* vs. *blood sandals*
A popular brand of sandals has straps that appear to be filled with different liquids. Joanie decides to buy the water/blood sandals.
14. *metal money* vs. *glass money*
The boy scouts have raised money by collecting items for recycling. The metal/glass money alone was enough to buy new uniforms.
15. *dog woman* vs. *snake woman*
At the beach a pair of men liken two women to animals basking in the sun. One man says, "I wouldn't mind dating the dog/snake woman."
16. *television vacation* vs. *computer vacation*
Seth and Joel's mother declares that they must abandon their indoor pastimes and go play outside. She says, "It's time for you to take a television/computer vacation."
17. *surfboard wave* vs. *bottle wave*
At Zuma Beach, the waves often dump garbage and other objects on the beach. Steve and David notice that, apparently, a surfboard/bottle wave had recently come ashore.
18. *tiger tree* vs. *pirate tree*
Robin and Gretchen play a game in which different trees have different imaginary identities. Gretchen says, "I'll meet you out by the tiger/pirate tree."
19. *museum exit* vs. *water exit*
In Los Angeles, Pamela calls her husband for advice about the correct freeway exit to take to carry out her errands. She asks, "Can you help me find the museum/water exit?"
20. *lecture pen* vs. *example pen*
Diane gets advice from a friend about what she can use as a pointer during a lecture. After the lecture, Diane says, "Thanks for the idea about the lecture/example pen."

REFERENCES

Callahan, M. (1996, December 9). Screen gem. *New York,* **29,** 20, 25.

Clark, H. H. (1983). Making sense of nonce sense. In G. B. Flores d'Arcais & R. Jarvella (Eds.), *The process of understanding language* (pp. 297–331). New York: Wiley.

Coolen, R., Van Jaarsveld, H. J., & Schreuder, R. (1991). The interpretation of isolated novel nominal compounds. *Memory and Cognition,* **19,** 341–352.

Coolen, R., Van Jaarsveld, H. J., & Schreuder, R. (1993). Processing novel compounds: Evidence for interactive meaning activation of ambiguous nouns. *Memory and Cognition,* **21,** 235–246.

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **23,** 71–87.

Gerrig, R. J. (1989). The time-course of sense creation. *Memory and Cognition,* **17,** 194–207.

Gerrig, R. J., & Murphy, G. L. (1992). Contextual influences on the comprehension of complex concepts. *Language and Cognitive Processes,* **7,** 205–230.

Kučera, H., & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown Univ. Press.

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language,* **29,** 259–288.

Wisniewski, E. J. (1996). Construal and similarity in conceptual combination. *Journal of Memory and Language,* **35,** 434–453.

Wisniewski, E. J., & Love, B. C. (1998). Relations versus properties in conceptual combination. *Journal of Memory and Language,* **38,** 177–202.

(Received August 17, 1998)
(Revision received March 17, 1999)