JOURNAL OF MEMORY AND LANGUAGE 28, 375–385 (1989)

# On the Functional Equivalence of Literal and Metaphorical Interpretations in Discourse

BOAZ KEYSAR

*Princeton University*

Are functionally equivalent operations required for understanding literal and metaphorical meanings? Context stories were constructed in which sentences such as "He is a baby" had both literal *and* metaphorical "truth" values. In Experiment 1, sentence verification was most difficult when literal and metaphorical meanings conflicted, indicating that both meanings were generated. In Experiment 2, people were asked to comprehend the text rather than to verify truth. Comprehension was easiest when both literal and metaphorical meanings were plausible in context. These results suggest that nonliteral meanings are generated automatically and are integrated with context whenever a coherent interpretation can be formed. © 1989 Academic Press, Inc.

How do people understand metaphors such as: "Man is a wolf"? Traditionally, such metaphors are thought to be processed via sequential operations. A literal interpretation of the form *X is a Y* is said to be derived first and then recognized as false or anomalous. Because speakers are expected to follow conversational maxims, a search for an alternative nonliteral meaning is triggered due to the apparent violation of the truth maxim (Grice, 1975; Searle, 1979). The listener then transforms the sentence into a simile of the form *X is like a Y* and searches for a resemblance between *X* and *Y*. Therefore, the sentence is said to be comprehended as a true comparison rather than as a false identity statement. As Miller (1979) argues, because *Man is a wolf* is a violation of categories, readers "can understand the author's grounds for using the sentence by reconstructing the conceptual basis of the comparison, *Man is like a wolf,* and they can then proceed to interpret that comparison statement" (p. 229).

This view of how nonliteral meanings are understood has two major implications: First, it assumes that a literal interpretation of a sentence always precedes a nonliteral reading—suggesting that metaphorical interpretation requires extra processing compared to literal interpretation. Several findings are inconsistent with this assumption. Ortony, Schallert, Reynolds, and Antos (1978) and Inhoff, Lima, and Carroll (1984) have shown that comprehension of metaphorical sentences in context does not take longer than comprehension of comparable literal sentences. Gibbs (1979, 1984, 1986) presented analogous findings for different types of nonliteral speech such as indirect requests and idioms.

A second implication of such sequential models is that a nonliteral interpretation is optional because it requires a triggering condition such as the violation of cooperative maxims. However, Glucksberg, Gildea, and Bookin (1982) have shown that the computation of metaphorical meanings is not optional in this sense. Instead, metaphorical meanings seem to be computed

This research was supported by NSF Grant BNS 85-19462 to Princeton University, Sam Glucksberg principal investigator. I thank Howard Strauss for providing the computer program and William Prinzmetal, Cristina Cacciari, and Sachi Kumon for their insightful comments on earlier drafts. Special thanks to Linda Ginzel for help all along and to Kay Locitzer for a superb job in preparing stimulus materials and assisting in all phases of this research. I am especially indebted to Sam Glucksberg for his invaluable guidance. Reprint requests should be sent to Boaz Keysar, Department of Psychology, Princeton University, Princeton, NJ 08544.



0749-596X/89 $3.00
Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.



even when the task requires only literal interpretations. In their study, subjects were asked to verify the literal truth value of sentences such as "Some desks are junkyards." Subjects were slower to respond NO to a literally false sentence when it also had a figurative interpretation than to literally false sentences that had no such interpretation, e.g., "Some desks are roads." Glucksberg et al. (1982) concluded that metaphorical meaning was computed involuntarily and interfered with the decision.

Even with the findings outlined above, the question of how metaphors are understood is not yet answered. Metaphors may not take longer to process, but they still require different processes for comprehension. Furthermore, Dascal (1987) has questioned Glucksberg et al.'s (1982) interpretation of their metaphor interference effect. Dascal argued that the results could have been an artifact of using false sentences. Thus, the difficulty in verifying literally false sentences could be interpreted as supporting a sequential "literal first" model;

> One could as well suggest that the reason is the difficulty in determining what does the sentence *literally* mean, which is a necessary condition for assessing its truth value . . . one could imagine that, although the subjects have been told to pay attention only to the literal meaning, they automatically generate or infer *also* a metaphorical interpretation, precisely because of the difficulty of generating a literal reading in the first place (p. 277).

This seems to be a valid criticism and is consistent with the findings. In Glucksberg et al.'s experiments, subjects were faster to respond to true than to false sentences. Also, Gildea and Glucksberg (1983) argued that readers seem to use everyday discourse strategies in the laboratory; principles such as Grice's (1975) cooperative maxims and Clark and Haviland's (1977) given-new contract may be ingrained strategies that are rarely, if ever, inhibited. Subjects may involuntarily apply the "be truthful" maxim to false sentences and interpret them metaphorically if a metaphorical interpretation is available. Therefore, when subjects try to interpret false sentences, this may trigger the application of discourse principles. If so, Glucksberg et al.'s results may not reflect obligatory processing of metaphorical meanings. Instead, the time difference between the two types of false sentences (metaphorically meaningful vs. metaphorically not meaningful) may be attributable to false sentences triggering a search for alternative meanings.

If the metaphorical interference effect is an artifact of verifying false sentences, then such interference should not occur in the absence of triggers, i.e., with true sentences. The first experiment reported here is a conceptual replication of the Glucksberg et al. metaphor interference study with literally true as well as literally false sentences.

To accomplish this, sentences that could be interpreted as true or false on both literal and metaphorical grounds were utilized, e.g.,

1. My son is a baby.

This sentence may be literally true or false depending on the chronological age of the son. Interpreted metaphorically it may also be either true or false: true if the son exhibits infantile behavior and false if he is relatively mature and independent. Note that the literal and metaphorical interpretations are logically orthogonal. Depending on the situational context, such sentences may be literally and metaphorically true (L+/M+), literally and metaphorically false (L−/M−), literally false and metaphorically true (L−/M+), and literally true and metaphorically false (L+/M−).

Sentences were presented in a context that was relevant both to their literal and metaphorical interpretations. The context could render the sentences either literally true or false (e.g., the subject would be either an infant or an adult) and either metaphorically true or false (e.g., baby-like or mature behavior).

If the results of Glucksberg et al. reflect the obligatory construction of metaphorical interpretations, then we expect interference for both literally true and literally false sentences: When a test sentence is literally false (L−), a metaphorically true (M+) interpretation should produce interference. Similarly, when a test sentence is literally true (L+), we should expect interference when the context renders the sentence metaphorically false (M−). That is, interference should occur whenever literal and metaphorical meanings are incongruent. This pattern of results would indicate that metaphorical interpretations are indeed constructed in an obligatory manner.

Dascal's alternative explanation predicts a different pattern. He argued that with literally false sentences it may be difficult to determine "what does the sentence *literally* mean, which is a necessary condition for assessing its truth value." Therefore, when sentences are literally false, we would expect metaphorical interpretations to be triggered. Accordingly, this may result in interference when a sentence is literally false but metaphorically true. Because they can easily be interpreted literally, however, no such interference should be expected when sentences are literally true but metaphorically false (i.e., incongruent).

## Experiment 1

The Stroop-like interference paradigm developed by Glucksberg et al. (1982) was employed with sentences that could be interpreted both literally and metaphorically. Depending on context, the literal meaning of each sentence could be either congruent with its metaphorical meaning (L+/M+; L−/M−) or incongruent (L+/M−; L−/M+). If both literal and metaphorical interpretations are generated, interference should occur whenever literal truth is at odds with metaphorical truth. For example, when a sentence is literally false, a metaphorically true interpretation should cause interference. Similarly, when a literal interpretation is true, then interference should occur when the sentence is metaphorically false. Thus, a congruity effect is predicted: Verification should take longer following incongruent than congruent context stories.

### *Method*

*Materials.* Test sentences of the form "*X* is a *Y*" (e.g., "Bob Jones is a magician") were used. For each sentence, a set of four context sections was constructed. A context story was generated by combining two context sections—a literally related section with a metaphorically related section. As can be seen in Table 1, each context section provided relevant information for *either* the literal or metaphorical truth of the test sentence.

A literally related context section rendered the test sentence either literally true (L+) or literally false (L−). Similarly, a metaphorically related context story rendered the test sentence either metaphorically true (M+) or metaphorically false (M−). Context sections that provided information about a metaphorical interpretation were irrelevant to the literal interpretation and vice versa. Four different context *stories* for each test sentence were constructed by pairing appropriate context sections. A sample of four such sections is provided in Table 1.

TABLE 1
CONTEXT SECTIONS FOR THE TEST SENTENCE "BOB JONES IS A MAGICIAN"

| |
|---|
| *Literally True* (L+) |
| Bob Jones is an expert at such stunts as sawing a woman in half and pulling rabbits out of hats. He earns his living travelling around the world with an expensive entourage of equipment and assistants. |
| *Literally False* (L−) |
| Bob Jones is maestro and manager of a famous orchestra. They are known for their drama and style. He earns his living travelling around the world, but the expenses of a major orchestra are not minor. |
| *Metaphorically True* (M+) |
| Sometimes it seems as if Bob's money is made of rubber because he stretches it so far. How does he create such a healthy profit despite these expenses? |
| *Metaphorically False* (M−) |
| Although Bob tries to budget carefully, it seems to him that money just disappears into thin air. With such huge audiences, why doesn't he ever break even? |

*Note.* Sections were combined to form four types of context stories: L+/M+; L+/M−; L−/M−; L−/M+.



Thus, a test sentence had two interpretations in each story—literal and metaphorical, which were either congruent or incongruent. The two congruent interpretations were:

(1) Literally true and metaphorically true (L+/M+) and
(2) Literally false and metaphorically false (L−/M−).

The two incongruent interpretations were:

(3) Literally false and metaphorically true (L−/M+) and
(4) Literally true and metaphorically false (L+/M−).

Each of the 12 test sentences had four context stories, one of each type. For example, the (L+/M−) context story for the test sentence "Bob Jones is a magician" was:

> Bob Jones is an expert at such stunts as sawing a woman in half and pulling rabbits out of hats. He earns his living travelling around the world with an expensive entourage of equipment and assistants. Although Bob tries to budget carefully, it seems to him that money just disappears into thin air. With such huge audiences, why doesn't he ever break even?

*Pretesting of materials.* Prior to generating each story, test sentences were rated for literal truth and for consistency given each context *section*. This procedure had two purposes: One goal was to obtain judgments about the status of a test sentence vis-a-vis each context section. For example, to ensure that the test sentence following a literally false (L−) section would be judged as literally false and that this test sentence would also be perceived as more consistent with an M+ than with an M− context section. The second goal was to ensure that metaphorically related context sections did not influence the truth value of literal interpretations. This is important because it was hypothesized that readers would take longer to verify the literal truth of a sentence following an L+/M− context than an L+/M+ story context. It was essential, therefore, to demonstrate that metaphorically related context sections (M+ & M−) did not alter the literal truth values of test sentences (e.g., M− sections did not render test sentences literally false).

Forty Princeton University undergraduates served as pretest subjects. All were native English speakers and none had participated in any similar studies. Each context section, followed by a test sentence, was presented on a separate page. Subjects rated test sentences for (1) literal truth, (2) consistency with context, and (3) how much it made sense in context. Seven-point scales were used. For literal truth, higher numbers indicated higher perceived truthfulness; similarly, higher numbers indicated higher consistency and "sense" for the second and third rating scales, respectively. Each subject read context sections of all four types, but only one type per test sentence.

Ratings confirmed that test sentences were perceived as intended in the different context sections (See Table 2). Subjects in-

TABLE 2
MEAN RATINGS OF TEST SENTENCES VIS-A-VIS CONTEXT SECTION TYPE (STANDARD ERRORS IN PARENTHESES)

| Context type | L+ | M+ | M− | L− |
|---|---|---|---|---|
| Literal truth | 6.17 | 3.66 | 3.42 | 1.67 |
| | (.27) | (.39) | (.34) | (.25) |
| Consistency | 6.67 | 5.00 | 3.00 | 2.33 |
| | (.25) | (.30) | (.34) | (.41) |

*Note.* For literal truth, higher numbers indicate higher perceived truthfulness. For consistency, higher numbers indicate that the test sentence was more consistent with preceding context.

terpreted test sentences as true with L+ context sections and as false with L− context sections. The modal rating for each context section was computed and averaged across section type. The mean literal truth rating with L+ context sections was 6.17 compared to 1.67 for L− contexts. In contrast, the literal truth value of test sentences was not affected by the metaphorically related context sections. Sentences with M+ and M− context sections were rated near the midpoint for literal truth, with mean responses of 3.66 and 3.42, respectively. A one-way analysis of variance for context section type (L+, M+, L−, M−) yielded a significant effect for ratings of literal truth, $F(3,33) = 48.22$, $p < .0001$, $MSe = 0.854$. Newman–Keuls tests revealed that literal context sections differed significantly from one another and from metaphorically related contexts ($p < .01$). The metaphorically related context sections did not differ from one another. This indicated that metaphorically related information did not affect the literal truth value of test sentences.

A one-way analysis of variance for ratings of consistency was also significant, $F(3,33) = 28.81$, $p < .0001$, $MSe = 1.616$. Though the test sentences did not differ in "literal truth" ratings with M+ and M− context sections, they were judged as more consistent with M+ sections (mean = 5.00) than with M− sections (mean = 3.00). Test sentences were also more consistent with L+ context (mean = 6.67) than with L− (mean = 2.33). All pair-wise comparisons were significant by Newman–Keuls tests, $p < .01$. Ratings of whether test sentences "made sense" paralleled ratings of consistency.

Thus, the materials met the desired criteria: The 12 test sentences were rated as more consistent with M+ than M− context sections, yet did not differ with regard to literal truth ratings.

*Subjects.* Seventy-seven Princeton University undergraduates participated for pay. All were native English speakers and none had participated in a similar experiment. Data from four subjects who suspected the goal of the experiment and a fifth subject who failed to follow instructions were discarded. Subsequently, data from 72 subjects were analyzed.

*Design and procedure.* Each subject received all 12 test sentences, three sentences in each of the four different context conditions. Test sentences were presented in one of three randomly determined orders and each subject responded only once to each test sentence.

A 2(Literal Context: L+ or L−) × 2(Congruity: Yes or No) within-subjects design was used. An Epson Equity 2 computer controlled the experiment; stimuli were presented on a video monitor and subjects used a keyboard to make their responses. The experiment was conducted in a sound-attenuated booth.

First, the context story was presented on the screen. Subjects were instructed to read the story and then to press the space bar with their left hand when they were ready for the test sentence which then appeared on the screen. They were instructed to determine quickly and accurately whether the test sentence "is literally true or strongly implied (as such) given the preceding paragraph." Subjects used their right hand to indicate responses by pressing either of two keys marked YES or NO. An initial practice of 13 items preceded the experiment. The test session contained 8 fillers and 12 experimental items. Fillers were of the same form as experimental items and required an equal number of YES and NO responses.

### *Results and Discussion*

Data for each context type were trimmed so that latencies greater than 3 S.D. from the mean were considered as errors (this amounted to 1.6% of the data). For each subject, reaction times for correct responses within context type were averaged, resulting in four mean reaction times, one per context type. These data were submitted to a 2(Literal Context) × 2(Congruity) analysis of variance.



The results strongly support the hypothesis. As predicted, congruent stories were verified faster than incongruent (means = 885 ms & 967 ms, respectively), $F(1,71) = 7.42$, $p < .01$, $MSe = 45,472$. There was no effect for Literal Context ($p > .2$), and more importantly, Congruity did not interact with Literal Context, $F < 1$ (See Table 3). This indicates that Congruity had comparable effects for both literally true and literally false sentences. An analysis over items revealed the same pattern: Congruity had an effect, $F(1,11) = 7.09$, $p < .05$, $MSe = 24,862$ and did not interact with Literal Context ($F < 1$). Thus, the Congruity effect generalizes over items for both literally true and false sentences.

Because error rates were relatively high (12% overall), accuracy data were submitted to a 2(Literal Context) × 2(Congruity) analysis of variance. This was done in order to determine whether the pattern of error rates also reflected the Congruity effect. As can be seen in Table 4, the accuracy data support the hypothesis as well. Subjects erred more after incongruent than congruent contexts (rates = 16% and 8%, respectively), $F(1,71) = 11.33$, $p < .01$, $MSe = 0.032$. Again, Literal Context had no effect ($p > .2$) and more importantly, Congruity did not interact with Literal Context, $p > .2$. These error-rate data provide additional support for the hypothesis that metaphorical meanings are generated irrespective of the truth of literal interpretations. Finally, the error-rate data argue against a speed–accuracy trade-off interpretation of these results: When subjects were faster, they were also more accurate.

TABLE 3
MEAN VERIFICATION LATENCY (MILLISECONDS) AS A FUNCTION OF LITERALLY RELATED CONTEXTS AND CONGRUITY OF LITERAL AND METAPHORICAL INTERPRETATIONS (STANDARD ERRORS IN PARENTHESES)

| | Literal context | |
|---|---|---|
| Congruity | True | False |
| Yes | 874 | 897 |
| | (36) | (22) |
| No | 954 | 981 |
| | (44) | (29) |

TABLE 4
ERROR RATES FOR SENTENCE VERIFICATION AS A FUNCTION OF LITERALLY RELATED CONTEXT AND CONGRUITY OF LITERAL AND METAPHORICAL INTERPRETATIONS (STANDARD ERRORS IN PARENTHESES)

| | Literal context | |
|---|---|---|
| Congruity | True | False |
| Yes | .11 | .06 |
| | (.022) | (.016) |
| No | .16 | .16 |
| | (.025) | (.025) |

These results strongly support the position that metaphorical interpretations are constructed in an involuntary manner. As predicted, interference occurred whenever the literal and metaphorical meanings were incongruent. In contrast, the predictions that follow from Dascal's (1987) alternative explanation were not realized: First, literally true sentences were not reliably faster overall than literally false sentences. Second, a metaphor-interference effect was found even for literally true sentences, where there is no triggering condition. These results argue strongly against the standard model which posits metaphor interpretation to be optional, viz., contingent on a "defective" literal interpretation.

If metaphor interpretation is obligatory, then how is it accomplished? According to the standard model, assertions of the form *X is a Y* are transformed into the simile form, *X is like a Y*. Thus, for a sentence such as (1) *My son is a baby*, the literal interpretation would assign "my son" to the literal category of infants, whereas the metaphorical interpretation would assert that "my son is *like* a baby." The categorical assertion and the assertion of similarity are pragmatically incompatible. Category

membership is incompatible with assertions of similarity, e.g., if

(2a) Copper is a metal

is acceptable, then

(2b) Copper is *like* a metal

is not acceptable. Similarly, if one asserts and believes that someone is an actual baby, then one cannot simultaneously assert that the person is "like" a baby.

If a metaphor is transformed into a simile ("he is like a baby"), then within the context of this experiment, this leads to a counterintuitive prediction: Sentences that are true both literally and metaphorically should be difficult to verify because they pose a functional incongruity, i.e., category membership and similarity. This implies that the L+/M+ condition is functionally an incongruent condition because a true categorization assertion such as "My son is a baby" excludes an assertion of similarity (i.e., he is like a baby). Given this logic, the L+/M+ conditions should have longer verification latencies than the L−/M− condition because in this sense, only the L−/M− condition is truly congruent. The data show no indication of this; in fact, the L+/M+ context yields the fastest verification times. These data, then, support neither the optionality assumption nor the transform-to-simile assumption of the standard sequential model of metaphor comprehension.

## Experiment 2: Comprehension without Verification

In normal discourse, people do not generally make explicit decisions about the truth values of sentences. Instead, as each segment of spoken discourse or text is processed it is integrated with prior discourse in order to form a coherent text representation (Kintsch & van Dijk, 1978; Rumelhart, 1975). The time needed to accomplish this depends on a number of factors: comprehensibility of the sentence itself, coherence with prior text, and so on.

Any given sentence can be coherent with prior text either literally or metaphorically. If only one type of interpretation is coherent, then comprehension should be slower than if both types are coherent. This follows from the simple assumption that as soon as a coherent interpretation is found, a reader will accept that interpretation and proceed to the rest of the text. This suggests that when people are asked simply to comprehend text interference should not result from incongruent literal and metaphorical meanings. Instead, whenever either type of meaning "makes sense," that meaning should be sufficient and text processing should proceed without attempts to see if alternative meanings also make sense.

The purpose of Experiment 2 was to extend the findings of Experiment 1 to the more normal language processing task of simply comprehending text. According to a sequential model, metaphorical interpretations should not be generated whenever literal interpretations would suffice. This implies that literally false but metaphorically true sentences (L−/M+) should take longer to understand than literally true sentences. This also implies that literally true sentences that also have metaphorically true (L+/M+) interpretations will be no easier to understand than literally true but metaphorically false (L+/M−) sentences. The hypothesis that underlies this paper makes different predictions: When two interpretations are possible, comprehension should be easiest. Either literally true or metaphorically true interpretations should be sufficient for successful comprehension.

If this hypothesis is correct, then both Literal Context and Metaphorical Context should have effects on comprehension latency. Comprehension of test sentences should be faster after literally true than literally false contexts, and similarly it should be faster after metaphorically true than false contexts. These predictions were tested by measuring comprehension time instead of verification latency, utilizing exactly the same materials as in Experiment 1.



### *Method*

*Subjects.* Fifty-six Princeton University undergraduates participated for pay. All were native English speakers and none had participated in similar experiments. The data from eight subjects were discarded: three because they suspected the goal of the experiment and five because they failed a comprehension quiz that was used to determine whether subjects paid attention to the text.

*Design and procedure.* The materials from Experiment 1 were used: 12 test sentences, each with one of four context stories. However, in Experiment 2 the task as well as the presentation of materials differed. Each story was presented line by line, and subjects were instructed to press the space bar on a keyboard as soon as they comprehended each line. Comprehension time was recorded only for test sentences. Each session included one practice item after which subjects had the option of asking clarification questions before continuing with the experiment. As in Experiment 1, eight fillers were used in addition to the 12 experimental items. Items were separated by "end of story" messages which allowed for short breaks between items.

In order to ensure that subjects read each line, they were told that an occasional quiz would be introduced to test the comprehension of the preceding story. Quizzes were used following two experimental items. In each quiz, subjects indicated their responses on an answer sheet to four yes/no comprehension questions that were presented on the screen. No time pressure was involved but accuracy was stressed. The format of the answer sheet led subjects to believe that additional quizzes would be introduced. This was done in order to keep subjects alert yet avoid too many quizzes. Data from five subjects who did not answer at least 6 out of the 8 quiz questions correctly were discarded.

The experiment followed a 2(Literal Context: L+ or L−) × 2(Metaphorical Context: M+ or M−) within subjects design, thus yielding the same four cells (context types) as in Experiment 1. Except for the differences outlined above, the procedure, random assignment of subjects and the equipment were exactly the same as in the first experiment.

### *Results and Discussion*

The data were trimmed as in Experiment 1, and the resulting data were submitted to a 2(Literal Context) × 2(Metaphorical Context) analysis of variance.

As predicted, both Literal and Metaphorical Context facilitated comprehension. On average, subjects were 264 ms faster following L+ than L− context sections (See Table 5), indicating that when a literal interpretation was plausible, comprehension was easier. The possibility of a metaphorical interpretation facilitated comprehension as well: Subjects were 175 ms faster on average following M+ than M− context sections. With subjects as the random variable, the Metaphorical Context effect was reliable, $F(1,47) = 12.33$, $p < .01$, $MSe = 110{,}183$. Literal Context had a reliable effect as well, $F(1,47) = 47.60$, $p < .001$, $MSe = 72{,}015$. Importantly, Literal and Metaphorical Context did not interact ($p > .1$). This indicates that Metaphorical Context had comparable effects within the two Literal Context conditions.

An analysis with items as the random

TABLE 5
MEAN COMPREHENSION TIME (MILLISECONDS) FOR TEST SENTENCES AS A FUNCTION OF LITERALLY AND METAPHORICALLY RELATED CONTEXTS (STANDARD ERRORS IN PARENTHESES)

| | Literal context | |
|---|---|---|
| Metaphorical context | True | False |
| True | 791 | 993 |
| | (28) | (45) |
| False | 904 | 1230 |
| | (37) | (59) |

variable revealed the same pattern. The effect of Literal Context was reliable, $F(1,11) = 24.46$, $p < .001$, $MSe = 33{,}738$. More importantly, Metaphorical Context had a reliable effect, $F(1,11) = 13.30$, $p < .01$, $MSe = 27{,}490$ and did not interact with Literal Context ($p > .1$).

Thus, whenever it was possible to interpret the test sentence metaphorically, comprehension was easier.

## General Discussion

Both experiments provide evidence for the hypothesis that metaphorical meanings are computed involuntarily. In Experiment 1 evaluation of truth conditions was required: Metaphorical interpretation interfered when the context rendered literal and metaphorical meanings incongruent. This indicates that metaphorical meaning was evaluated simultaneously with literal meaning. In the second experiment, rather than a judgment task, subjects were timed for comprehension latency. No interference occurred with this task. Instead, each interpretation, literal and metaphorical, contributed independently to ease of comprehension. Whenever a literal or metaphorical interpretation was possible, comprehension was facilitated. Together these findings argue that, when available, metaphorical meanings are automatically computed. This was true both when readers verified the truth value of sentences and when they simply comprehended them.

Metaphor interference occurred in the first experiment both for literally false and literally true sentences. These results cannot be explained by Dascal's (1987) argument that literal falsehood of sentences triggers the construction of an alternative metaphorical interpretation. When sentences are literally true no such trigger is present because no discourse convention is violated. The results clearly indicate that the construction of metaphorical interpretations is obligatory; i.e., they are not constructed only after a literal interpretation is evaluated and found wanting.

The results of Experiment 2 generalize this argument from a decision task to routine reading. When people read normally, they attempt to construct a sensible interpretation of the text rather than evaluate the truth value of individual sentences. A metaphorical interpretation was constructed during reading not only when a sentence was literally false but also when it was true: Reading was facilitated whenever it was possible to interpret a sentence metaphorically. Readers, then, seem to be computing multiple contextually plausible interpretations whether they be literal or figurative.

It appears that readers evaluating the truth of sentences (Experiment 1), simultaneously consider the metaphorical interpretation as true or false. This can be contrasted with views that use truth value to distinguish between literal and metaphorical interpretations. Traditionally, only literal interpretations are said to be taken as true or false. End (1986) expresses this view succinctly: "A literal expression is either true or false if interpreted literally. A metaphor, on the other hand, is false or nonsensical if interpreted literally, but is meaningful if understood figuratively" (p. 328). Thus, a literal interpretation has two-valued meaning, a metaphor is simply "meaningful." However, the results of Experiment 1 may indicate that when people's goal is to evaluate truth, they employ a pragmatic measure of truth that includes metaphorical as well as literal truth. They appear to take nominative metaphors as expressing something that is either true or false, not as merely meaningful.

The notion of pragmatic true receives further support from findings reported by Glucksberg, Gildea, and Bookin (1982). Metaphors with an existential quantifier, e.g., "Some surgeons are butchers," were perceived as better metaphors than were those with a universal quantifier, e.g., "All surgeons are butchers." Only "good" metaphors yielded an interference effect (i.e., longer to reject as literally false.) Furthermore, when the goodness of the "All" metaphors was improved (e.g., "All rumors are diseases") both "Some" and "All" forms caused interference. These results appear to involve an evaluation of metaphorical truth. The "Some" metaphors were metaphorically true: Some surgeons *are* butchers, but not *all* of them. Therefore, no interference should be expected with these "All" metaphors, i.e., "bad" metaphors. In contrast, the "All" metaphors which were "good" were metaphorically true: All rumors *are* diseases, and indeed these metaphors produced interference. It seems, then, that pragmatic truth may have been an important factor in these experiments. Coupled with the results from Experiment 1 of this paper, it may be concluded that readers do not necessarily perceive literal truth as having a different status than metaphorical truth. Instead, both are evaluated with regard to pragmatic truth.

Theories that do not employ a notion of pragmatic truth need to explain how some literally false (or anomalous) sentences are comprehended as metaphors. A widely accepted account is that people transform a metaphor of the form *X is a Y* into a simile of the form *X is like a Y*. This transformation would yield a true interpretation, as argued by Davidson (1978): "The most obvious semantic difference between simile and metaphor is that all similes are true and most metaphors are false. The earth is like a floor . . . but it is not a floor" (p. 39). However, as noted above, such a view yields a peculiar prediction. The literal interpretation of *X is a Y* is incompatible with the corresponding similarity statement when both meanings are true (L+/M+). For example, when the context describes Bob as literally a magician, it is unacceptable to say that he is also *like* a magician even if he is metaphorically a magician (e.g., a financial whiz). Therefore, according to such a view, a literally true and metaphorically true interpretation should be functionally incongruent. If indeed metaphors are computed automatically and a metaphor is comprehended as a simile, then interference should have occurred in the L+/M+ condition in Experiment 1. The data showed exactly the opposite yielding the shortest verification latency in this condition (i.e., in interference). This indicates that the metaphorical interpretation could not have been of simile form.

If a metaphor is not transformed into a simile, then what procedure does its interpretation require? The symmetrical effects in these experiments suggest that literal and metaphorical interpretations of *an X is a Y* assertion require exactly the same treatment. One possibility, suggested by Glucksberg and Keysar (1989), is that metaphors of the form *X is a Y* are comprehended as class-inclusion statements and not as implicit similes. For example, the utterance *My husband is a baby* is not transformed into a simile, but instead is understood by assigning *husband* to a class of entities that is exemplified by infants and toddlers (although it might include dependent people of all ages, animals that require constant care, etc.) However, if this sentence were about an infant and interpreted literally, then the class referred to by the word "baby" would include a different set of entities. In this case, it would name a category that includes only members of a very young age.

Thus, the literal and figurative interpretations of *My husband is a baby* are identical in form; both express $x \in \{Y\}$, $x$ is a member of the class named $Y$. The metaphorical interpretation would be true when $X$ can indeed be a member of $Y$ and false otherwise. In this sense, sentences of this form are comprehended as class inclusions, whether they are taken literally or metaphorically; in neither case do they express (false) identity or (true) similitude.

As opposed to the standard model, this categorization-based approach is consistent with the results of both experiments. The operation required for successful interpretation is identical for both metaphorical and literal interpretations. In Experiment 1,

metaphor interference would result from competing class-inclusion interpretations. In the second experiment, the task was such that any successful class inclusion would be sufficient for comprehension, and instead of interference, mutual facilitation was found.

To sum up: Metaphorical and literal interpretations are functionally equivalent in comprehension. First, metaphorical meanings are computed in an obligatory manner during reading. Just as literal meanings, they provide an interpretation when they make sense in context. Second, metaphorical interpretations do not take on an implied simile form. Like literal interpretations, they need not require a transformation. Therefore, models that assume that literal interpretations in comprehension have priority are inappropriate, specifically, models that assume that metaphorical interpretation occurs only as a secondary or optional stage. While such models may be useful for philosophical, linguistic, or computational accounts of language, they do not adequately describe the way people actually use figurative language.

## References


Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension*. Norwood, NJ: Ablex.

Dascal, M. (1987). Defending literal meaning. *Cognitive Science*, **11**, 259–281.

Davidson, D. (1978). What metaphors mean. In S. Sacks (Ed.), *On Metaphor*. Chicago: University of Chicago Press.

End, L. J. (1986). Grounds for metaphor comprehension. In I. Kurcz, G. Shugar, & J. Danks (Eds.), *Knowledge and language*. Amsterdam: North-Holland.

Gibbs, R. (1979). Contextual effects in understanding indirect requests. *Discourse Processes*, **2**, 1—10.

Gibbs, R. (1984). Literal meaning and psychological theory. *Cognitive Science*, **8**, 575–304.

Gibbs, R. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, **9**, 17–30.

Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, **22**, 577–590.

Glucksberg, S., & Keysar, B. (1989). *Understanding metaphorical comparisons: Beyond similarity*. Manuscript submitted for publication.

Glucksberg, S., Gildea, P., & Bookin, H. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, **21**, 85–98.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts*. New York: Academic Press.

Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension in reading. *Memory & Cognition*, **12**, 558–567.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, **85**, 363–394.

Miller, G. A. (1979). Images and models: Similes and metaphors. In A. Ortony (Ed.), *Metaphor and Thought*. Cambridge: Cambridge University Press.

Ortony, A., Schallert, I., Reynolds, R., & Antos, S. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior*, **17**, 465–477.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding*. New York: Academic Press.


(Received November 4, 1988)
(Revision received January 17, 1989)