Synonymy and anomaly

RICHARD P. HONECK
University of Cincinnati, Cincinnati, Ohio 45221

and

ROBERT R. HOFFMAN
University of Minnesota, Minneapolis, Minnesota 55455

The comprehension of anomalous "strings" is frequently explained in terms of their "reduction" to grammatical and literal interpretations. Although degree of departure from grammaticality (a form of anomaly) is a potent psychological variable, the actual interpretability of anomalies has not been examined. The present study establishes some criteria, based on the concept of synonymy, for deciding if a sentence is a genuine interpretation of an anomaly. Specifically, Hoffman and Honeck's (1976) bidirectionality paradigm was used, in which in Phase 1, subjects ranked potential interpretations in terms of their similarity to poetry lines taken from e. e. cummings; in Phase 2, different subjects ranked these lines and foil lines with respect to the interpretations. The results met all of the criteria, indicating that the presumed interpretations were genuine and that cummings' lines are interpretable. Discussion centered on the inadequacy of the reduction view and on the need to incorporate factors (inference, world knowledge, etc.) that allow practically unlimited flexibility in theories of understanding.

To some theorists, deviant linguistic strings can be meaningless (Berggren, 1962; Katz, 1964; Quine, 1967). Others argue that "deviants" such as metaphors and anomalies can be interpreted by "reduction" to a grammatical and literal linguistic form (Kintsch, 1974; Smith, Rips, & Shoben, 1974). Empirical research specifically on anomalies has focused on memory tests and on judgments of the degree of departure from grammaticality (Coleman, 1965; Danks & Glucksberg, 1973; Downey & Hakes, 1968; Epstein, 1972; Malgady & Johnson, 1977; Marks & Miller, 1964; Mistler-Lachman, 1975; Moore, 1972).

Interpretability has not been examined in these studies. Experiments (e.g., Fillenbaum, 1970) in which people's "interpretations" of anomalies have been elicited suggest that interpretations are more grammatical than their anomaly bases, supporting the reduction view. However, if an "interpretation" is not also semantically similar to the base, then its grammatical structure is formally irrelevant to the reduction view. Nevertheless, there is some suggestion that anomalies can be interpreted. Shanon (1973) found common patterns in people's corrections of ungrammatical sentences containing relatively simple violations in rules of agreement. About 75% of Pollio and Burns' (1977) subjects gave acceptable interpretations of the anomalous sentences they used. Also, the subjects'

recall of anomalous sentences did not differ statistically from their recall of "natural sentences" if they interpreted the sentences before recalling them (see also, Pollio & Smith, 1979). Similarly, Harris (1979) reports that metaphors, which often violate selection restriction rules, are not necessarily more difficult to recall then their literal counterparts. And both Honeck, Riechmann, and Hoffman (1975) and Verbrugge and McCarrell (1977) report excellent recall of proverbs and metaphors, respectively, when these are prompted by conceptually related "grounds." Finally, it is common experience that deviant poetry can be interpreted.

The above research indicates through indirect means (recall, learning rates, subjects' acceptability judgments, etc.) that anomalies can be interpreted. However, no systematic criteria for determining whether an interpretation of a deviant sentence is in fact an interpretation issue from this research. One way to begin to establish such criteria involves the concept of synonymy. A set of sentences (potential interpretations) are interpretations of a base sentence if: (1) the potential interpretations order consistently with respect to their semantic similarity to the base, (2) conversely, the base orders consistently, in the context of other potential bases, with respect to its semantic similarity to the potential interpretations, (3) the judgments in Alternatives 1 and 2 yield "judgmental parity" (i.e., the function relating them should be essentially linear), (4) when the similarity of the potential interpretations to their base decreases, the variance increases in the selection of the base from among the set of potential bases.

Hoffman and Honeck (1976) found these criteria

adequately described the results of their study of the semantic similarity between proverbs and their potential interpretations. The present study used their paradigm to determine whether potential interpretations of highly anomalous lines of e. e. cummings' poetry would conform to the above criteria. Since these interpretations did not overlap their bases in content words, deep structure, or propositional structure, a test of the reduction view of anomaly was also made possible.

METHOD

Phase 1

Subjects. Twelve male and 12 female college students at the University of Cincinnati received credit in an introductory psychology course for their participation.

Materials. The stimuli for Phase 1 were 16 problems. Each problem included a base anomaly and four interpretations of it. The bases were lines of poetry from cummings (1954), selected so as to represent possible combinations of violations of phrase structure, subcategorization, and selection restriction rules. The interpretations were adapted from a corpus of over 200 "interpretations" that had been provided by subjects in a separate study (Hoffman, Note 1). Based on the experimenters' judgments, the four interpretations were designed to represent conceptually good, mediocre, poor, and unrelated instances of the base's meaning. The unrelated interpretations, constructed by the experimenters, neither confirmed nor denied the meaning of the base. To allow a test of the reduction view and of certain theories of meaning, the main criterion for the materials was that the sentences in each problem differ in content words, deep structure, and propositional structure. Example sentence sets appear in Table 1.

Design and Procedure. Each of two groups of subjects received a different random ordering of the 16 problems and of the four interpretations within problems. The subjects were given test booklets containing instructions to rank the four interpretations in terms of their semantic similarity to the line of poetry. A rank of 1 indicated most similar in meaning and 4 indicated least similar.

Phase 2

Subjects. The subjects were 48 male and 48 female introductory psychology students at the University of Cincinnati who participated for class credit.

Materials. Three anomalies, also from cummings (1954), were added to each problem set of sentences to serve as foils (see Table 1). They were selected so as to differ from each other and from the sentences already in a set in terms of content words, deep structure, propositional structure, and overall meaning.

Design and Procedure. Within each of four groups of subjects, half (six males and six females) received one random ordering of the 16 problems and half received another such ordering. Block randomization schemes were employed such that for a given problem, one group received the good interpretation, another group received the mediocre interpretation, another the poor, and another the unrelated interpretation. Each group received an equal number (four) of good, mediocre, poor, and unrelated interpretations across the 16 problems. The base appeared an equal number of times in each of the four possible positions across problems. The subjects were instructed to rank the four lines of poetry (the base and three foils) in terms of their semantic distance from a "standard" sentence, actually an interpretation.

RESULTS

The raw data were forced-choice rankings. The

Table 1
Two Representative Sentence Sets

Base:	There's no sharpest neat word for the thing.
Good Interpretation:	Some terms or feelings are not easy to convey to someone else.
Mediocre Interpretation:	For any particular event, there is not a best way to describe it.
Poor Interpretation:	We hope others don't misinterpret the intentions of what we say.
Unrelated Interpretation:	Language is crucial for all of our activities.
Foil:	What's beyond logic happens beneath will.
Foil:	The soiled canary cage of largest existence.
Foil:	If science could fail, a mountain's a mammal.
Base:	People move love hurry in a gently.
Good Interpretation:	We sometimes engage in passion quickly but softly.
Mediocre Interpretation:	Savor the feeling that seems to change so rapidly.
Poor Interpretation:	Mankind often confuses work and relaxation.
Unrelated Interpretation:	Luxuries come to be taken for granted.
Foil:	I think things which were supposed to be out of my reach.
Foil:	No sunbeam ever lies.
Foil:	Even a pencil has fear.

Note—The first base violates selection restriction rules, the second violates phrase structure, subcategorization, and selection restriction rules.

summary statistic was summed rank since the groups in both phases had an N of 24. The statistic maps inversely onto its hypothetical counterpart, semantic similarity. Thus, a lower summed rank indicates greater similarity.

The extent to which the subjects agreed in their rankings in Phase 1 was assessed by Kendall's (cited in Hays, 1973) technique. The W values ranged from .12 to .77, with 10 of the 16 problems generating values greater than .50. All but one of the corresponding chi squares were significant at the .01 confidence level, with the exception significant at the .05 confidence level (df = 3). A comparison of the a priori (experimenter-defined) good, mediocre, poor, and unrelated classifications with their subject-defined Phase 1 ranks showed 72% agreement.

For Phase 2, as the interpretation becomes less related to the base, the subjects should be less likely to select the base from among the foils as being closest in meaning to the interpretation. The base was given a lower summed rank than the foils for 16 out of the 16 good interpretations, 14 out of 16 mediocres, 13 out of 16 poors, and 11 out of 16 unrelated interpretations. Finally, there were 64 Phase 2 problems, and of the W values calculated on each, 52 of the corresponding chi squares were significant at the .01 confidence level (df = 3). Thus, the subjects showed high agreement in their rankings, despite the fact that the sentences (base
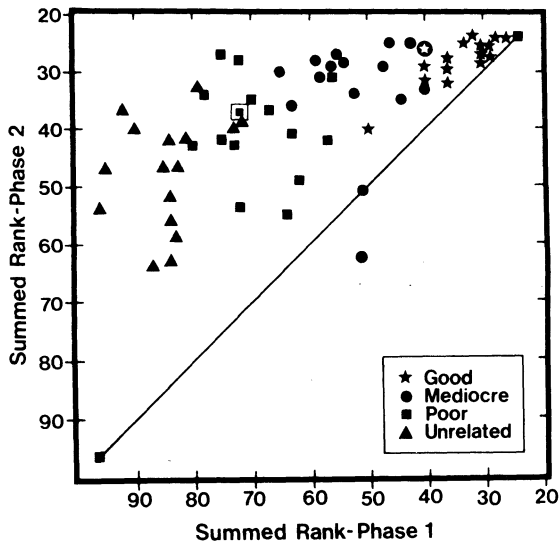
Figure 1. A scatter-plot comparison of the Phase 1 and Phase 2 data. The summed ranks assigned to the base given each of the interpretations (good, mediocre, poor, and unrelated) in Phase 2 are coordinated with the summed ranks assigned to each of the interpretations in Phase 1. The major diagonal connects minimum possible summed rank (24, 24) with maximum possible summed rank (96, 96).

and three foils) were chosen a priori to be semantically unrelated.

The two phases are compared in Figure 1. For each of the 16 problems, there are four points in this space. Both the Phase 2 summed ranks and variability appear to increase as semantic distance (Phase 1) increases. The Pearson r between the phases was .67, which is significantly different from 0 at the .01 confidence level (df = 63), and the slope of the best-fitting straight line was .36. That is, the bases in Phase 2 were given summed ranks directly related to, but slightly lower than, the ranks assigned to their corresponding interpretations in Phase 1. The a priori good, mediocre, poor, and unrelated classification was used to group the abscissa values into four sets, treating Phase 2 as a dependent variable. There was only a highly significant linear component [linear $F(1,60) = 464.83$, MSe = 48.90, $p < .001$]. Eta squared was .49, and omega squared was .44. These estimate the percentage of variance in Phase 2 accounted for by linear regression on Phase 1. According to Duncan's range tests, all four categories were significantly different from each other at the .01 confidence level (df = 60).

DISCUSSION

The results were completely consistent with the criteria outlined in the introduction. The subjects agreed in their rankings of the interpretations for every problem in Phase 1, the anomaly bases were selected in accordance with their distances from their interpretations in Phase 2, and a linear form of parity held between the phases, with an increasing inability to select the correct base as the similarity between base and

interpretation decreased. We conclude that the interpretations of cummings' lines were genuine and, therefore, that these lines were not meaningless. Indeed, they were quite meaningful, much as the Pollio and Burns (1977) mechanically generated anomalies could be considered meaningful and interpretable by their subjects.

The results also argue against the reduction view of anomaly. The relationships between interpretations and their bases were expressly constructed so as to extend beyond the theoretical reach of linguistic theory. That is, these theories contain no mechanism for deriving any of the interpretations as grammatical restatements of their bases. Indeed, a few of the interpretations themselves contained anomalous or metaphorical elements. It is clear that world knowledge, complex inferential processes, and affective content must be added to linguistic knowledge. In general, we would argue that the bases and interpretations are synthesized and matched on the basis of an abstract, holistic, nonlinguistic "conceptual base" (Honeck, 1973b; Honeck et al., 1975).

Finally, the results are consistent with and extend Hoffman and Honeck's (1976) and Honeck's (1973a) findings that synonymy between sentences can emerge in the absence of deep structural or propositional structural overlap. These and other findings (Bransford & Johnson, 1972; Verbrugge & McCarrell, 1977) require that great flexibility be built into future theories of linguistic understanding, beyond the level of recursive syntactic rules. In particular, theories must come to grips with what Hoffman and Honeck (1976) call "semantic infinity": the ability of people to create conceptual relationships between inputs (words, sentences, pictures, etc.) that seem unrelated. The major findings of the present study demonstrate this principle, but the incidental finding that the subjects consensually ranked "unrelated" items in Phase 2 is even more dramatic.

REFERENCE NOTE

1. Hoffman, R. R. Conceptual base hypotheses and the problem of "meaningless" sentences. In W. Kintsch (Chair), *Recent psycholinguistic research with metaphors.* Symposium presented at the annual convention of the American Psychological Association, San Francisco, California, August 1977.

REFERENCES

BERGGREN, D. The use and abuse of metaphor. *Review of Metaphysics*, 1962, **16**, 450-472.

BRANSFORD, J. D., & JOHNSON, M. H. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 717-726.

COLEMAN, E. B. Responses to a scale of grammaticalness. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 521-527.

CUMMINGS, E. E. *Complete poems.* New York: Harcourt Brace & Co., 1954.

DANKS, J. H., & GLUCKSBERG, S. Psychological scaling of linguistic properties. *Language and Speech*, 1973, **12**, 118-140.

DOWNEY, R. G., & HAKES, D. T. Some psychological consequences of violating linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 158-161.

EPSTEIN, W. Retention of sentences, anomalous sequences, and random sequences. *American Journal of Psychology*, 1972, **85**, 21-30.

FILLENBAUM, S. A. A note on the "search after meaning": Sensibleness of paraphrases of well-formed and malformed expressions. *Bulletin of the Psychonomic Society*, 1970, **8**, 67-68.

HARRIS, R. J. Memory for metaphors. *Journal of Psycholinguistic Research*, 1979, **8**, 61-71.

HAYS, W. L. *Stastistics for the social sciences*. New York: Holt, Rinehart, & Winston, 1973.

HOFFMAN, R. R., & HONECK, R. P. The bidirectionality of judgments of synonymy. *Journal of Psycholinguistic Research*, 1976, **5**, 173-183.

HONECK, R. P. Interpretive versus structural effects on semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 448-455. (a)

HONECK, R. P. Semantic similarity between sentences. *Journal of Psycholinguistic Research*, 1973, **2**, 137-151. (b)

HONECK, R. P., RIECHMANN, P. F., & HOFFMAN, R. R. Semantic memory for metaphor: The conceptual base hypothesis. *Memory & Cognition*, 1975, **3**, 409-415.

KATZ, J. J. Semisentences. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, N.J: Prentice-Hall, 1964.

KINTSCH, W. *The representation of meaning in memory*. Hillsdale, N.J: Erlbaum, 1974.

MALGADY, R. G., & JOHNSON, M. G. Recognition memory for literal, figurative, and anomalous sentences. *Bulletin of the Psychonomic Society*, 1977, **9**, 214-216.

MARKS, L. E., & MILLER, G. A. The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 1-5.

MISTLER-LACHMAN, J. Queer sentences, ambiguity, and levels of processing. *Memory & Cognition*, 1975, **3**, 395-400.

MOORE, T. E. Speeded recognition of ungrammaticality. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 550-560.

POLLIO, H. R., & BURNS, B. C. The anomaly of anomaly. *Journal of Psycholinguistic Research,* 1977, **6**, 247-260.

POLLIO, H. R., & SMITH, M. K. Sense and nonsense in thinking about anomaly and metaphor. *Bulletin of the Psychonomic Society*, 1979, **13**, 323-326.

QUINE, W. V. O. On a suggestion of Katz. *Journal of Philosophy*, 1967, **64**, 54-62.

SHANON, B. Interpretation of ungrammatical sentences. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 389-400.

SMITH, E. E., RIPS, L. J., & SHOBEN, E. J. Semantic memory and psychological semantics. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8). New York: Academic Press, 1974.

VERBRUGGE, R. R., & McCARRELL, N. S. Metaphoric comprehension: Studies in reminding and resembling. *Cognitive Psychology*, 1977, **9**, 494-533.