

An Intra-verbal Explication of the Nature of Metaphor

FRANK KOEN¹

University of Michigan, Ann Arbor, Michigan

Thirty subjects selected a metaphor or a literal word to complete each of 36 sentences, under instructions to choose the word that more accurately encoded a group of ideas suggested by a set of four cue words. Each sentence was presented under three conditions: (a) with a set consisting of words more often associated with the metaphorical word; (b) with a set of words more often associated with the literal word; and (c) with a set which were common associations of both metaphor and literal words. Under condition (a), the mean number of choices of the metaphor was 8.8 out of 12; under (b) the literal word was chosen an average of 7.6 times out of 12; under (c) the two response words were chosen equally often. The metaphors were also classified into five categories and were rated for goodness as figures of speech. These results were related to the other findings. The implications of these effects for an explication of the nature of metaphor were discussed.

The phenomenon of metaphor involves two terms—a figurative word or phrase which is actually used in a given linguistic context, and the literal term which might be expected in the context. The metaphor itself may be described as a statistically improbable but “appropriate” response—appropriate in the sense that it meets the semantic and grammatical requirements of the utterance. For a metaphor to be at once comprehensible and effective, the hearer must know what literal term is implied and must sense the relationship between the literal and figurative terms—a relationship involving both similarities and differences. Both the author and the hearer of a metaphorical expression, then, take account of an implicit literal term in responding to an explicit figurative one. They are dealing in shared categories. This study

proposes that metaphor is a psycholinguistic phenomenon that effects a modification of the implicit literal concept which was adumbrated by the context; that it does this by encoding a related but different conceptual realm; and that the nature of this relationship can be assessed by means of word associations.

When a metaphor is substituted for a literal term, the content of the utterance is changed. A concept is encoded which is different from, but overlaps that of a corresponding literal word. The nature of the overlap is indicated by associations common to both; differences, by associations unique to each. The argument, then, is as follows: (a) associations given by one group of Ss to a literal word (and to its corresponding metaphor) are members of distinguishable but related conceptual domains; (b) to the degree that the domains are related they will include common associations; (c) these domains are culturally acquired and widely distributed; therefore (d) a second group of Ss can use these associations as clues to re-

¹ Currently on leave of absence at Harvard University. The research upon which this paper is based was supported in part by National Science Foundation Post-doctoral Fellowship #43,042. The author wishes to thank Dr. Roger W. Brown for his help in the design of the experiment, and his useful suggestions in the preparation of the manuscript.

cover the literal word (or metaphor) to which the associations were originally given as responses.

If the foregoing is valid, it may be predicted that, when Ss must choose between metaphorical and literal response words to complete a given sentence, and are instructed to base the choice on how well each encodes a conceptual realm represented by a group of cue words, that: (1) If the cue words are associations more frequently given to the metaphor, the metaphor will be chosen more often; (2) if the cues are associations more frequently given to the literal word, it will be chosen more often, and (3) if the cues are associations given equally often to both metaphor and literal word, the two response words will be chosen equally often.

METHOD

Stimuli

The stimuli were developed by a two-stage process from 42 sentences each containing a word judged by *E* to be metaphorical in usage. The sentences were obtained from popular magazines, from a volume of poetry, and five were composed by *E*.

Operationally Defining the Literal Word. The literal word (hereinafter called LIT) was defined in the following manner: The metaphorical word (or MET) was deleted and a blank space left in its position in each sentence. Booklets were prepared, consisting of one page of instructions, three pages of practice items (not tabulated), and 42 pages, each containing one sentence repeated five times. The order of pages was randomly varied in each booklet. Ten Ss were instructed to supply for each sentence "five straightforward, ordinary words that will fit the sentence and be understood by anyone who reads it." The point was specifically made that the procedure was not a test of imagination or originality—that simple factual words were desired. The 50 "literal" words thus obtained were tabulated and the word most frequently given for each sentence was accepted as LIT, subject to the restriction that it had been given by at least five Ss. This condition eliminated six sentences, leaving a total of 36 LITs (and their corresponding METs)—9 adjectives, 13 verbs, and 14 nouns. The median number of different responses to a given stimulus item was 28, with a range from 15 to 40.

Associations. The second stage consisted of obtaining associations to each MET and LIT in the same

sentence context. The 36 items were assigned by a table of random numbers to two sub-lists, *P* and *Q*. Booklets were prepared, consisting of one page of instructions, three pages of practice items (not scored), and 36 pages each containing one sentence with five blank spaces immediately under the stimulus word. Half the booklets consisted of 18 *P* items containing METs (e.g., "The sandpiper ran along the beach, leaving a row of tiny *stitches* in the sand"), and 18 *Q* items with LITs (e.g., "The stream *flowed* over some smooth rocks and disappeared around a sharp bend"). The *P* items were presented with LIT ("The sandpiper ran along the beach leaving a row of tiny *marks* in the sand"), and the *Q* items with MET ("The stream *muttered* over some smooth rocks and disappeared around a sharp bend") in the other half of the booklets. The order of pages was randomly varied in each booklet. Twenty Ss were instructed to give five associations to each underlined word. Stress was laid on the influence of contexts on the meaning of the word—that associations were to reflect what the word meant in the specific sentence. The 50 associations to each MET and the 50 to the corresponding LIT thus obtained were tabulated in three categories: those given to both MET and LIT, and those given only to one or the other. The median number of common associations was 5, with a range from 2 to 14; the median number unique to MET, 38 (range 31 to 47); the median unique to LIT, 38 (range 32 to 44).

The Experimental Stimuli. Three sets of "cue" words were mechanically drawn for each sentence. An "M" set consisted of the two most frequent associations common to MET and LIT and the two most frequent associations unique to MET; an "L" set, the same two common associations, and the two most frequent associations unique to LIT; and a "K" set, four common associations, two of which were those used in the "M" and "L" sets. The 36 sentences were assigned by means of a table of random numbers to sub-lists A, B, and C. Thirty booklets were prepared, combining each item with each of its corresponding cue-word sets (M, K, or L) equally often. The stimulus array presented a given *S* consisted of three pages of instructions and illustrations, three pages of practice items (not scored), and 36 pages each with one stimulus item—12 each with M, K, and L cue-word sets. A single item consisted of a sentence with both MET and LIT present (e.g., "The sandpiper ran along the beach leaving a row of tiny ^{STITCHES} _{MARKS} in the sand"), and an M (prints, indentation, imprints, thread), a K (prints, scratches, imprints, trail) or an L (prints, dents, imprints, legs) set of cue words. There was also a four-point scale for use in recording the difficulty of the choice between the two words.

The order of items was randomly varied in each booklet.

In addition, two sets of 3 x 5 cards were prepared, each card containing one of the 36 sentences with MET (but without LIT) in place. These cards were used in the categorization of the METs into conceptual groups and in Q sorts of "goodness."

Design

A repeated-measures treatment by levels design was used with three conditions (defined by the set of cue words with which the item was presented) and three groups of Ss. The groups were defined by the particular sublist (A, B, or C) that was presented under conditions M, K, and L. In addition, a group of five judges sorted the 36 items into five categories of metaphor. This procedure was included to determine whether metaphors could be reliably classified and, if so, whether different kinds of METs would exhibit similar results in the experimental situation here reported. Another group of 11 judges performed Q sorts to determine how good the items were considered as figures of speech. The results of these sorts were correlated with the other data obtained for each item.

Subjects

All Ss were paid Harvard and Radcliffe undergraduate volunteers and were randomly assigned to the various groups, subject only to the requirements of the sequential arrangement used in developing the experimental stimuli. No S served in more than one phase of the experiment. The proportion of sexes in each group was allowed to vary randomly.

Procedure

Three one-hour experimental sessions were conducted, with groups ranging in size from 6 to 13. The 30 Ss were randomly assigned to one of three groups and each experimental session included representatives of each group. The S's task was to choose one of two words to best complete each stimulus sentence. Instructions stressed that the set of cue words that accompanied the sentence was to be the sole criterion upon which to base a choice, and that predilections toward graphic or imaginative words were to be curbed in the interests of the most exact encoding possible of the "group of ideas" represented by the four cue words. The Ss were also instructed to record the difficulty of each choice.

Five other Ss served (independently) as judges in a "categorization sort" of the metaphors. Each judge performed four successive dichotomous sorts, dividing the METs into (a) physical terms used to describe psychological phenomena (e.g., "The soldier was a

green recruit"); (b) transposed psychological terms (e.g., "I'll buy that idea"); (c) psychological terms describing physical phenomena (e.g., "The dark clouds promised a storm"); (d) transposed physical terms within a given sense modality (e.g., "The meadow had a hem of daisies"); and (e) synesthetic physical terms (e.g., "He wore a loud tie").

An additional group of 11 judges made independent Q sorts of the items along a nine-step scale from "best figure of speech" ("most imaginative, original, creative, and effective") to poorest. The same symmetrical distribution of items was required of all judges, although instructions encouraged the shifting of items from one category to another until the judge was satisfied with the final placement.

RESULTS

The Ss in the three experimental groups were required to choose between a MET and a LIT to complete a given sentence, using as a criterion that word which better encoded a set of four cue words. The results were analyzed in terms of the MET choices. Table 1

TABLE 1
MEANS AND STANDARD DEVIATIONS OF NUMBER OF
MET AND LIT CHOICES UNDER THREE
EXPERIMENTAL CONDITIONS^a

Response chosen		Condition		
		M	K	L
MET	Mean	8.8	5.8	4.4
	SD	1.6	2.4	1.5
	Mean	3.2	6.2	7.6
LIT	SD	1.7	2.4	1.5

^a Across 30 Ss responding to 12 items under each condition.

gives the mean number of MET and LIT choices made by the 30 Ss under each experimental condition. Analysis of variance indicated that the difference among conditions was highly significant, $F(2, 54) = 83.90, p < .001$; as was the difference among groups, $F(2, 27) = 6.29, p < .01$; and the Groups x Conditions interaction, $F(4, 54) = 20.58, p < .01$. Differences between the mean number of MET choices under conditions M, K, and L were evaluated by *t*-tests for paired

observations. The number of choices when the sentences were presented with the M sets of cue words was significantly greater than with the K sets; the K sets in turn evoked more MET choices than the L sets ($p < .01$ in both cases). The differences between sublists A, B, and C in number of MET choices across all conditions and groups were assessed by t -tests for paired observations; all values were less than 1.0.

The significant difference among experimental groups was unexpected, since Ss had been randomly assigned, and all groups were approximately equally represented in each experimental session. Further investigation indicated that the difference among groups, as well as the significant interaction, was the result of the chance assignment to one of the sub-lists of a majority of the items that evoked an unusually large number of MET choices. The analysis of variance and Table 1 provide strong support for all three predictions of the experiment.

"Goodness" of Metaphor. Eleven judges performed Q sorts of the items along a nine-step scale from best to worst metaphor. An alpha coefficient computed as an index of reliability was .76, although the average correlation between pairs of judges was only .25. Relationships involving these "goodness" judgments should probably be viewed with some caution. It is interesting to note, however, that Pearson $r = -.33$ ($p < .05$) between the mean goodness of the MET and the number of times it was chosen across all experimental conditions.

Kinds of Metaphor. The results of the categorization sort indicated that 27 of the 36 items were reliably classified as one of five kinds of metaphor (binomial test, confidence levels of .02 or better). Since the number of items in some of the classifications was quite small, a Friedman two-way analysis of variance was performed to determine whether differences existed between categories in frequency of MET choice. None did, $\chi_r^2(3) = 3.72$, $p > .30$. Some categories did reveal dif-

ferences in other respects, however. METs which were classified as synesthetic were considered better figures of speech (in the Q sort) than those whose semantic shift remained within a single modality, $t(29) = 6.05$, $p < .01$.

DISCUSSION

It may be profitable to consider the semantic and grammatical requirements of a given slot in a sentence as defining a conceptual domain. Some of the related semantic dimensions might involve such points as evaluation, time, human versus non-human and various sensory qualities. This is much like the proposal advanced by Katz and Fodor (1963). The appropriate grammatical considerations would, of course, include tense, number, and gender, among others. These latter are under the control of structural markers in the sentence. Any word that meets these referential and syntactic requirements should then be comprehensible in that slot. In addition, it is suggested that each content word in a language is a probabilistic member of many such domains, and that the semantic aspects of these domains can be explored by means of the words associated with them. A term that can be metaphorically used in a given context shares intra-verbal bonds with any word that is literally applicable in that slot, in the sense that they evoke some common associations when responded to in the same environment. Under these circumstances, metaphor and literal word may be said to share a domain.

When a figurative term is substituted for a literal one, new areas of content are introduced into the utterance, and the nature of these areas is indicated by the associations that are made to MET but not to LIT. It is in this way that MET reflects the influence of its more usual verbal contexts. As Ogden and Richards (1946) observed, "The metaphorical aspects of the greater part of language, and the ease with which any word may be used metaphorically, further indicate the

degree to which . . . words have gained contexts through other words" (p. 214). It may well be due to such implicit relationships that MET was chosen with such high frequency in condition M, and relatively seldom in condition L.

Asch and Nerlove (1960) found that, with increasing age (and, it should be added, increasing vocabulary mastery), children could state better the relationships involved in the application of physical terms (e.g., sweet, soft, warm) to descriptions of people. It is difficult to account for their results as an example of the simple extension of word usage from physical to psychological realms. However, their findings are readily interpretable in terms of increasing systematization and manipulation of intra-verbal bonds. Similarly, the explications of metaphor in terms of sensory correlations (Brown, 1958) or of functional similarities (Asch, 1958) between the referents of the metaphor and literal term would appear to be only partial answers inasmuch as both failed to take account of the *linguistic* processes implicit in the operation.

It was suggested earlier that the author of a metaphorical expression selects certain components of an event to be represented by a nonliteral term. The problem is how the selection is made. It is not enough to point to a set of dominant features common to the referents of the figurative and literal terms, because it is possible to attend to many different aspects of a complex stimulus. Properties that are salient at any given moment may not be solely a function of perceived physical similarities, but are influenced by the relations between the verbal terms we use to encode our environment, and the particular utterance.

Metaphor may be viewed as a psycholin-

guistic phenomenon rather than a purely perceptual one. Glanzer and Clark's (1962) "verbal loop" hypothesis appears relevant to this point. They suggest that perception is a two-stage affair, consisting of (a) internal translation of sensory input into words and (b) storage of, and subsequent overt response in terms of, the verbalization. To the degree that this process is generalizable, it appears necessary to consider the influence of semantic and grammatical relationships in dealing with the effects of perception. In the case of metaphor, where perception of similarities is said to occur, verbal connections might be expected to play an important role in operationally defining "similarity" itself. While metaphor may well have its genesis in perceived functional analogies between the referents of figurative and literal terms, in adult speech its character appears to be simply and accurately assessed through a culturally standardized system of intra-verbal connections.

REFERENCES

- ASCH, S. E. The metaphor: a psychological inquiry. In R. Tagiuri and L. Petrullo (Eds.) *Person, perception and interpersonal behavior*. Stanford: Stanford Univer. Press, 1958. Pp. 86-94.
- ASCH, S. E., AND NERLOVE, H. The development of double function terms in children. In B. Kaplan and S. Wapner (Eds.) *Perspectives in psychological theory*. New York: International Universities Press, 1960. Pp. 47-60.
- BROWN, R. W. *Words and things*. Glencoe, Ill.: The Free Press, 1958.
- GLANZER, M., AND CLARK, W. H. Accuracy of perceptual recall: An analysis of organization. *J. verb. Learn. verb. Behav.*, 1962, **1**, 289-299.
- KATZ, J. J., AND FODOR, J. A. The structure of a semantic theory. *Language*, 1963, **39**, 170-210.
- OGDEN, C. K., AND RICHARDS, I. A. *The meaning of meaning*. New York: Harcourt Brace, 1946.

(Received March 16, 1964)