Journal of Psycholinguistic Research, Vol. 16, No. 6, 1987

Effects of Categorical Dissimilarity and Affective Similarity Between Constituent Words on Metaphor Appreciation

Takashi Kusumi[1]

Accepted September 16, 1987

This study examines the effects of categorical and affective similarities between the constituent words of sentence (Noun-A is like Noun-B) on metaphor appreciation. In the preliminary studies using 48 words to be used later, the two kinds of similarity were measured separately by either sorting or semantic differential technique. In the main study, a set of five rating tasks concerning three types of formulaic sentences (literal, metaphorical, and anomalous) was given for each of 28 subjects in relation to (a) similarity between the words in a pairing, (b) similarity of the words in a particular formulaic sentence, (c) sentence comprehensibility, (d) sentence novelty, and (e) metaphor aptness. The application of path analysis for the above data revealed the following: (1) Categorical dissimilarity affects novelty [N], (2) affective similarity influences comprehensibility [C], and (3) [N] and [C] affect metaphor aptness.

It has been shown that, in the various forms of figurative speech (e.g., metaphor, simile), the similarity and dissimilarity between constituent terms play a fundamental role in their appreciation (e.g., Ortony, 1979a; Tourangeau & Sternberg, 1981, 1982). Typically, metaphorical sen-

A portion of the results was reported at the First Annual Conference of the Japanese Cognitive Science Society, Kyoto, July 1984. The author wishes to acknowledge the critical assistance and support of his advisors, Yoshitaka Umeoka, Hisataka Saiga, and Minoru Yabuuchi. The author also wishes to thank Katsuko Morigiwa, Shinsuke Shimojo, Yoshihisa Hamamura, Susumu Yamaguchi, Kiyomi Yamashita, and Koji Murata for their valuable suggestions on improving earlier drafts.

[1]Department of Psychology, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171, Japan.

0090-6905/87/1100-0577$05.00/0

tences (as with various forms of figurative language) contain two essential terms: *topic* that is being commented upon, and *vehicle* that is used to talk about the topic. In the metaphor "The sky (topic) is a mirror (vehicle)," for example, the constituent terms (i.e., *sky* and *mirror*) are categorically dissimilar in the sense that *sky* is a member of *natural object* category and *mirror* is a member of *artifact* category. They are, on the other hand, affectively or perceptually similar in the sense that *sky* and *mirror* have some shared features or attributes (e.g., bright, shiny).

According to Tourangeau and Sternberg (1981, 1982), there are three major views concerning similarity and aptness in metaphor: *anomaly view*, *comparison view*, and *interaction view*. The first view is that metaphors are linguistically anomalous utterances. This view emphasizes categorical dissimilarity between topic and vehicle as a basis for metaphor aptness owing to the resulting tension and novelty. The second view is that metaphors make comparisons, the grounds for the comparison being the shared features of topic and vehicle (e.g., Ortony, 1979b; Tversky, 1977). This view emphasizes affective or perceptual similarity between the two terms as a source of metaphor aptness because of the ease of interpretation. The third view is that a metaphor represents an interaction of topic and vehicle. This view emphasizes similarity and dissimilarity equally. Tourangeau and Sternberg have proposed *domains-interaction view*, incorporating elements of the three earlier views. Its details will be described later.

To explain the perceived similarity or dissimilarity between the constituent terms of metaphors, we need to grasp the underlying structure of the terms. The models for this that have been proposed may be classified as set-theoretical models and spatial models.

The set-theoretical models express the similarity between two terms as a linear combination of their common as well as distinctive features (Tversky, 1977). Ortony (1979b) extended this model to account for nonliteral similarity and metaphoricity by noting the relative degree to which the two terms in the metaphor shared salient features. Kusumi (1985) confirmed some key elements of Ortony's model by showing that (a) the saliences of common features and of the distinctive features predict the observed similarity between the terms well, and (b) three types of sentences (metaphor, reversed metaphor, and anomalous sentence), which have different levels of metaphoricity, are distinguished by their relative degrees of salience of common features.

The spatial models express the dissimilarity between two terms as a metric distance between the respective points in some coordinate space. Tourangeau and Sternberg (1981) applied the spatial models to account

for both similarity and dissimilarity between terms. In their model, the coordinate space has a hierarchical structure with higher-order and lower-order spaces. In the higher-order hyperspace, the distance between the points expresses the between-domain dissimilarity, or the degree to which the domains occupied by the terms are themselves similar (Ortony, 1979b, calls this dissimilarity "domain incongruence".) The hyperspace contains the lower-order subspaces (or categories) as points embedded within it. In the subspaces, the distance between the points denotes the within-domain similarity, or the degree to which the terms occupied similar positions with respect to their own domain. Their Experiment 2 confirmed that aptness of metaphors is positively related to the between-domain dissimilarity and to the within-domain similarity, but the former was weak.

The present study examines a more specific problem that has developed from the domain-interaction view and the spatial model approach: How could two types of distances of metaphor terms in the dual semantic spaces affect aptness mediated by such cognitive variables as judged similarity, comprehensibility, and novelty? The present study focused on the three limitations indicated by Tourangeau and Sternberg (1981) in their study, and made the following improvements.

1. Instead of the artificially constructed metaphors with relatively low aptness, which were employed by Tourangeau and Sternberg (1981), the present study uses literary metaphors, literal and anomalous sentences that have various degrees of aptness.
2. They used the semantic differential task applied to category names to investigate the hyperspace, whereas the present study uses a sorting task applied to the categorical similarity of terms to investigate the *categorical meaning space* more directly. In addition, they assumed the existence of two-dimensional analogous subspaces for each category, using the semantic differential task, but the present study assumes a uniform three-dimensional *affective meaning space* (Osgood, 1962) across categories.
3. They did not offer a causal process to explain how two types of similarity could affect aptness in metaphors, but the present study examines some causal models by making use of path analysis and discriminant analysis. Path analysis is employed to test the validity of the hypothetical causal model; multiple discriminant analysis is employed to determine the efficiency of the variables in the path model to correctly classify metaphorical, literal, and anomalous sentences.

In summary, the purpose of the present study was twofold: (a) to provide a comprehensive model of the semantic domain of metaphor terms, and (b) to test some causal models regarding how semantic similarity between the two terms affects comprehending and appreciating metaphors.

Accordingly, the principal hypotheses of this study are the following: (a) Dissimilarity between topic and vehicle is a result of their categorical incongruence, whereas similarity between the two is due to their affective or perceptual congruence. (b) Conceptual dissimilarity between topic and vehicle increases novelty, and affective similarity increases comprehensibility; they both, in turn, have a positive effect on aptness of metaphors.

METHOD

Subject

The subjects were 230 Japanese undergraduate volunteers of Gakushuin University and Tokyo Woman's Christian University.

Materials

Ninety-six simple active declarative sentences of the form A (Noun) is/are like B(noun) (in Japanese: *A ha B no youda*) were used. This form provided consistency with many of the empirical studies in the field and minimized context effects and stylistic variations. Pairs of nouns were chosen from a pool of 48 nouns to construct the following three sets of sentences (see Appendix for a list of the pairs): (a) 16 metaphorical sentences made of similes that were rated highly metaphorical in a previous study (Kusumi, 1985)[2] (e.g., The dew is like a veil); (b) 16 literal sentences constructed by the nouns in the same category (e.g., The lake is like the sea); (c) 16 anomalous sentences generated by randomly paired nouns (e.g., The smile is like a cage). For each set, reversed

[2]A total of 24 noun pairs composed of 20 Japanese literary metaphors collected by Nakamura (1977)—including several Japanese modern writers such as Yasunari Kawabata and Yukio Mishima—and 4 English literary metaphors used by Malgady and Johnson (1976). The 24 metaphors were rank-ordered with respect to metaphoricity, as defined by mean rating on a 7-point scale (Kusumi, 1985); the 16 metaphors of high metaphoricity (mean upper 4.9) were selected for the present study.

sentences were constructed by switching the topic-vehicle assignments of the 16 sentences (e.g., reversed simile: The veil is like a dew).

Procedure

Sorting Task. Each of the 48 Japanese nouns was typed on a 3 × 5-inch card. Fifty subjects were given the deck of 48 cards and instructed to sort the cards into two or more piles "on the basis of similarity of categorical meaning." The procedure was essentially the same as described by Miller (1969).

Semantic Differential. Forty subjects rated the 48 nouns (out of sentence) on 37 7-point semantic differential scales (Table I). The scales were made in accordance with the listed features of each noun obtained in Kusumi (1985). On each page of the booklet, 19 of 37 scales for each noun were presented in random order to reduce the subjects' task load while still permitting reliable estimates of the factors. For each noun, three different versions of the list of scales were assigned to different subjects. Ten relevant scales (the features listed with high frequency on each noun; Kusumi, 1985) were included in all three versions, and 9 out of the other 27 irrelevant scales were assigned randomly to one of the three versions.

Rating Task. Five rating tasks were employed: (a) similarity between topic and vehicle in the sentence, (b) similarity of topic and vehicle without a sentence, (c) sentence comprehensibility, (d) "interest"[3] of sentence, (e) aptness of metaphor. In all, 140 subjects were randomly assigned to one of five groups of equal size.

The task of the first two groups was to rate the similarity of topic to vehicle on 11-point scales (1 = very dissimilar, 11 = very similar). The first group read each of the sentences and then judged the similarity of the underlined nouns (topic and vehicle), whereas the second group looked at each of the pairs of nouns (joined by a dash) drawn from the sentences and then judged their similarity. The remaining three groups rated the sentences on three 11-point scales, labeled hard-easy (with higher numbers indicating sentences that are easier to understand), interesting-dull, apt-not apt (with higher number indicating better metaphors).

[3] "Interest" is translated from the Japanese term *omoshirosa*, which also has as its connotation novelty and attractiveness.

Table I. Factor Loadings on the 34 Semantic Differential Scales

Scales	Factors		
	1	2	3
Pleasant–unpleasant	.87	.25	.22
Beautiful–ugly	.87	.36	.17
Bright–dark	.87	−.15	.29
Transparent–opaque	.84	.29	.18
Dorodoro–sarasara[a]	−.83	−.07	.04
Good–bad	.82	.34	.12
Like–dislike	.82	.33	.27
Fragrant–foul	.77	.14	.22
Shiny–dull	.77	.21	.40
Relieved–anxious	.76	.03	.10
Happy–unhappy	.75	−.17	.20
Elegant–inelegant	.74	.33	−.05
Fearful–benign	−.72	.15	−.22
Black–white	−.69	.11	−.10
Necessary–unnecessary	.65	.30	.13
New–old	.58	−.14	.44
High–low	.57	.06	−.11
Large–small	.54	.52	.21
Sweet–salty	.51	−.28	.29
Dry–wet	.47	−.32	−.29
Heavy–light	−.39	.38	−.38
Blue–yellow	.11	.79	−.07
Deep–shallow	−.05	.78	.11
Eternal–temporal	.21	.75	.04
Mysterious–obvious	.01	.72	.35
Noisy–quiet	−.17	−.70	.03
Big–little	.26	.60	.05
Green–red	.02	.39	−.25
Covert–overt	−.30	.31	−.19
Soft–hard	.13	.05	.87
Active–inactive	.14	−.16	.76
Rounded–angular	.13	.25	.75
Free-constrained	.48	.28	.57
Hot–cold	.30	−.44	.48

[a]Dorodoro–sarasara is Japanese onomatopoeia that has some point of resemblance to thick–thin, rough–smooth, and cloggy (muddy)–fluent.

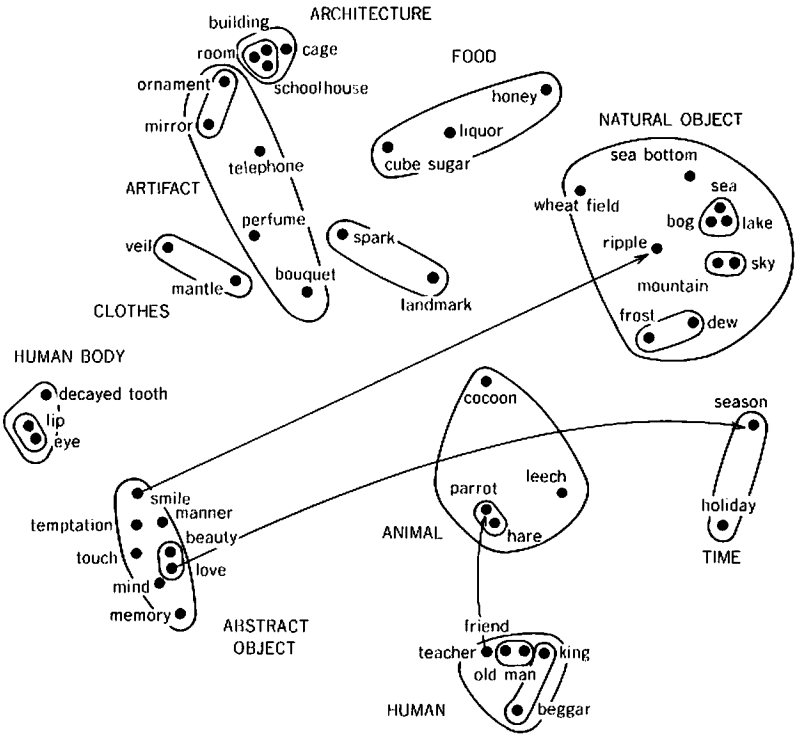


Fig. 1. Categorical meaning space from the sorting task based on the categorical similarity of metaphor terms. Points are plotted in a multidimensional space, the solid curves indicate clusters in a tree, and the arrows indicate some topic-vehicle relations of the metaphors.

RESULTS

Categorical Meaning Space

Soring data obtained from 50 subjects were pooled and then the number of subjects putting a pair of words into the same cluster was taken as a measure of the proximity of those two words.

Both nonmetric multidimensional scaling and hierarchical clustering techniques were applied to the proximity data: The multidimensional technique employed was that of Takane (1980), available as the computer program SMACOF; the clustering was achieved by the farthest-neighbor method.

The one-, two-, and three-dimensional stress values resulting from the multidimensional scaling were .450, .293, and .179, respectively.

Figure 1 presents the cluster solution superimposed on the two-dimensional Euclidean solution of SMACOF. The 48 nouns represented as points in the two-dimensional space were classified into 11 categories. The two-dimensional solution provided a more clear-cut interpretation of the resulting configuration than did the three-dimensional one. Most of the metaphors linked two nouns together, which are categorically dissimilar, such as *smile* (abstract object) and *ripple* (natural object). The distance between topic and vehicle was calculated with the assumption that the Euclidean distance between the points was its categorical distance.

Affective Meaning Space

Principal-component factor analysis was applied to the two-mode (34 scales[4] × 48 concepts) mean data matrix. In the scale mode, the first three factors accounted for 66.7% of the total variance. The first three Varimax-rotated factors appear in Table I. These factors correspond to Osgood's (1962) E-P-A (Evaluation, Potency, Activity) dimensions. Factor 1 definitely appears to be an Evaluation dimension, with high loadings on such scales as pleasant-unpleasant, beautiful-ugly, and good-bad. Factor 2 appears to be a Potency dimension, with high loadings on deep-shallow, big-little. Factor 3 appears to be an Activity dimension, with high loadings on such scales as active-inactive and rounded-angular.

Figure 2 displays the factor scores of each noun as points on these three dimensions. The figure gives the perspective representation of the affective meaning space. Most of the metaphors link two nouns together that are affectively similar, such as *love* and *season*. The distance between topic and vehicle was calculated with the assumption that the Euclidean distance between points was its affective distance.

Path Analysis

All ratings were averaged across subjects and submitted to statistical analyses.

The path analysis provides a useful means for estimating the strengths of the causal connections among variables. Table II presents the

[4]Three scales (concrete-abstract, natural-artificial, animate-inanimate), which might be related more to a categorical meaning, were excluded.

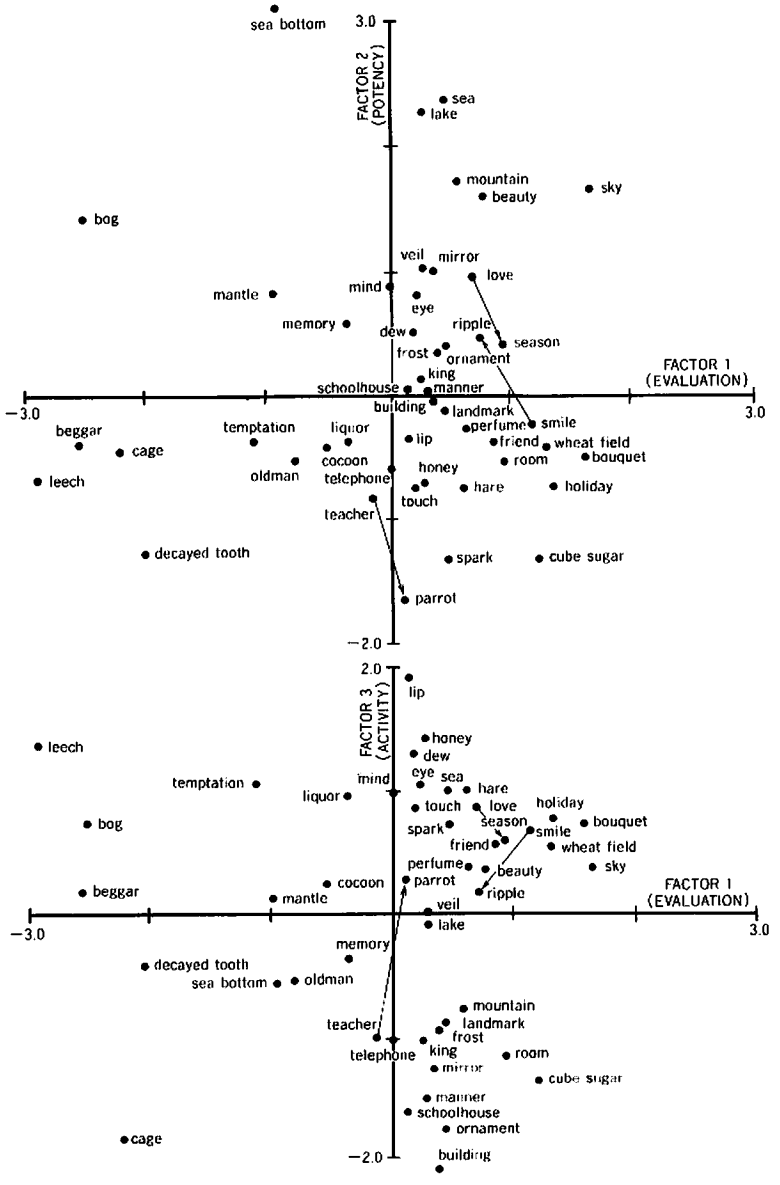


Fig. 2. Affective meaning space based on the semantic differential task of metaphor terms. Points are plotted in a three-dimensional factor-analytic space: plot of dimensions one and two (upper), one and three (lower), and the arrows indicate some topic-vehicle relations of the metaphors.

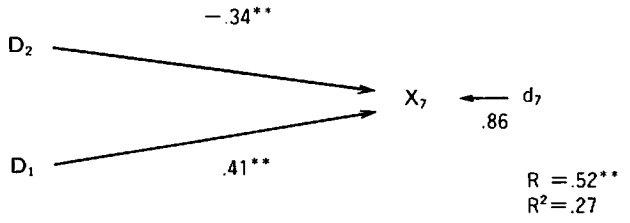
Table II. Correlations Among Aptness of Metaphor and the Other Six Variables[a]

Variables	D_1	D_2	S_3	S_4	X_5	X_6	X_7
D_1 Categorical distance	—	.08	−.17	−.37[c]	−.02	.66[c]	.32[c]
D_2 Affective distance	.08	—	−.45[c]	−.44[c]	−.34[c]	−.08	−.33[c]
S_3 Similarity within sentence	−.09	−.48[c]	—	.84[c]	.90[c]	.15	.72[c]
S_4 Similarity out of sentence	−.36[c]	−.44[c]	.80[c]	—	.67[c]	−.17	.42[c]
X_5 Comprehensibility	.05	−.36[b]	.92[c]	.63[c]	—	−.33[c]	.83[c]
X_6 "Interest" of sentence	.72[c]	−.11	.20	−.23	.35[b]	—	.66[c]
X_7 Aptness of metaphor	.39[c]	−.31[b]	.77[c]	.40[c]	.84[c]	.68[c]	—
M	.99	2.11	4.21	3.81	4.81	4.08	3.81
SD	.78	.97	2.25	2.15	2.34	1.35	1.97

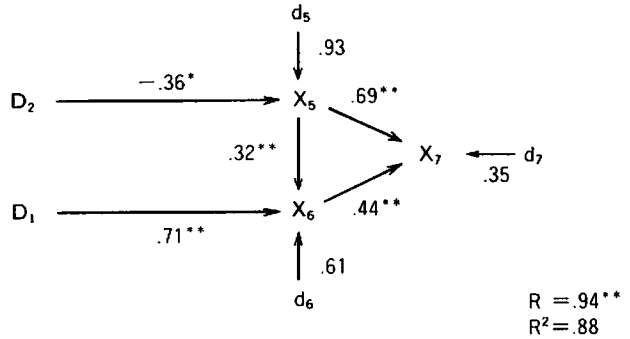
[a]Correlations above the diagonal are based on the full set of sentences (N = 96); those below the diagonal are based on a subset, omited the reversed-sentences from the full set (N = 48).
[b]$p < .05$.
[c]$p < .01$.

product-moment correlations among the seven variables used in this analysis. For two of the seven variables, the affective distance and the categorical distance, the distances between the two terms in the original sentence ("A is like B") and the reversed sentence ("B is like A") were symmetric. Therefore, the correlations for the 48 original sentences were computed independently (below the diagonal in Table II). For comparison purposes, Table II also presents the correlations based on the full 96 sentences (above the diagonal). The comparison of the correlations for the subset with the correlations for the full set revealed that most of differences were extremely small. The following analyses was based on the correlations for the 48 original sentences. All of the correlations between aptness and the other variables were significant.

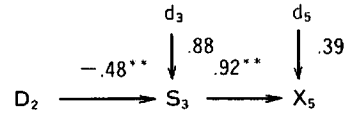
Figure 3 shows the basic results in path-analytic form (for the method, see James, Mulaik, & Brett, 1982). In this analysis, four levels of variable sets were assumed: (1) semantic distance between topic and vehicle (categorical distance (D_1) and affective distance (D_2)); (2) similarity of topic to vehicle (S_3, S_4); (3) sentence comprehensibility (X_5) and "interest" of sentence (X_6); (4) metaphor aptness (X_7). These causal orderings of variables are consistent with the principal hypothesis on metaphor appreciation. The variables of set (1) are *exogeneous variables*, which are determined by the structure of knowledge outside the model, acting as cause but not to be influenced by other variables. The variables of sets (2)–(4) are *endogeneous variables*, which are



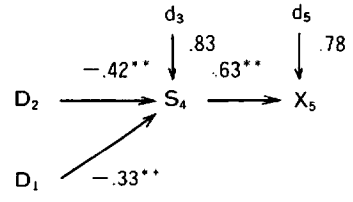
(a) Two-level model



(b) Three-level model



(b') Four-level model 1 (Part of (b))



(b'') Four-level model 2 (Part of (b))

Fig. 3. (a)-(b) Path diagrams for aptness of metaphor (X_7), influences of categorical distance between topic and vehicle (D_1), affective distance of the two (D_2), sentence comprehensibility (X_5), "interest" of sentence (X_6). (b')-(b'') Path diagrams for X_2 in (b), influences of topic-vehicle similarity within sentence (S_3) or similarity out of sentence (S_4). (All path coefficients are significant, $*p < .05$, $**p < .01$.)

determined by variables within the model. In Figure 3, one-way arrows between two variables indicate significant paths of the causal model; the structural coefficients describe the hypothetical causal relationships between the two. The path coefficients were calculated through multiple regression analyses.

As shown in Figure 3(a), the simplest model, two types of semantic distance (D_1, D_2) were significant predictors of metaphor aptness (X_7). This model indicated that increase in categorical distance (D_1) and decrease in affective distance (D_2) had a positive effect on metaphor aptness (X_7). This result was consistent with that of Tourangeau and Sternberg (1981). The multiple regression coefficient was significant but low (.52), and the two variables accounted for only 27% of the "aptness" variance.

In the modified model shown in Figure 3(b), categorical distance (D_1) affected aptness (X_7) indirectly through "interest" (X_6), whereas the direct effect of D_1 on X_7, P_{17}, was very small ($-.03$). In addition, affective distance (D_2) affected aptness (X_7) indirectly through comprehensibility (X_5), the direct effect of D_2 on X_7, P_{27}, also being very small (.09). The effect of X_5 on X_7 was decomposed into direct effect ($P_{57} = .69$) and indirect effect ($P_{56} \times P_{67} = .32 \times .44 = .14$) via "interest" ($X_6$). The multiple regression coefficient was significant and very high (.94), and the two variables (X_5,X_6) accounted for 88% of the "aptness" variance.

The small *d*'s associated with the endogenous variables are the *disturbance terms*. The arrow from d_5 to X_5 indicates other sources of variation are also affecting X_5. This residual path coefficient was large (.93) in Figure 3(b). Modified (b) models—(b′), (b″)—assume that this other source affecting comprehensibility (X_5) is judged similarity (X_3, X_4).

As shown in Figure 3(b′), the similarity between topic and vehicle within sentence (S_3) having substantial direct impact on comprehensibility (X_5), the residual path coefficient of d_5 was small (.39). In addition, S_3 was only influenced by affective distance (D_2) (the direct path from categorical distance (D_1) to S_3, P_{13}, was only $-.05$), and eliminated from Figure 3(b'). In Figure 3(b''), both distances (D_1, D_2) influenced the out-of-sentence similarity between topic and vehicle (S_4). These results suggest that similarity between the terms within sentences is judged only from affective meaning, whereas their similarity out of sentence is judged for both affective and categorical meanings. The former (S_3) can be regarded as metaphorical similarity, the latter (S_4), literal similarity.

To summarize, the path analyses indicated two causal processes: (a) The greater the categorical distance between topic and vehicle, the more interesting the metaphor was judged to be; and (b) the smaller the affective distance of the two, the more comprehensible the metaphor. They both, in turn, have a positive effect on aptness of the metaphor.

Multiple Discriminant Analysis

Subsequent to the path analysis, multiple discriminant analysis was used to determine the efficiency of the variables in the path model to correctly classify metaphorical, literal, and anomalous sentences. The predictor variables were entered in sets consistent with the four levels of variables in the path model (Figure 3): (1) semantic distance between topic and vehicle (D_1, D_2); (2) similarity of topic to vehicle (S_3, S_4); (3) characteristic of sentence (X_5, X_6); (4) aptness of metaphor (X_7). At each step the variable set entered was tested for the significance of its contribution beyond that associated with all causally antecedent variable sets.

Table III summarizes the results of the multiple discriminant analysis. The first four columns indicate means for each type of sentence and univariate F ratio. The latter is a one-way analysis of variance test for equality of group means on a single discriminating variable. This test showed that overall differences between types of sentences were significant, except for D_2. In particular, for S_3, variable sets 3 (X_5, X_6) and 4 (X_7), the mean for metaphorical sentences was higher than that for the other types of sentences (by orthogonal comparison test, $t(45) = 5.04$, 7.03, 8.40, 9.38, respectively, $p < .01$).

The fifth column indicates Wilks's lambda, which is an inverse measure of the discriminating power in the original variables that have not yet been removed by the discriminant functions. The major decreases in Wilks's lambda occurred with the addition of the first two variable sets. Furthermore, the increase in correct classification percentage (the sixth column) of the first two steps (ranging from 77.1% to 89.6%, with prior probabilities set at .333 for each type of sentence) suggests reasonable discriminating power in the variable set 1 and 2. The addition of the variables in the third and fourth step had a lesser impact on classification accuracy than that in the first two steps. Furthermore, when each variable set was entered singly, the correct classification percentages for each variable set (ranging from 72.9% to 83.3%, the seventh column) also suggested reasonable discriminating power.

Table III. Discriminant Analysis of Literal (L), Metaphorical (M), and Anomalous (A) Sentences

Step/variable set added	Means and univariate F ratio				Wilks	% correct classification	
	L	M	A	F	Λ	Cumulation	Part
1. Semantic distance between topic and vehicle						77.1	77.1
D_1 Categorical distance	.15	1.56	1.27	36.8[a]	.379		
D_2 Affective distance	1.99	1.78	2.56	3.0	.334		
2. Similarity of topic to vehicle						89.6	83.3
S_3 Similarity within sentence	4.50	5.31	1.62	28.3[a]	.183		
S_4 Similarity out of sentence	6.02	4.61	2.01	28.0[a]	.117		
3. Characteristics of sentence						91.7	79.2
X_5 Comprehensibility	5.03	6.92	2.48	36.8[a]	.096		
X_6 "Interest"	2.75	5.44	4.05	46.0[a]	.083		
4. Aptness of metaphor						91.7	72.9
X_7 Aptness of metaphor	2.98	6.05	2.41	44.9[a]	.071		

[a] $p < .01$, $df = 2/45$.

To summarize, correct identification of 91.7% of the literal, metaphorical, and anomalous sentences through discriminant analysis suggests a reasonably good discriminating efficiency among the variables in the path model.

DISCUSSION

The present study examined the causal process of metaphor comprehension and its aesthetics by path analysis. The hypothesized direction of influence from semantic distance (or dissimilarity) through cognitive variables to aptness of metaphors was confirmed. In particular, high categorical dissimilarity of topic to vehicle directly influenced "interest" of sentence, which in turn had a positive effect on aptness of metaphors. On the other hand, high affective similarity of the two directly influenced sentence comprehensibility, which also in turn had a positive effect on aptness of metaphors. These two types of effect appear to be necessary to explain the following: The former type of effect is closely related to

personification and metaphor, which is based on categorical transformation (e.g., human to object, abstract to concrete); the latter type of effect is closely related to synaesthesia and phonetic symbolism.

The dichotomy between categorical and affective meanings corresponds to common or diagnostic components and supplementary components (Nida, 1975). In the present study, the two types of meaning were investigated independently. The first, the categorical meaning, was represented by a spatial model of multidimensional scaling and a hierarchical model of cluster analysis based on sorting data of the terms. The second, the affective meaning, was represented by a spatial model of factor analysis based on semantic differential data. The distances among the terms in these two types of semantic space could predict cognitive variables concerning judged similarity of terms, sentence comprehensibility, "interest" of sentence, and aptness of metaphors.

Furthermore, this study examined the recognition of metaphors by discriminant analysis. The analysis was used to determine which of the variable levels in the causal model contributes to correctly classifying metaphorical, literal, and anomalous sentences. The discriminant analysis correctly identified approximately 90% of the three types of sentence at the first two variable levels, knowledge level (D_1, D_2) and judged-similarity level (S_3, S_4). These levels are regarded as sources for determining aptness of metaphors. In particular, metaphorical sentences were characterized by high categorical dissimilarity and high affective similarity. Literal sentences were characterized by very high categorical similarity and intermediate affective similarity. Conversely, anomalous sentences were characterized by high categorical and affective dissimilarity. The classification accuracy of this study was higher than that of Malgady (1977) (87.3% average), who used truth value and topic-vehicle similarity as predictor variables.

There could be other factors affecting aptness in metaphors. A series of extensive studies on the relationships between imagery and metaphor quality has been conducted by Marschark, Katz, and Paivio (1983), on the basis of a *dual-coding model* (Paivio, 1979). They used 10 multiple regression analyses for each of 10 variables based on the ratings of 98 constructed metaphors (in the form "A(Noun phrase) is/are B(Noun phrase)") in their study 2 but did not employ the path analysis. In light of their assumption and data (means, standard deviations, and correlation coefficients for 10 rating scales), the present study introduced path analysis, in order to reveal the hidden causal relationships among the variables.

Figure 4(a) indicates a hypothetical causal model for metaphor

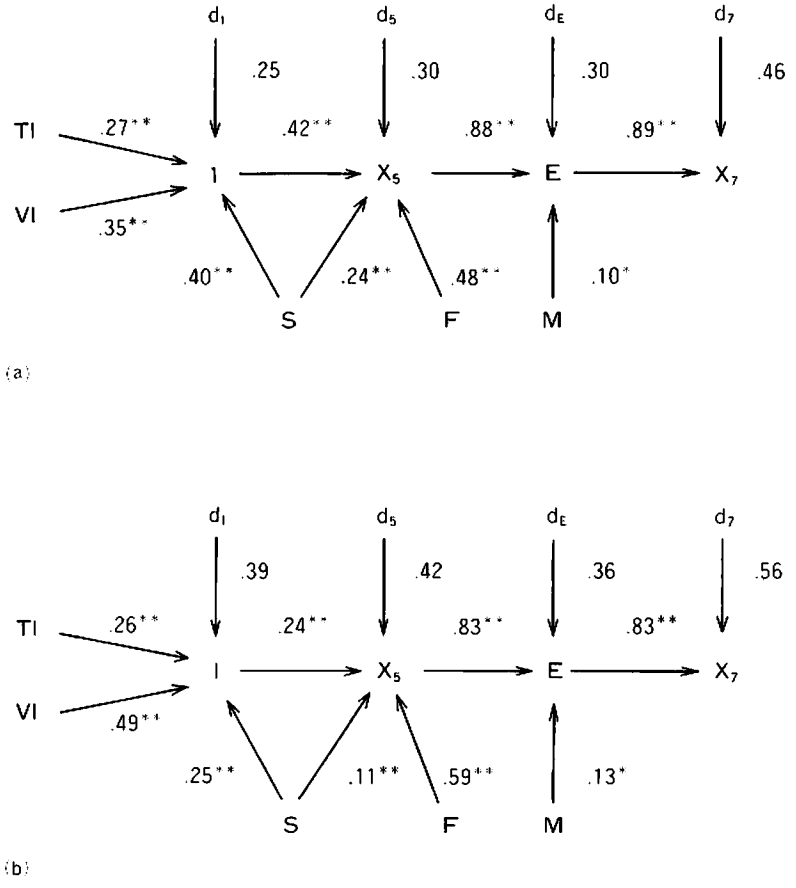


Fig. 4. Path diagrams for metaphor goodness (X_7), influences of overall metaphor imageability (I), topic imageability (TI), vehicle imageability (VI), semantic relatedness of topic to vehicle (S), felt familiarity (F), comprehensibility (X_5), degree of metaphoricity (M), and ease of metaphoric comprehension (E). (All path coefficients are significant, $*p < .05$, $**p < .01$.) Number of alternative interpretations did not have any significant path coefficient to the other variables and was eliminated from the pass. The data in (a) from Marschark et al. (1983, Study 2); the data in (b) from Katz et al. (1985).

goodness (X_7). The most important result was that overall metaphor imageability (I) indirectly affected metaphor goodness (X_7) via comprehensibility (X_5) and ease of metaphoric interpretation (E). Overall imageability (I) was directly influenced by semantic relatedness of the topic and vehicle (S), vehicle imageability (VI), and topic imageability

(TI), VI being relatively more important than TI. These results were consistent with Paivio's (1979) predictions.

Katz, Paivio, and Marschark (1985) further extended the methodology of Marschark et al. (1983) to 204 poetic metaphors (which were drawn from poetic works and were rewritten into the form "A is/are B"). This study likewise applied the path analysis to their data. The causal relationship in Katz et al. (Figure 4(b)) was essentially identical to that in Marschark et al. (Figure 4(a)), except that the effect of overall imageability (I) on comprehensibility (C) in the former study was less than that in the latter. This might be due to a strong influence of familiarity (F) on comprehensibility (C) of poetic metaphors.

According to the previous studies, there are four major views concerning metaphor processing, i.e., anomaly, comparison, interaction, and dual-coding view. This study incorporates the insights of these views into the following two relations: (a) Semantic relatedness of topic and vehicle can be classified into two types—the relatedness of categorical meaning and of affective meaning. The former affects metaphor "interest," the latter affects comprehensibility. They both have an effect on metaphor goodness. (b) Imageability of topic and vehicle and familiarity of the metaphoric ground affect metaphor goodness via comprehensibility.

The role of context in metaphor processing is also an important issue. In the present study, subjects judged aptness of metaphor in the isolated sentence. But people are usually assumed to comprehend a metaphor within the context rather than the sentence. For example, an anomalous sentence combining categorical and affective dissimilar terms is harder to comprehend in the isolated sentence. An anomalous sentence can be comprehensible if it is put in the appropriate context, which decreases affective dissimilarity between the terms. Then, the anomalous sentence becomes apt metaphor combining categorical dissimilar and affective similar terms. Such a contextual effect will be interpretable within the present path-analytic model (Figure 3(b)). In this model, the context affects aptness of metaphor (X_7) indirectly through affective distance of the terms (D_2) and comprehensibility (X_5).

Although the present study used Japanese literary metaphors, literal sentences, and anomalous sentences with varying degrees of comprehensibility and aptness of metaphor, the results of this study were congruent with those of Tourangeau and Sternberg (1981), who had used a different method of measurement of meaning and artificial metaphor. It, however, clarified in further detail the more specific causal relationships that determine aptness of metaphors.

In conclusion, dissimilarity and similarity between topic and vehicle are explained by categorical and affective meaning of the two. In particular, dissimilarity of categorical meaning of the two terms increases "interest" of sentence, and similarity of affective meaning increases sentence comprehensibility. They both, in turn, have a positive effect on aptness of metaphors.

APPENDIX

Stimuli Used in the Experiment

Metaphor Pairs

Temptation (*yûwaku*)-Honey (*mitsu*), Perfume (*kôsui*)-Bouquet (*hanataba*), Wheat field (*mugibatake*)-sea(*umi*), Mountain(*yama*)-King(*ôsama*), Eye(*me*)-Lake(*mizuumi*), Sky(*sora*)-Mirror(*kagami*), Mind (*kokoro*)-Bog(*numa*), Smile(*bishô*)-Ripple(*sazanami*), Love(*ai*)-Season (*kisetsu*), Old man(*rôjin*)-Ornament(*okimono*), Dew(*tsuyu*)-Veil(*bêru*), Teacher(*kyôsi*)-Parrot(*ômu*), Manner(*reigi*)-Cage(*ori*), Building(*tatemono*)-Cube sugar(*kakuzatô*), Friend(*yûjin*)-Liquor(*sake*), Lip(*kuchibiru*)-Leech(*hiru*)

Literal Pairs

Lake-Sea, Love-Mind, Building-Schoolhouse(*kôsha*), Mantle (*manto*)-Veil, Bog-Lake, Honey-Cube sugar, Mirror-Ornament, Frost (*simo*)-Dew, Smile-Manner, Mountain-Sky, Cage-Room(*heya*), Eye-Lip, Friend-Old man, Ripple-Sea bottom(*kaitei*), Cocoon(*mayu*)-Leech, Parrot-Hare(*usagi*)

Anomalous Pairs

Building-Liquor, Telephone(*denwa*)-Sea, Frost-Beggar(*kojiki*), Sky-Cube sugar, Schoolhouse-Lake, Love-Parrot, Manner-Spark (*hibana*), Smile-Cage, Dew-King, Teacher-Mantle, Room-Honey, Touch (*kanshoku*)-Ornament, Holiday(*kyûjitsu*)-Hare, Perfume-Cocoon, Old man-Leech, Eye-Season
(Words in parentheses indicate Japanese.)

REFERENCES

James L.R., Mulaik, S.A., & Brett, J.M. (1982). *Causal analysis: Assumptions, models and data*. Beverly Hills: Sage.

Katz, A.N., Paivio, A., & Marschark, M. (1985). Poetic comparisons: Psychological dimensions of metaphoric processing. *Journal of Psycholinguistic Research, 14*, 365–383.

Kusumi, T. (1985). Similarity between constituent words or phrases of metaphors: Effect of feature salience in comprehending metaphor. *Japanese Journal of Psychology, 56*, 269–276. (In Japanese with English summary)

Malgady, R.G. (1977). Discriminant analysis of psychological judgments of literal and figurative meaningfulness versus anomaly. *Journal of Psychology, 95*, 217–221.

Malgady, R.G., & Johnson, M.G. (1976). Modifiers in metaphors: Effects of constituent phrase similarity on the interpretation of figurative sentences. *Journal of Psycholinguistic Research, 5*, 43–52.

Marschark, M., Katz, A.N., & Paivio, A. (1983). Dimensions of metaphor. *Journal of Psycholinguistic Research, 12*, 17–40.

Miller, G.A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology, 6*, 169–191.

Nakamura, A. (1977). *Hiyu hyougen jiten* [Japanese dictionary of metaphorical expression]. Tokyo: Kadokawa Shoten.

Nida, E.A. (1975). *Component analysis of meaning: An introduction to semantic structure*. The Hague: Mouton.

Ortony, A. (1979a). The role of similarity in similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought*. Cambridge, England: Cambridge University Press.

Ortony, A. (1979b). Beyond literal similarity. *Psychological Review, 86*, 161—180.

Osgood, C. (1962). Studies on the generality of affective meaning systems. *American Psychologist, 17*, 10–28.

Paivio, A. (1979). Psychological processes in the comprehension of metaphor. In A. Ortony (Ed.), *Metaphor and thought*. Cambridge, England: Cambridge University Press.

Takane, Y. (1980). *Tajigenshakudohou* [Multidimensional scaling]. Tokyo: University of Tokyo Press.

Tourangeau, R., & Sternberg, R.J. (1981). Aptness in metaphor. *Cognitive Psychology, 13*, 27–55.

Tourangeau, R., & Sternberg, R.J. (1982). Understanding and appreciating metaphors. *Cognition, 11*, 203–244.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–352.