Bulletin of the Psychonomic Society
1981, 17(2), 69-72

Comprehension of semantic relationships and the generality of categorization models

ROGER J. S. CHAFFIN
Trenton State College, Trenton, New Jersey 08625

and

DOUGLAS J. HERRMANN
Hamilton College, Clinton, New York 13323

Six groups of subjects rated word pairs for the degree to which they exemplified one of six semantic relationships. The relationships that subjects were instructed to rate were antonymy, synonymity, subordination, superordination, coordination, and similarity. Stimulus pairs represented antonyms, synonyms, subordinates, superordinates, and coordinates. The pairs representing each stimulus relationship varied across four levels of typicality, ranging from good examples of the relationship to unrelated pairs. The highest rating in each group was given to the stimulus relationship corresponding to the relationship being judged (e.g., antonyms received the highest rating under antonym judgment instructions). This interaction was strongest for high-typicality pairs and decreased across the levels of typicality. Semantic decision models cannot explain these results unless the models are modified so that decisions are based on relationship similarity, the degree to which a stimulus pair exemplifies the relationship subjects are instructed to judge.

Current models of semantic decisions are based principally on the study of category membership decisions. A subject is presented a pair of words and decides whether one word is a member of the category named by the other word. The latency of these judgments has been found to vary as a function of the semantic similarity of the two words. Semantic similarity facilitates "yes" decisions and impedes "no" decisions (Kintsch, 1980). The models that have been proposed to explain these similarity effects are all based on the assessment of item similarity, the similarity between the two stimulus words. The mechanisms differ in different models: the probability of an additional judgment, proportion of positive and negative evidence, or order of search in a hierarchy (Cohen, 1977; Danks & Glucksberg, 1980; Kintsch, 1980; Smith, 1977). Despite these differences in mechanism, the functional variable in each case is item similarity, expressed as overlap, number of shared attributes, or distance in a network.

Another kind of similarity relevant to semantic decisions is relationship similarity, the similarity of the relationship between two stimulus words to the relationship being judged. Herrmann, Chaffin, Conti, Peters, and Robbins (1979) found that relationship similarity rather than item similarity influenced latency in antonym and synonym decision tasks. One group of subjects decided whether or not stimulus pairs were antonyms; another group classified pairs as synonymous or not synonymous. Subjects in both groups were presented five kinds of stimuli: antonyms (e.g., hot-cold), synonyms (e.g., large-big), pseudoantonyms (e.g., popular-shy), pseudosynonyms (e.g., slow-late), and unrelated pairs (e.g., dirty-curious). The results for "no" decisions showed that latency was a function of relationship similarity. For the antonym group, "no" responses were slowest for pseudoantonyms and approximately the same for synonyms, pseudosynonyms, and unrelated words. For the synonym group, "no" responses were slowest for pseudosynonyms and approximately the same for antonyms, pseudoantonyms, and unrelated words. Since pseudoantonyms were similar to the relationship judged in the antonym task and pseudosynonyms were similar to the relationship judged in the synonym task, the results indicated that relationship similarity slowed negative responses. In contrast, the results also showed that "no" response times were not affected by item similarity. As just mentioned, in the antonym task, the latencies of "no" responses to synonyms, pseudosynonyms, and unrelated pairs did not differ, although the differences between the three stimulus types in item similarity were extreme. Likewise, in the synonym task, antonyms, pseudoantonyms, and unrelated pairs did not differ in latency despite differences in item similarity (cf. Chaffin, Note 1). Item similarity did not, therefore, affect response time.

The purpose of the present experiment was to

The authors would like to thank Mary Crawford and Ed Sarafino for helpful comments on an earlier draft of this article. Requests for reprints should be sent to Roger J. S. Chaffin, Department of Psychology, Trenton State College, Trenton New Jersey 08625.

0090-5054/81/020069-04$00.65/0

Table 1
One of the Sets of Stimulus Pairs, Listed by Stimulus Condition and Typicality Level

Typicality Level	Stimulus Condition				
	Antonyms	Coordinates	Subordinates	Superordinates	Synonyms
High	wet-dry	gin-scotch	pear-fruit	crime-robbery	cheery-merry
Low	stable-frantic	football-fishing	wool-fiber	clothing-belt	modest-shy
Pseudorelation	happily-poorly	cancer-brain	anger-illness	bird-moth	sacred-rare
Unrelated	false-swift	branch-priest	mountain-tool	vehicle-noon	lightly-mighty

demonstrate, first, that judgments about relationships other than antonymy and synonymity are determined by relationship similarity (Chaffin, Russo, & Herrmann, in press; Chaffin & Andrews, Note 2). A second purpose was to determine whether the criteria used by subjects in making relationship judgments are those criteria that are suggested by logical analysis of the relationships (Lyons, 1968).

Subjects rated the degree to which stimulus pairs typified a particular semantic relationship. The stimulus pairs represented five semantic relationships: antonymy (e.g., wet-dry), synonymity (e.g., cheery-merry), coordination (e.g., gin-scotch), subordination (e.g., a pear is a fruit), and superordination (e.g., fruit is a category whose members include pear). Subjects judged the pairs in each of the five stimulus conditions for the degree to which they exemplified one of six semantic relationships. The relationships that subjects were instructed to judge included the same five relationships represented by the stimulus conditions, and, in addition, one group of subjects rated the stimulus pairs for similarity per se. Thus the same stimulus pairs were rated by six groups of subjects, each group rating the pairs for a different relationship. The stimulus pairs in each stimulus condition represented four levels of typicality.

METHOD

Subjects

One hundred and twenty undergraduates of Buena Vista College participated in the study during a class in biology, education, psychology, or sociology.

Materials

Booklets were constructed that presented rating instructions (described below) followed by two sets of 20 stimulus pairs, each set on a separate page. One set is presented in Table 1. In each set, four pairs represented each of the five semantic relations: antonyms, synonyms, coordinates, superordinates, and subordinates. The four pairs representing each relationship were selected so that one pair was a good example of the relationship (high typicality), one was a poor example (low typicality), one was a "near miss" (pseudorelation), and one was clearly not an example of the relationship (unrelated). Unrelated pairs were of the same grammatical category as the other pairs representing a relationship. Stimuli were selected to represent the form classes typical of a relationship. Within sets, words were matched in mean written frequency across the logical relation types and across the levels of typicality (mean = 111; Kučera & Francis, 1967).

The word pairs in a set were randomly ordered on a page. Two random orders were used, one the reverse of the other. The random orders and the order of the two lists were counterbalanced across subjects in each group.

Procedure

Each subject rated the 40 word pairs for the degree to which they exemplified one of the five semantic relationships. Each semantic relationship (antonymy, synonymity, coordination, superordination, and subordination) was used by 20 subjects. In addition, 20 subjects rated the 40 pairs in terms of similarity. There were thus six groups of 20 subjects, 120 in all.

The instructions specified the semantic relationship to be rated, gave examples (except in the case of the similarity relationship), and instructed subjects to decide how well each pair conformed to their idea of a good example of the relationship. For the subordination and superordination relations, the instructions also specified the order in which the subordinate and superordinate should occur.

For each set of stimulus pairs, subjects were instructed to first assign a score of 100 to the pair that best represented the relationship and a score of 10 to the pair that represented it worst and then to rate the remaining pairs. No two pairs in a set could have the same score.

Analysis

The ratings were analyzed with 6 by 5 by 4 (judgment condition, stimulus relationship, typicality level) analyses of variance, in which subjects and stimulus pairs were treated as random factors (Clark, 1973). The .05 level of significance was used in all analyses. The mean-square error terms from the analysis of subject means are reported along with min F′ values.

RESULTS AND DISCUSSION

Mean ratings of relationship similarity are shown in the five sections of Table 2. Each section presents ratings as a function of judgment and stimulus condition. The top four sections represent decreasing levels of typicality (high, low, pseudorelation, and unrelated). The bottom section provides mean ratings across typicality levels. The means for the judgment conditions at each typicality level are provided in the bottom row of each section. The means for the stimulus conditions at each typicality level are presented in the final column of each section. The main effects of judgment and stimulus condition are given in the margins (row and column) of the last section of Table 2; the main effect of typicality level is represented by the means at the bottom right of each of the top four sections. To facilitate examination of the table, the ratings in each panel that represent agreement of judgment and stimulus condition have been italicized.

The analysis of variance showed that all main effects were significant [judgment condition, min $F'(5,195) = 3.14$, $MSe = 1,386$; stimulus condition, min $F'(4,27) =$

Table 2
Mean Relationship Similarity Ratings as a Function of Judgment Condition and Stimulus Condition for Four Levels of Typicality and Collapsed Across the Four Levels of Typicality

Stimulus Condition	Judgment Condition*						
	1	2	3	4	5	6	7
				Typicality Level: High			
Antonyms	*96.6*	55.7	39.6	31.0	17.4	41.4	46.9
Coordinates	56.6	*85.7*	49.4	45.3	51.4	72.8	60.2
Subordinates	38.7	81.3	*93.8*	60.4	76.1	93.1	73.9
Superordinates	33.4	84.3	67.9	*92.1*	75.3	86.5	73.3
Synonyms	27.3	71.2	59.8	67.9	*91.0*	81.1	66.4
Mean	50.5	75.6	62.1	59.3	62.2	75.0	64.1
				Typicality Level: Low			
Antonyms	*82.5*	42.4	36.3	28.9	17.1	37.7	40.8
Coordinates	47.6	*71.4*	54.1	43.8	49.4	66.5	55.5
Subordinates	39.2	64.1	*78.2*	65.0	59.1	65.4	61.8
Superordinates	36.4	73.0	59.9	*83.2*	51.2	64.5	61.4
Synonyms	41.7	76.7	63.5	65.9	*80.9*	84.5	68.8
Mean	49.5	65.5	58.4	57.4	51.5	63.7	57.7
				Typicality Level: Pseudorelation			
Antonyms	*70.9*	45.7	33.4	27.6	22.8	33.9	39.0
Coordinates	43.2	*50.3*	47.0	48.1	32.9	51.2	45.4
Subordinates	38.9	39.7	*41.7*	40.8	31.9	39.8	38.8
Superordinates	40.3	45.8	36.2	*40.8*	32.5	43.4	39.8
Synonyms	38.0	49.7	49.7	47.0	*50.5*	45.5	46.7
Mean	46.3	46.2	41.6	40.9	34.1	42.7	42.0
				Typicality Level: Unrelated			
Antonyms	*37.0*	22.1	30.4	23.8	19.6	20.1	25.5
Coordinates	37.4	*25.1*	27.9	27.8	19.8	22.3	26.7
Subordinates	37.8	28.7	*22.9*	20.2	21.4	19.5	25.1
Superordinates	34.6	21.7	20.6	*19.0*	20.3	19.1	22.5
Synonyms	51.1	31.0	30.4	23.2	*24.9*	27.7	31.4
Mean	39.6	25.7	26.4	22.8	21.2	21.7	26.3
				Typicality Level: Overall			
Antonyms	*71.8*	41.5	34.9	27.8	19.2	33.3	38.1
Coordinates	46.2	*58.1*	44.6	41.2	38.4	53.2	46.9
Subordinates	38.7	53.4	*59.1*	46.6	47.1	54.4	49.9
Superordinates	36.2	56.2	46.2	*58.8*	44.8	53.4	49.2
Synonyms	39.5	57.1	50.8	51.0	*61.8*	59.7	53.3
Mean	46.5	53.3	47.1	45.1	42.3	50.8	47.5

**1 = antonymy; 2 = coordination; 3 = subordination; 4 = superordination; 5 = synonymity; 6 = similarity; 7 = mean.*

5.58, MSe = 417; typicality, min $F'(3,29) = 60.51$, MSe = 485]. Two of the two-way interactions were significant: the Stimulus by Judgment interaction [min $F'(20,281) = 10.44$, MSe = 417] and the Typicality by Judgment interaction [min $F'(15,292) = 4.51$, MSe = 485]. The interaction of stimulus and typicality failed to reach significance [min $F'(12,24) = 1.92$, $p < .10$, MSe = 223]. The three-way interaction was also significant [min $F'(60,197) = 2.99$, MSe = 223]. Selected comparisons for main effects were significant if the difference was greater than 6.3 for the judgment variable, 4.9 for the stimulus variable, and 3.9 for the typicality variable. Comparisons of simple main effects were significant in the two-way interactions in the following manner. For the interaction of judgment with stimulus conditions, comparisons were significant across judgment conditions if a difference exceeded 9.3 and across stimulus conditions if the difference exceeded 8.3. In the case of the Judgment by Typicality interaction, differences in judgment conditions greater than 8.3 and differences in typicality levels greater than 10.1 were significant. Finally, comparisons in the three-way interaction were significant for differences greater than 21.3 for typicality and stimulus conditions and 16.6 for judgment conditions.

Three results indicated that the rating judgments involved the processing of relationship similarity. First, the highest rating in each judgment condition was given to the stimulus condition that corresponded to the relationship being judged. Antonyms received the highest rating in the antonym judgment condition, synonyms in the synonym judgment condition, and so on. This result indicated that subjects evaluated the degree to which the relationship of a stimulus pair met the criteria for the particular relationship the subject was instructed to judge. Second, the above interaction varied as a function of typicality, being strongest at the high-typicality level, decreasing in strength across typicality levels, and disappearing for unrelated pairs. This three-way interaction is attributable to relationship similarity, which varied most widely for high-typicality pairs and decreased in variability as typicality decreased, reaching a uniformly low level with the unrelated pairs. Third, typicality effects occurred for each of the five stimulus relationships. Typicality is therefore a property of all semantic relationships, not just of category members. Typicality is the degree to which a stimulus pair typifies a particular semantic relationship (i.e., the relationship similarity of the pair).

These three results indicate that subjects employed judgment criteria that were appropriate to the relationship they were instructed to judge. However, in some cases the criteria used by subjects in making their judgments were not identical to those that would be suggested by the semantic properties of the relationship (Perfetti, 1967; Riegel & Riegel, 1963). In the coordinate judgment condition, subjects ignored the nature of the inclusion relationship between stimulus pairs and rated subordinates and superordinates as high in coordinate properties as they rated the coordinate pairs. In the antonym judgment condition, subjects used polarity and contrast as their sole criteria, rating antonyms highest, followed by coordinates. Although polarity requires that a dimension of meaning be shared by two words (Clark & Clark, 1977), sharing a dimension of meaning was not one of the judgment criteria used in the antonym judgment condition; synonyms, subordinates, and superordinates also share dimensions of meaning but were rated as low as were unrelated words. Correspondingly, in the synonym, subordinate, and superordinate judgment conditions, subjects distinguished between the shared meaning required by polarity and that required for synonymity, subordination, and superordination. In these three judgment conditions, antonyms were rated as

low as were unrelated words, although shared meaning must have been among the criteria used by subjects in each of the three judgment conditions. The judgment criteria used by subjects were thus not necessarily identical to those indicated by semantic analysis of the relationship.

The present results demonstrate that the criteria used by subjects in making semantic judgments depend on the semantic relationship the subjects are asked to judge and on the particular interpretation the subjects give to the relationship in that context. Semantic judgments must, therefore, allow different criteria to be used for different semantic judgments. It may be possible to modify current models of semantic memory to achieve this goal by simply replacing item similarity with relationship similarity as the effective variable in each model. It is more likely, however, that the decision process will differ for each relationship just as the decision criteria differ.

REFERENCE NOTES

1. Chaffin, R. *Effects of denotative and connotative similarity in judgments of synonymity.* Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1980.

2. Chaffin, R., & Andrews, K. *Part-whole relationships: Another look at class inclusion.* Paper presented at the Hamilton College Conference for Research in Psychology, June 27, 1980.

REFERENCES

CHAFFIN, R., RUSSO, A., & HERRMANN, D. An effect of relationship similarity on categorization latency. *Journal of General Psychology,* in press.

CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12**, 335-359.

CLARK, H. H., & CLARK, E. V. *Psychology and language.* New York: Harcourt Brace Jovanovich, 1977.

COHEN, G. *The psychology of cognition.* New York: Academic Press, 1977.

DANKS, J. H., & GLUCKSBERG, S. Experimental psycholinguistics. In M. P. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 31). Palo Alto, Calif: Annual Reviews, 1980.

HERRMANN, D. J., CHAFFIN, R. J. S., CONTI, G., PETERS, D., & ROBBINS, P. H. Comprehension of antonymy and the generality of categorization models. *Journal of Experimental Psychology: Human Learning and Memory,* 1979, **5**, 585-597.

KINTSCH, J. Semantic memory: A tutorial. In R. S. Nickerson (Ed.), *Attention and performance VIII.* Hillsdale, N.J: Erlbaum, 1980.

KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English.* Providence, R.I: Brown University Press, 1967.

LYONS, J. *Introduction to theoretical linguistics.* London: Cambridge University Press, 1968.

PERFETTI, C. A. A study of denotative similarity with restricted word associations. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6**, 788-795.

RIEGEL, K. F., & RIEGEL, R. M. An investigation into denotative aspects of word meaning. *Language and Speech,* 1963, **6**, 5-21.

SMITH, E. E. Theories of semantic memory. IN W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 5). Hillsdale, N.J: Erlbaum, 1977.

(Received for publication December 19, 1980.)