Computational Linguistics

A. G. OETTINGER, Editor

Contextual Correlates of Synonymy

HERBERT RUBENSTEIN AND JOHN B. GOODENOUGH
Decision Sciences Laboratory, L. G. Hanscom Field, Bedford, Masschusetts.

Experimental corroboration was obtained for the hypothesis that the proportion of words common to the contexts of word *A* and to the contexts of word *B* is a function of the degree to which *A* and *B* are similar in meaning. The tests were carried out for variously defined contexts. The shapes of the functions, however, indicate that similarity of context is reliable as criterion only for detecting pairs of words that are very similar in meaning.

Introduction

This study is concerned with the relationship between similarity of context and similarity of meaning (synonymy). More specifically, we asked how the proportion of words common to contexts containing word A and to contexts containing word B was related to the degree to which A and B were similar in meaning. The existence of this positive relationship is a basic assumption of statistical association methods in information retrieval [6]. These methods assume that pairs of words which have many contexts in common are semantically closely related and consequently that items of information in which either word occurs will tend to be relevant to the same query. To the extent that this assumption is invalid statistical methods cannot be reliable; our purpose here is to investigate the validity of this assumption.

The more sophisticated of these methods distinguish various orders of association. Thus if the sentence is taken as the unit of context any two words A and B which occur in the same sentence are first-order associates. If some other word C occurs with B in a sentence in which A is absent, then A and C are said to be second-order associates since both have B as a first-order associate. Our concern in this paper is the assumption that the degree of semantic similarity existing between a pair of words is indicated by the frequency with which they stand in first-order association to the same words. More specifically we test what is actually a converse of the assumption; namely, that the more similar in meaning two words are, the greater will be the number of first-order associates they have in common.

It is not a new notion that words which are similar in meaning occur in similar contexts. Joos [3] defined the meaning of a morpheme as "the set of conditional probabilities of its occurrence in context with all other morphemes." From this definition it was a slight step to Harris' view [2]: "If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the [contextual] distributions of A and B are more different than the distributions of A and C."

While it seems evident that words which are very similar in meaning will indeed be very similar in contextual distribution and, contrarily, words which are completely dissimilar in meaning will be very dissimilar in their contextual distributions, the nature of the relationship for intermediate values of synonymy cannot be even conjectured.

The general approach is to compare contexts from a corpus of sentences especially generated for a small set of words (hereafter called '*theme words*') with human judgments of the degree of synonymy existing between pairs of these theme words. Similarity of contextual distribution is measured by the overlap between the sentence sets of the themes being compared. However context is defined in four different ways, each definition specifying a different set of conditions under which words in the sentence set are to be considered first-order associates to the theme word. Thus the context is defined as:

(1) all words within a sentence,

(2) all content words within a sentence that fall in a certain frequency range according to the Lorge Magazine Count [8],

(3) all content words which stand in closest proximity to the theme word in the grammatical schema of each sentence,

(4) all words judged to be most closely associated with the theme, where words A and C are considered to be closely associated if the occurrence of word C is judged as strongly implying the occurrence of A and vice versa.

Using 65 pairs of words (which range from highly synonymous pairs to semantically unrelated pairs) the

This is ESD-TR-65-218 of the AF Electronic Systems Division, Air Force Systems Command. This research was performed at the Decision Sciences Laboratory as part of Project 2806, Man-Computer Dynamic Interaction. Further reproduction is authorized to satisfy the needs of the U. S. Government.

Some of the data analysis and the writeup were carried out while the authors were on leave at Harvard University. Since the first author was a Research Fellow at the Center for Cognitive Studies, financial support under ARPA SD-187 should be acknowledged.

relation is shown between similarity of meaning and the amount of overlap for each definition of context.

The 65 word pairs consist of ordinary English words. It was felt that since the phenomenon under investigation was a general property of language there was no necessity to study technical vocabulary. Furthermore, using technical vocabulary would raise the practical difficulty of finding enough competent people to serve as judges of synonymy for such words.

Procedures

Synonymy Judgments. The purpose of this procedure was to obtain judgments on how similar in meaning one word was to another. There were 65 pairs of nouns (*theme pairs*) presented for judgment. Each subject was given a shuffled deck of 65 slips of paper, each slip containing a different theme pair. The subject was given the following instructions:

1. After looking through the whole deck, order the pairs according to amount of "similarity of meaning" so that the slip containing the pair exhibiting the greatest amount of "similarity of meaning" is at the top of the deck and the pair exhibiting the least amount is on bottom.
2. Assign a value from 4.0–0.0 to each pair—the greater the "similarity of meaning," the higher the number. You may assign the same value to more than one pair.

Two groups of college undergraduates were paid to serve as subjects. Group I, consisting of 15 subjects, met for two sessions two weeks apart. In the first session they gave synonymy judgments on 48 pairs of themes including 36 of the 65 pairs finally selected for the study. In the second session they gave synonymy judgments on the 65 pairs finally selected. Thus there were 36 theme pairs used in both sessions. These pairs enabled us to compute the intra-subject reliability in judging synonymy. The product-moment correlation was computed between the first and second judgments on these 36 pairs for each subject. Despite the fact that these pairs were intermingled among different pairs in the two sessions the correlation turned out quite high; the average over all 15 subjects was $r = .85$.

A second group of 36 subjects (Group II) from the same general population as Group I participated only in the second session but gave judgments on all 65 pairs. Accordingly for each of the 65 pairs of words judged by both groups of subjects in the second session a mean judgment was calculated for Groups I and II independently. The correlation between these two sets of means came to $r = .99$. Thus it seemed reasonable to pool the judgments of the two groups. The synonymy values shown in Table 1, and used in all the discussions below, are the means of the judgments collected at the second experimental session from both groups totaling 51 subjects.

Generation of the Corpus. The purpose of this procedure was to obtain a set of contexts for each of the theme words. Each set consisted of 100 sentences written for each different word in the 65 theme pairs. There are 48 different nouns represented in the theme pairs. They are shown in Table 2.

The theme pairs always contained one word from the column *A* and one from the column *B*. Fifty undergraduates were paid to write two sentences using each of the theme words in column *A* and fifty other undergraduates using the theme words in column *B*. (These subjects had *not* participated in making the synonymy judgments.) The subjects were told that the themes were to be used as nouns only and that each sentence was to be at least 10 words long. (The average length actually turned out to be

TABLE 1. JUDGED SYNONYMY OF THEME PAIRS

cord	smile	0.02	hill	woodland	1.48
rooster	voyage	0.04	car	journey	1.55
noon	string	0.04	cemetery	mound	1.69
fruit	furnace	0.05	glass	jewel	1.78
autograph	shore	0.06	magician	oracle	1.82
automobile	wizard	0.11	crane	implement	2.37
mound	stove	0.14	brother	lad	2.41
grin	implement	0.18	sage	wizard	2.46
asylum	fruit	0.19	oracle	sage	2.61
asylum	monk	0.39	bird	crane	2.63
graveyard	madhouse	0.42	bird	cock	2.63
glass	magician	0.44	food	fruit	2.69
boy	rooster	0.44	brother	monk	2.74
cushion	jewel	0.45	asylum	madhouse	3.04
monk	slave	0.57	furnace	stove	3.11
asylum	cemetery	0.79	magician	wizard	3.21
coast	forest	0.85	hill	mound	3.29
grin	lad	0.88	cord	string	3.41
shore	woodland	0.90	glass	tumbler	3.45
monk	oracle	0.91	grin	smile	3.46
boy	sage	0.96	serf	slave	3.46
automobile	cushion	0.97	journey	voyage	3.58
mound	shore	0.97	autograph	signature	3.59
lad	wizard	0.99	coast	shore	3.60
forest	graveyard	1.00	forest	woodland	3.65
food	rooster	1.09	implement	tool	3.66
cemetery	woodland	1.18	cock	rooster	3.68
shore	voyage	1.22	boy	lad	3.82
bird	woodland	1.24	cushion	pillow	3.84
coast	hill	1.26	cemetery	graveyard	3.88
furnace	implement	1.37	automobile	car	3.92
crane	rooster	1.41	midday	noon	3.94
			gem	jewel	3.94

TABLE 2. LIST OF THEME WORDS

A		*B*	
asylum	gem	automobile	midday
autograph	glass	bird	monk
boy	graveyard	cemetery	pillow
brother	grin	forest	rooster
car	mound	fruit	sage
coast	noon	hill	serf
cock	oracle	implement	shore
cord	slave	jewel	signature
crane	tool	journey	smile
cushion	voyage	lad	stove
food	wizard	madhouse	string
furnace	woodland	magician	tumbler

about 13.5 words long.) Thus sets of 100 sentences were obtained for each of the 48 themes.

By having the sentences for the themes in each column written by different subjects we avoided obtaining the spuriously high overlap which might result if the sentences for both members of a theme pair were written by the same subjects. Systematic effects of the ordering of the themes upon sentences generated by the subjects were avoided by presenting a different ordering of the themes to each subject and by placing only one theme word on each page of the booklet in which the sentences were written. The fact that each subject contributed only two sentences to each set made the peculiarities of any particular subject negligible.

Measures of Overlap. Many measures have been proposed for characterizing the amount of overlap in two contextual distributions. A summary is given by Giuliano and Jones [1] and Kuhns [4]. We have chosen a simple formula whose interpretation seems intuitively clear.

Let S_A be the set of all words used in the sentences written for theme word A. Let A_x be a subset of S_A defined according to some condition, x. Then the measure of overlap under condition x is defined as

$$M_x = \frac{N(A_x B_x)}{\text{Min } [N(A_x), N(B_x)]}.$$

Expressed verbally M_x is the number of words shared under condition x divided by the number of items in A_x or in B_x, depending on which is the lesser. (This denominator is the maximum number of items that could be shared under condition x.)[1]

The measures used in this study are derived from two conditions: a *type* condition and a *token* condition. The type condition defines a subset A_y consisting of the different word types which occur in S_A. This condition which defines the type measure M_y is developed from the assumption that only the number of different word types in the overlap is significant and that the frequencies with which these types occur in S_A or S_B is irrelevant. The token condition which defines the token measure M_k, on the other hand, does take into account the frequencies of occurrence of words in S_A and S_B by including a given word type in A_k (or B_k) just as often as that type occurred in S_A (or S_B). Thus the token measure M_k reflects the similarity in the frequency distributions of word types. Table 3 illustrates the calculation of M_y and M_k for a sample of the contextual distributions of the themes *gem* and *jewel*.

TABLE 3. SAMPLE CALCULATION OF M_y AND M_k

S_{gem}	S_{jewel}
priceless	priceless
priceless	priceless
	priceless
jewelry	jewelry
jewelry	
ray	ray
	faceting

$\text{Gem}_y \cap \text{Jewel}_y$: priceless, jewelry, ray

$N(\text{Gem}_y) = 3, \quad N(\text{Jewel}_y) = 4$

$$M_y = \frac{N(\text{Gem}_y \cap \text{Jewel}_y)}{\text{Min } [N(\text{Gem}_y), N(\text{Jewel}_y)]} = \frac{3}{3} = 1.00$$

$\text{Gem}_k \cap \text{Jewel}_k$: priceless, priceless, jewelry, ray

$N(\text{Gem}_k) = 5, \quad N(\text{Jewel}_k) = 6$

$$M_k = \frac{N(\text{Gem}_k \cap \text{Jewel}_k)}{\text{Min } [N(\text{Gem}_k), N(\text{Jewel}_k)]} = \frac{4}{5} = .80$$

[1] A slightly different formula replaced the denominator of M_x with $N(A_x) + N(B_x) - N(A_x \cap B_x)$ and yielded qualitatively similar results.

Unrestricted Context. The context of a theme word is defined here as *all* the words—content words and function words—that occur in its sentence set. Figure 1 shows overlap versus judged similarity of meaning for this definition of context. The lower curve is composed of the M_y values and the upper curve is composed of the M_k values. The smooth curves, in this and later figures, are third-order least-square fits, shown simply to characterize the trend of the values.

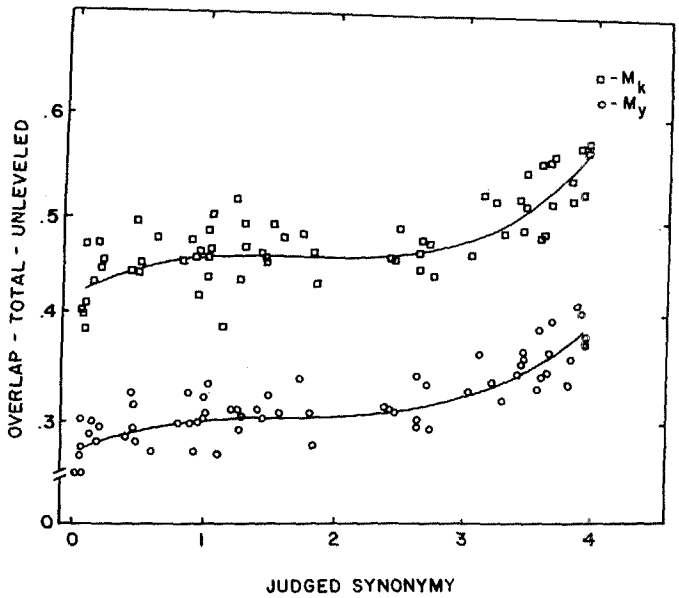

FIG. 1. Context unrestricted. Each point represents one of 65 pairs of theme words. The contexts from which the overlap is derived are all the content and function words in the sentence sets in which the theme words occur. The parameter is the condition used to calculate overlap.

The plots clearly support the hypothesis that the more similar words are in meaning, the more similar they are in their contextual distributions. It is apparent however that this relationship is strongest for the highly synonymous pairs, i.e., those with a judged synonymy greater than 3.0. For the intermediate values 1.0–2.7 overlap is almost constant. This constancy may mean that the subjects did not react to the distance 1.0–3.0 on the scale as they did to the distance 2.0–4.0. However, subjects' judgments were reliable in the range 1.0–3.0 since the correlation between subjects' judgments in this interval on two trials

averaged .67 and the correlation for the mean judgments of the two groups in this interval was .97. Thus, although we do not know whether we can truly represent the synonymy judgment on a ratio scale, it seems clear that the subjects were at least consistent in the area 1.0–3.0. (The standard deviation for subject judgments in this interval ranged from .70 to 1.30 and was greater than at the extremes of the scale, but this is to be expected.)

The picture presented by Figure 1 remained essentially unchanged when morphological differences among the words in the sentence sets were leveled; e.g., *walk*, *walks*, *walking* were all considered occurrences of *walk*; *boy*, *boys*, *boy's*, *boys'* were all considered occurrences of *boy*; *pretty*, *prettier*, *prettiest*, *prettily* were all considered occurrences of *pretty*. The result of such leveling was to reduce the average number of types per sentence set from 478 to an average of 420 and to increase the overlap primarily for M_y and only slightly for M_k. Figure 2 shows that the effect of morphological leveling was merely to raise the curves shown in Figure 1 an approximately constant amount over all synonymy values.

In both Figures 1 and 2 the slope of M_k is greater than the slope of M_y for synonymy values above 3.0. This implies that the number of tokens in the overlap increases more rapidly than the number of types—in other words

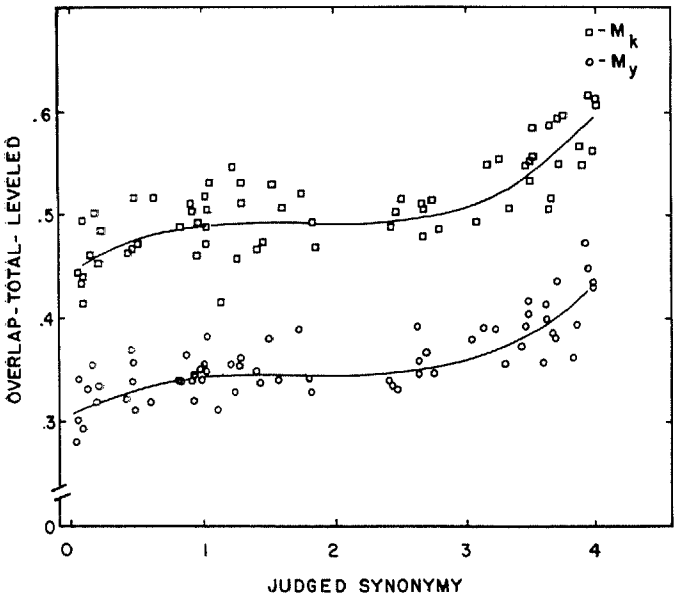

FIG. 2. Effect of leveling. The conditions are the same as in Figure 1 except that the contexts were morphologically leveled before the overlap was calculated.

TABLE 4. CORRELATIONS (r) BETWEEN OVERLAP MEASURES FOR UNRESTRICTED CONTEXT

	M_{ku}	M_{yl}	M_{kl}
M_{yu}	.87	.96	.87
M_{ku}		.87	.98
M_{yl}			.88

u = unleveled; l = leveled

that the number of occurrences per type in the overlap is greater for pairs of greater synonymy.

All four measures are very closely correlated as can be seen in Table 4.

Having shown that the amount of overlap is indeed a function of the semantic similarity existing between a pair of words, we must now take up the question that is of primary interest for the associative method of information retrieval: To what extent can the degree of synonymy existing between a pair of words be inferred from their contextual overlap? It is quite apparent from Figures 1 and 2 that one can reasonably expect to make only a broad division on the basis of overlap—namely, between theme pairs having a degree of synonymy less than 3.0 and those having a synonymy value greater than 3.0. To quantify this observation we used the following criterion.

Inference Power. The overlap measure can be used as a statistic to test the hypothesis that a given word pair has (or would have) a judged synonymy value less than 3.0, i.e., a medium or low synonymy pair. We can estimate the probability of rejecting this hypothesis when it is in fact true (Type I error) by considering the percentage of known low synonymy pairs that exceed a particular overlap value. If the probability of Type I error is fixed (in our study we chose 1 percent and 5 percent levels), this determines a *critical value* for the overlap statistic. We now calculate the percentage of the known high synonymy pairs whose overlap is greater than this value and take this percentage as an estimate of the probability of rejecting the hypothesis when it is in fact false. This is the *percentage of correct inferences* (or the *inference power*) of the overlap measure (at the x percent error level) for a particular definition of context.

In our study 20 theme pairs had a judged synonymy greater than 3.0 and so the proportion of these pairs that exceeds the critical overlap value is the estimate of inference power for a particular test. To estimate Type I error we considered all possible theme word pairs such that each word of the pair had sentence sets written by different groups of subjects, i.e., 576 pairs. Judgments were actually obtained for only 65 of these pairs. The remaining 511 pairs could not be submitted for judgment to subjects for practical reasons, but it seemed clear that these pairs were less synonymous than the pairs which were judged to have synonymy values greater than 3.0. Thus using the overlap values for 556 low synonymy pairs we determined a critical value which was exceeded by 1 percent or 5 percent of these pairs and then determined the proportion of the 20 high synonymy pairs that exceeded the critical value.

Table 5 shows the percentages of correct inferences (and critical values) for overlap based on unrestricted context as well as on the other contexts studied.

Note that the percentage of correct inferences is about the same when the context is limited to content words.

Context Defined by Word Frequency. It might be

thought that a more sensitive reflection of the relationship between similarity of context and synonymy would be obtained if we considered only those words of context whose occurrence was conditional on the occurrence of the theme word. Such words would tend to occur in the overlap of two themes only if some significant semantic feature was common to both themes.

In particular, it seems obvious that the higher the frequency of occurrence of a word in the language as a whole or in some reasonably large corpus, the less likely will its occurrence be conditional upon the occurrence of some particular word. This follows from the observation that a word has a high frequency of occurrence usually as the result of its use in a wide variety of situations.

We investigated the relationship between word frequency and amount of overlap by partitioning the content words (nouns, verbs, adjectives and adverbs) into three word frequency intervals: 0–150, 151–1000, frequencies greater than 1001 (occurrences per 4.5 million tokens of the Lorge Magazine Count). Function words (articles, prepositions, conjunctions, pronouns, etc.) were partitioned out as a fourth set (for list of function words, see Miller et al. [5].)

Figure 3 shows that there is only a slight tendency for

TABLE 5. PERCENT CORRECT INFERENCES FOR TWO ERROR LEVELS

Context Definition	*Unleveled*				*Leveled*			
	M_y		M_k		M_y		M_k	
	1%	*5%*	*1%*	*5%*	*1%*	*5%*	*1%*	*5%*
Unrestricted								
CW & FW	75 (34)	95 (32)	75 (54)	95 (52)	70 (38)	90 (36)	80 (57)	85 (55)
CW alone	75 (22)	95 (19)	90 (20)	95 (17)	— —	— —	— —	— —
Lorge Frequency								
0–150	85 (9.1)	90 (6.6)	85 (8.1)	95 (6.4)	70 (8.5)	95 (6.3)	80 (7.5)	85 (5.6)
151–1000	70 (22)	80 (20)	75 (20)	95 (16)	50 (25)	70 (23)	85 (22)	95 (17)
1001 and above	5 (59)	15 (56)	35 (49)	60 (44)	5 (62)	35 (57)	55 (48)	75 (45)
FW	0 (82)	10 (79)	45 (8.1)	55 (80)	5 (82)	5 (79)	35 (83)	50 (81)
Grammatical	90 (15)	95 (12)	95 (13)	95 (11)	— —	— —	— —	— —
Association	90 (14)	100 (4.8)	95 (11)	100 (3.7)	— —	— —	— —	— —

The figures in parentheses are the overlap values associated with the given percentages.

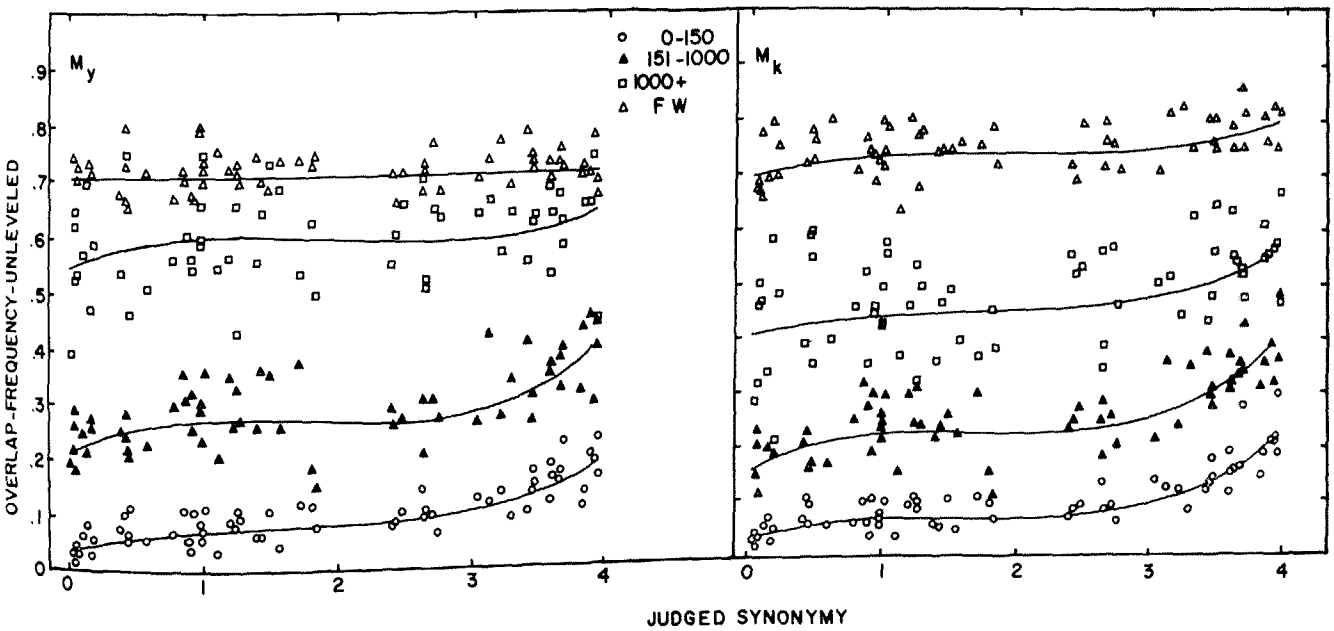

FIG. 3. Effect of partitioning context into frequency classes. Each point represents one of 65 pairs of theme words. The contexts were partitioned into four classes: function words (FW) and three content-word classes. The content-word classes represent three intervals according the Lorge Magazine Count (frequency of occurrence in 4.5 million tokens). Each curve represents the overlap obtained for one of the four partitions of context.

overlap consisting of content words of lower frequency to be more affected by synonymy than content words of higher frequency. As for function words, while the M_y measure shows zero slope (.002 with a standard error of .003) the M_k measure shows a slope which is significantly greater than zero (.017 with a standard error of .004). No explanation is readily available for this surprising result. We guessed that it might have been primarily due to the pronouns and prepositions, which of all function words would seem to have the greatest potential semantic connection to the theme words; however, even this conjecture was not borne out by an analysis of the data.

The results shown in Figure 3 are essentially unchanged if leveled data are used.

The percentages of correct inferences for these four Lorge frequency intervals are shown in Table 5. It is evident that content words with a frequency greater than 1000 are relatively useless for discriminating highly synonymous pairs. The statistical significance of the M_k slope for function words permits as high as 55 percent correct inference with 5 percent error. Apparently the small effect evident in Figure 3 has some inference power although there is no explanation why this is so.

Context Grammatically Defined. In this limitation on context the overlap was restricted to those words which were grammatically most closely related to the theme. Content words which were grammatically dependent upon the theme word, which was always a noun, or upon a pronoun standing for the theme were considered related. Thus the following were included: (1) nouns, adjectives or participles describing the theme, (2) the verb if the theme was the subject (and predicate noun if the verb was the copula or a copula substitute), and (3) the noun in a prepositional or possessive phrase modifying the theme. Also included were the following in the event that the theme word was the grammatically dependent form: (1) the verb of which the theme was the object, (2) the noun to which the theme was the appositive, and (3) the verb or noun modified by a prepositional or possessive phrase of which the theme was the head.

Figure 4 shows the relationship between the overlap of contexts defined in this manner and judged synonymy.

It is obvious that the slope is considerably steeper for this context definition than for either the unrestricted or the frequency-defined context. This difference is reflected in the greater discriminating power of grammatically defined context: percent correct inferences for one percent error, M_y equals 90 percent, M_k equals 95 percent; for five percent error, M_y and M_k both yield 95 percent (all based on unleveled data).

Context Defined by Association. On the notion that we might obtain a relation between overlap and synonymy which was more sensitive to differences in synonymy, especially at the lower end of the scale, we limited the context to those words in a sentence set that were judged to be highly associated to the theme.

To obtain these judgments lists of the word types occurring in each sentence set were prepared. One judge examined the lists for the column A theme words (Table 2) and another judge examined the lists for the column B themes. (The two judges were the first-named author and his wife.)

The judge looking at a list of words that occurred with a given theme asked himself: What is the likelihood that a situation which evoked the theme word would evoke the word type in question, and conversely, what is the likeli-

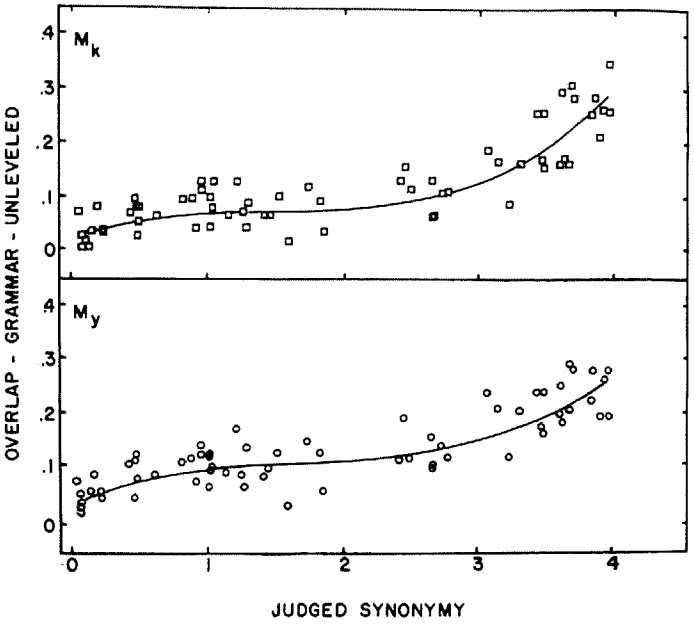

FIG. 4. Effect of limiting the context to words grammatically related to the theme. Each point represents one of 65 pairs of theme words.

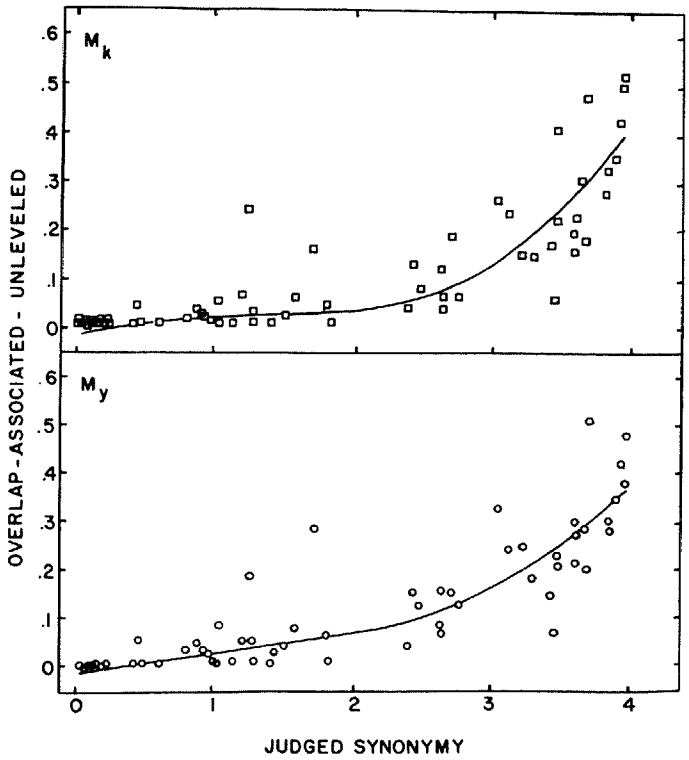

FIG. 5. Effect of limiting the context to words judged as being in high association relationship with the theme. Each point represents one of 65 pairs of theme words.

hood that a situation which evoked the word type would also evoke the theme word? Only if both likelihoods were judged "high" or "very high" were the words considered highly associated. For example, for the theme word *food*, the following words were among those judged to be highly associated: *appetite, appetizing, ate, bread, breakfast, cooked, delicious, digested, digestion, mealtime, nourishment, nutritional, refrigeration, restaurant, taste.*

It is unfortunate that only two judges were available for this experiment. However, having observed a high degree of reliability among a larger group of judges in a similar associative task, we are fairly confident that essentially the same results would have been obtained if a greater number of judges had been employed.

Figure 5 shows that the lower portion of the curves is still rather flat although the curves do rise on the high synonymy end much more sharply than they do under the other definitions of context.

The high discriminating power of associatively defined context is clearly shown by the percentage of correct inferences possible from examination of the overlap measures: for one percent error, M_y equals 90 percent, M_k equals 95 percent while for five percent error both measures yield 100 percent.

In view of the fact that the overlap from contexts defined this way is more discriminating it was considered whether such contexts could be obtained by computer processing. One possibility worth exploring seemed to be that highly associated words would tend to occur close to the theme word. Thus all the content words in the sentence sets were partitioned according to their position in the sentence relative to the theme. The proportion of the content words occurring at each position that had been judged highly associated to the theme was calculated and plotted against position.

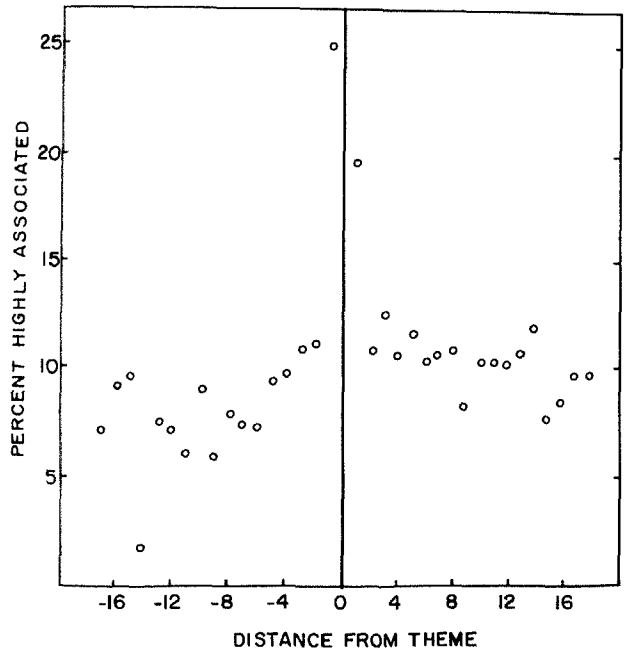

FIG. 6. Proportion of content words at various positions in the sentence that were judged to be in high association relationship with the theme. The negative values on the abscissa are the number of word places *before* the theme while the positive values are the number of word places *after* the theme.

Figure 6 shows only two positions have a distinctly higher proportion of highly associated words: 25 percent of the content words occurring directly before the theme and 20 percent of the content words occurring directly after belonged to the set of words which had been judged to be highly associated to the theme words.

Conclusions

1. The basic hypothesis investigated by the study is corroborated: there is a positive relationship between the degree of synonymy (semantic similarity) existing between a pair of words and the degree to which their contexts are similar.

2. It may be safely inferred that a pair of words is highly synonymous if their contexts show a relatively great amount of overlap. Inference of degree of synonymy from lesser amounts of overlap, however, is apparently uncertain since words of low or medium synonymy differ relatively little in overlap.

The first of these conclusions may be accepted in its full generality since it is, in all probability, a linguistic universal. The second conclusion, however, together with the findings on the effects of various definitions of context and on the particular shape of the overlap versus synonymy function should be accepted only tentatively for they may be properties of the language materials and procedures employed in this study.

Valid generalization of findings like these depends upon an increased knowledge of the effects of such factors as corpus size, degree of homogeneity of content and form and the size of the units in which information is packaged within the corpus.

RECEIVED MAY, 1965.

REFERENCES

1. GIULIANO, V. E., AND JONES, P. E. Studies for the design of an English command and control language system. ESD-TDR-63-673, Arthur D. Little, Inc., Cambridge, Mass., 1963.
2. HARRIS, Z. S. Distributional structure. *Word 10* (1954), 146-162.
3. JOOS, M. Description of language design. *J. Acoust. Soc. Amer. 22* (1950), 701-708.
4. KUHNS, J. L. The continuum of coefficients of association. in M. E. Stevens, L. Heilprin, & V. E. Giuliano, (Eds.), *Proceedings of the Symposium on Statistical Association Methods for Mechanized Documentation*, National Bureau of Standards, US Government Printing Off., Washington, D. C. In press.
5. MILLER, G. A., NEWMAN, E. B., AND FRIEDMAN, E. A. Length-frequency statistics for written English. *Inform. and Contr. 1* (1958), 370-389.
6. STEVENS, M. E., HEILPRIN, L., AND GIULIANO, V. E. (EDS.) *Proceedings of the Symposium on Statistical Association Methods for Mechanized Documentation*, National Bureau of Standards, US Government Printing Off., Washington, D. C. In press.
7. STILES, H. E. The association factor in information retrieval. *J. ACM 8* (1961), 271-279.
8. THORNDIKE, E. L., AND LORGE, I. *The Teacher's Word Book of 30,000 Words*. Columbia U. Press, New York, 1944.