The University of Michigan

Center for Human Growth and Development

A Semantic Featural Approach to Meaning Similarity and Association

by

Charles A. Perfetti

Report Number 18

Development of Language Functions

A Research Program-Project

May 18, 1967

Supported by the National Institute for Child Health and Human Development

Grant Number 5 PO1 HD 01368-03

ACKNOWLEDGMENTS

This report is a dissertation submitted to the University of Michigan in partial fulfillment of the requirements for the Doctor of Philosophy degree.

I wish to acknowledge the valuable assistance of the following people: Dr. David McNeill, for guidance, encouragement and stimulation; Dr. Klaus Riegel, Dr. Arthur Melton, Dr. David Birch, and Dr. J. C. Catford, for scholarly advice and helpful discussions. Dr. Charles E. Osgood, University of Illinois, for making available materials for two of the experiments and for discussing semantics with me; Miss Lynn Cooper, for assisting with the collection and analyses of data; Mrs. Judy Kingsley and Mrs. Janet Bachman for typing and other assistance in preparation of the report.

There were two sources of financial support for this work: Grant Number 5 PO1 HD 01368-03 from the National Institute of Child Health and Human Development and a predoctoral fellowship (4-FLMH-23,088-04) from the United States Public Health Service.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ii

LIST OF TABLES v

LIST OF FIGURES vii

LIST OF APPENDIXES viii

CHAPTER

1. INTRODUCTION: SEMANTIC FEATURES IN PSYCHOLOGICAL PERSPECTIVE 1

2. EXPERIMENT 1: THE MINIMAL CONTRAST HYPOTHESIS 20

   Design and Method of the Experiment 20
   Results 26
   Conclusions 35

3. EXPERIMENTS 2 AND 3: DIMENSION SALIENCE AND REDUNDANCY . . . 43

   Experiment 2 43
   Experiment 3 52
   Comparisons among the Three Experiments 60
   Conclusions 67
   Summary 72

4. EXPERIMENT 4: SIMILARITY JUDGMENTS 74

   Method 75
   Main Results 84
   Additional Analyses 88
   Summary 102

5. EXTENSIONS AND SPECULATIONS 104

   Meaning Similarity and Word Associations 104
   Further Aspects of Semantic Features 111
   Summary 118

APPENDIX Page

I. INSTRUCTIONS TO SUBJECTS FOR FIRST THREE EXPERIMENTS 120

II. FOUR SETS OF SIMILARITY JUDGMENTS AS PRESENTED TO Ss IN EXPERIMENT 4 WITH PERCENTAGE OF Ss SELECTING EACH COMPARISON 123

III. INSTRUCTIONS TO SUBJECTS FOR EXPERIMENT 4 132

REFERENCES 135

LIST OF TABLES

Table		Page
1	Feature Composition of the Artificial Words	22
2	Analysis of Variance Summary for Differences between Words in Rate of Acquisition	28
3	Frequency of Joint Occurrence of Features in Consecutively Recalled Items	30
4	Distribution of Associations in Experiment 1	31
5	Shared Features between Associates	34
6	Frequency of Double and Triple Contrasts for MCR Ss and for Non-MCR Ss.	40
7	Expected Distribution of Frequencies Under Single-Feature Mediation Model Compared with Observed Distribution for MCR Ss.	41
8	Feature Composition of the Artificial Words in Experiment 2	45
9	Summary of Analysis of Variance for Differences between Words in Rate of Acquisition - Experiment 2	48
10	Observed and Expected Frequencies of Joint Occurrence Features in Consecutively Recalled Words - Experiment 2	49
11	Distribution of Associations in Experiment 2	50
12	Feature Composition of the Artificial Words in Experiment 3	54
13	Summary of Analysis of Variance for Differences between Words in Rate of Acquisition - Experiment 3	56
14	Expected and Observed Frequencies for "Clusters" and "Non-clusters" in Recall - Experiment 3	57
15	Distribution of Associations in Experiment 3	58
16	Contingency Table Showing Relative Frequencies of (tangible) and (intangible) Responses to (tangible) and (intangible) stimuli	59

17 Summary Comparison of Overall Acquisition Rates for Experiments 1, 2, and 3 61

18 Two-way Analysis of Variance Summary for Acquisition Rate: Experiments x Words 62

19 Matrix of Significant Differences between Acquisition Rates of Individual Words Averaged over Three Experiments . . 63

20 Comparison of Associative Probabilities for Pairs Words in Experiments 1 and 3 65

21 Summed Associative Frequencies for Words Which Were MCRs in Either Experiment 1 or 3 but Not Both 67

22 Frequencies of Association Sharing One Feature and No Features for MCR Ss Compared with Non-MCR Ss - Experiment 3 . 69

23 Comparison of Expected and Observed Distribution of Non-MCR Responses for MCR Ss in Experiment 3 71

24 Feature Codings on 13 Dimensions 77

25 Mean RS for Fg Word 85

26 Mean RS for Cg Word Compared for Different Levels of Feature Sharing 87

27 Product-Moment Correlations between Four Predictive Variables and RS 92

28 Deviations of Median RS Scores from 50 94

29 RS Comparisons for Words Sharing with Words Contrasting on Pleasant-Unpleasant with Standard 96

30 Eighteen Judgment Items Not in Predicted Direction Coded on Evaluative Dimension 99

31 32 Associations Involving Words from Osgood's List 110

LIST OF FIGURES

Figure		Page
1	Representation of S-R process under single-feature mediation model	38
2	Associative probability as a function of the number of shared features between associates for Experiments 1 and 3	70

LIST OF APPENDIXES

Appendix		Page
I.	Instructions to Subjects for First Three Experiments	120
II.	Four Sets of Similarity Judgments as Presented to Ss in Experiment 4 with Percentage of Ss Selecting Each Comparison	123
III.	Instructions to Subjects for Experiment 4	132

Chapter 1

Introduction: Semantic Features in Psychological Perspective

The aim of this thesis is to explore the usefulness of the concept of semantic feature as a psychological description of semantically related associations, and, more generally, as a structural description of meaning similarity. This chapter will serve to define the concept of semantic feature and place it in psychological and linguistic perspectives.

The Componential Analysis of Meaning

The terms "semantic component" and "semantic feature" will be used interchangeably here. A third term which can be used in this regard is "semantic marker," an expression introduced by Katz and Fodor (1963) in their description of the abstract form of a semantic theory. According to Katz and Fodor, semantic markers are lexical categories ". . . by which we can decompose the meaning of a lexical item (on one sense) into its atomic concepts . . ." (p. 496). Examples of semantic markers from these authors will serve the present purpose: One sense of the word BACHELOR requires that it carry the semantic markers (Human) and (Male).[1]

---

[1] By way of notational convention, parentheses will be used to enclose semantic features, which can be considered values on semantic dimensions. To minimize confusion, dimensions will not be enclosed, although the abstractness of semantic features, denoted by the parentheses, should also be ascribed to semantic dimensions. That is, both features and dimensions are abstract in that they are part of the metalanguage, but are represented by English words.

In the Katz and Fodor system, the semantic markers plus the semantic distinguishers, which are special markers giving a unique sense characterization for a particular word, provide not only a componential description of a word, but also serve to distinguish among senses of the same word. Thus one sense of BALL carries the markers (Social Activity), (Large), and (Assembly), while another sense of BALL carries (Physical Object).

In general, the term "semantic feature" will be used here, and "semantic marker" will be reserved for the specific theoretical context of the Katz and Fodor theory. However, the theoretical status to be given here to the concept of semantic feature is similar to the status Katz and Fodor give to semantic marker: "A semantic marker is simply a theoretical construct which receives its interpretation in the semantic metatheory and is on par with such constructs as atom, gene, valence, and noun phrase" (p. 517).

While the foregoing can be taken as a working justification for the concept of semantic feature, its ultimate utility as a psychologically functioning unit remains an open, and, hopefully, empirical question. The support for the validity of semantic features has heretofore come from linguists and anthropologists, if not from psychologists. Thus, although there is little reference in psychological literature to the concept of semantic feature[2], one finds the assertion from linguist Uriel Weinreich that "It is hardly necessary any more to analyze or to justify the concept of semantic component." (Greenberg, 1963, p. 185). The papers which have given support to this contention all come from anthropologists or lin-

---

[2] There are, however, psychological analogs to semantic features, e.g., "attribute" as it is used in formal and psychological studies of concepts (c.f. Bruner, Goodnow, and Austin, 1956). Sensory associations, e.g., Underwood and Richardson, 1956, may be another analog.

guists--Conklin (1955), Lounsbury (1956), Goodenough (1956), Wallace and Atkins (1960), and Romney and D'Andrade (1964). Weinreich not only sees semantic components as "legitimate units of semantic description," but also suggests that certain semantic components may be universal--that, for example, features such as generation, sex, light vs. dark, dry vs. wet, etc. may be important components of meaning in all languages. Of special interest is the concept of synonymy in such a componential system. Weinreich defines as "immediate synonyms any pair of terms A and A' such that their designata differ by one component." This definition is particularly relevant to the idea of minimally contrasting word pairs, about which more will be said later. Most of the examples offered as illustrating componential analysis have come from highly structured sets of relationship-denoting words, such as kinship terms, and Weinreich appropriately recognizes that the bulk of lexical items is probably characterized by a much looser componential structure.

The work of anthropologists thus far has shown varying degrees of relevance to an individual psychological consideration of semantic components. The papers of Conklin (1955) and Lounsbury (1952) seem not to make a commitment to a cognitive structure of components, but rather emphasize categories of relationship within a cultural group. Other anthropological efforts, however, suggest stronger psychological claims. Goodenough (1956), Wallace and Atkins (1960) and Romney and D'Andrade (1964) apparently all share the view that the semantic components underlying kinship terms, and presumably other meaning relationships, provide real psychological definitions.

Componential Similarity

The kinship paradigm. Since kinship terminology (mostly English, although Lounsbury studied Pawnee) has provided the most frequently used paradigm for componential analysis, a description of this paradigm will illustrate componential analysis for a highly structured meaning system. Wallace and Atkins (1960) and Romney and D'Andrade (1964) have presented formal rules which specify the steps to be taken in Kin term analysis. For the present purpose only an informal description of the resultant analysis of Romney and D'Ardrade is necessary. The eight Kin terms of FATHER, MOTHER, SON, DAUGHTER, UNCLE, AUNT, NEPHEW, and NIECE can be analyzed, as a meaning system, into three componential dimensions: generation, lineage, and sex. Since two features comprise each dimension, the total of $2^3 = 8$ unique feature combinations is exhausted by the eight terms. FATHER would carry the features (ascending generation), (direct lineage), and (male), and would contrast with MOTHER only on sex, with SON on generation, and with UNCLE on lineage. Each term accordingly belongs to three sets of feature contrasts, each contrast distinguishing it from a different subset of terms. This paradigm is not the only possible one, and, in fact, a slightly more complicated one is necessary when additional terms, for example COUSIN, are added to the system. Romney and D'Andrade acknowledge that alternative analyses would lead to different psychological implications.

Exactly what are the psychological implications of such a componential analysis? Romney and D'Andrade make the general prediction that the more components any two terms have in common, the greater will be the similarity of response to these terms. This prediction is based on the implicitly psychological assumption that ". . . the components of a term

constitute the meaning of that term for an individual; hence the more components which are shared the more similar the meaning" (p. 154). The sharing of a component can occur on two levels: as a dimension, e.g., sex and lineage, and as a value (feature) on a dimension, e.g., (male) and (direct).

Romney and D'Andrade tested the psychological implications of this system with three experiments. In the first, subjects were required to list all the names for kinds of relatives and family members they could think of. Of main interest was the marked tendency for these lists to contain successive occurrences of terms contrasting only on the sex component, e.g., father-mother, son-daughter, uncle-aunt. These pairs were found in 98 percent of the cases in which both terms were given.

A second test was a semantic differential rating on the kinship terms. Here the hypothesis was that subjects would respond to the componential meaning of the terms, so that, for example, if subjects responded to the female component of WOMAN and SISTER on potency scales (e.g., heavy-light), then the potency scores for the two terms should be correlated. The results of this experiment were somewhat equivocal, with enough consistent groupings of terms under factors to lend support to the hypothesis, but with enough irregularities to be somewhat unsatisfactory. The main inconsistency was that "young" terms (descending generation) did not emerge as a factor, but occurred as two sex-separated factors--as if, the authors suggest, the components of youth and sex interacted to form a unique connotative meaning.

While the results of this second experiment were ambiguous, the results of the first experiment are open to another interpretation. The successive occurrence of sex-contrasts might merely reflect associative

strengths of the words, rather than component contrasting. FATHER - MOTHER, SON - DAUGHTER, etc. undoubtedly represent strongly associated word-pairs whose associative strength may be derived more from frequent co-occurrence than from component similarity. The result of the third experiment may be less equivocal. Subjects were presented with sets of kin term triads and asked to mark the term which was most different in meaning. By presenting all possible triads of the kin terms, Romney and D'Andrade could take the number of times each pair of terms were left together as a test of their similarity prediction. The data were clearly supportive for their prediction. Thus, for example, FATHER - MOTHER, and SON - DAUGHTER were classed together with very high frequency. Since sex-contrast produced pairs in this sorting task, the procedure could be extended to a set of same-sex terms--GRANDFATHER, FATHER, UNCLE, BROTHER, SON, GRANDSON, and COUSIN. Here it was found that selections were regularly made according to components of generation and lineage. The interpretation of this experiment was that the shared components between words accounts for their perceived similarity.

Potential and psychological similarity. At this point, it is appropriate to take note of the distinction drawn by Wallach (1958) between potential and psychological similarity. Potential similarity is defined by the physical attributes shared by a set of events, while psychological similarity is defined by an individual's employment of a rule for categorizing the events. The rule specifies which attributes must be conjointly present in two events in order for the individual to assign them to the same class, and, at the same time, specifies, or at least implies, the construction of contrast classes for this categorization. Psychological similarity thus depends upon some aspect of potential similarity, but the

particular aspect selected depends upon the contrast being constructed and the accompanying rule for assignment. Hence, the assignment rule operates to select what have been termed criterial attributes (Bruner, Goodnow and Austin, 1956) from among the physical or potential attributes. Although this description has been applied to concepts rather than to meaning, and although the two are not identical, word meaning may be considered a limiting case of the general form of the concept (Weinreich, 1963); hence, the foregoing description ought to be relevant to a componential analysis of meaning. In such an analysis the relevant attributes for similarity are to be found both in features of the word (grammatical components) and in features of the designata (semantic components).

In the kinship paradigm, there is potential similarity among the designata for all terms. The psychological similarity depends on the contrast being formed. When the contrast is made on the basis of lineal proximity, MOTHER and FATHER psychologically are more similar than are UNCLE and FATHER; when the contrast is made on sex, UNCLE and FATHER are more similar. When an individual selects one pair as more similar than another pair in the absence of an explicit rule, he is presumably employing an implicit rule which selects the criterial attributes. To the degree that individuals agree on the relative similarity of words, there is a consensus on attributive criteriality. Presumably there is a greater consensus for highly structured items such as kin terms than for the more loosely structured items in the general lexicon.

Although a componential analysis can apparently be fruitfully applied to the type of highly structured meaning system exemplified by kinship terms, whether it can be applied to the general lexicon may be another question. Moreover, the degree of literality to ascribe to a semantic

feature and the psychological process described to operate on it may pose serious questions. Justice is not done to the historical depth of these problems by posing them as new questions. In short, is componential structure an anachronistic reversion to the elementalism which plagued 19th Century scientific psychology and, before that, the philosophical efforts of the British associationists? While a lengthy discussion of philosophical associationism would be a digression, it is nonetheless well to recall the extremes to which associationism was driven by, particularly, James Mill. For Mill, an idea was a fusion of associations. All ideas, in fact, were seen as compounds of associated components: The idea of WALL, to use Mill's own example, is a complex idea composed of two component complex ideas, BRICK and MORTAR. Although Mill was concerned with ideas, it is clear that the analysis was also applicable to meanings. Hence, to avoid absurdity, the interpretation of semantic features, particularly of their combinatorial properties, must be less literal than that given by Mill. Moreover, any psychological process which might be proposed to operate on these components must not imply that the human subject responds to the word WALL by decomposing it into BRICK and MORTAR. At the same time, it must be recognized that Mill's componential analysis is quite different from the type discussed here. The distinction between the two is analogous to the familiar distinction between the physical components of color (wave-length mixtures) and the psychological attributes of color (e.g., brightness).

Componential Analysis of Word Association

The psychological implications of semantic componential analysis with which this investigation is primarily concerned are its possibilities as a description of psychological meaning similarity and, particularly, as an

approach to semantically related word associations. This section is a brief discussion of a selection of relevant literature.

Other semantic approaches to word associations have been attempts to uncover patterns of relationship between stimulus and response, a task which is shared by componential analysis. However, in contrast to componential analysis, the previous attempts have usually involved the classification of response words according to some physical, abstract, or syntactic relationship presumed to exist between their designata. For example, TABLE and DISH have the physical relationship of contiguity, DARK and LIGHT have the abstract relationship of contrast, while BOY and RUN have the syntactic relationship of noun-verb. This approach characterized the early efforts of Woodrow and Lowell (1916) which were the most comprehensive of many similar investigations. More recently, Riegel (1965) and Jenkins and Palermo (1964) have reversed the procedure of classifying responses. Their subjects give responses under categorical restrictions to produce contrasts, similars, verbs, parts, etc., to stimulus words. This procedure eliminates the arbitrariness of classifying responses. Its significance, however, may be more as a method of structuring word relationships than as a semantic analysis of associations, because subjects are providing elements of an a priori relationship rather than simply associating. A study which applied this method to meaning similarity will be described later.

Other semantic approaches to word associations are concerned, not with the S-R relationship, but with the nature of meaning as a response and its role in associations. Thus Bousfield has described meaning as implicit associations, while Osgood has described associations as implicit meaning responses (see Cofer, 1961). The theoretical controversies within the be-

havioristic approach, as well as the larger ones between behavioristic and linguistic approaches, as represented by the arguments of Chomsky (1959) and Fodor (1965) against the approaches of Skinner and Osgood, respectively, can be ignored at present. (For a recent exchange on the basic issues see separate articles by Osgood, Berlyne, and Fodor in The Journal of Verbal Learning and Verbal Behavior, 1966, 5, No. 4.)

An example of an experiment which is interpreted within the S-R framework is one by Pollio (1964). Subjects associated continuously to four words for four minutes. It was observed that responses tended to occur in bursts, and that the resultant clusters of words tended to be related. Words within a cluster were words which frequently evoked common responses (high associative overlap) and which were similar in connotative meaning. The results were interpreted as evidence for ". . . the existence of associative clusters . . . of strongly interconnected words, each of which evokes essentially similar meaning responses."

The relationship between meaning and association is also an important aspect of the work of Deese, Bousfield, Cofer, and others. The title of Deese's paper, "On the Structure of Associative Meaning" (Deese, 1962b) reveals the view that part of the psychological meaning of words can be measured by word associations. Deese's method produces a measure based on the percentage of associative responses given in common to two words. Ten methods of measuring word relatedness by means of associations and associative overlap have been reviewed by Marshall and Cofer (1963). Most of these measures have proved to be valid predictors of some associatively significant performance variable, such as clustering in free-recall and ease of paired-associate learning. This body of literature represents, not a semantic approach to associations, but an associative approach to semantic similarity.

A previous study by the author (Perfetti, in press) provides an example of an attempt to combine the method of word associations with the study of similarity. Briefly, the method involved the application of the restricted association tasks of Riegel (1965). Subjects gave superordinate, similar, contrast, part, function (verbs), and quality (adjectives) responses to noun-stimuli. Measures of associative overlap were taken across all tasks. The similarity of two words was assumed to be measured by the percentage of common and mutual responses evoked across selected tasks; i.e., the extent to which two words evoked the same superordinate, part, function, etc., was taken as an indication of their relative similarity. This approach is thus related to the present one in that it assumed associations reflect underlying semantic structures. The difference is that while both approaches use the notion of commonality, the former does so at the level of semantically related words, the present does so at the level of semantic components of words. Even this difference, however, is less than might be supposed. For example, the previous study showed superordinate overlap for TABLE and CHAIR--both evoke FURNITURE; certainly in a componential analysis one of the semantic components of both words would be (furniture). In general, the semantic components which most often appear as illustrations in the literature, and components which would probably most frequently be intuited by speakers of the language, at least for nouns, often seem to correspond to the relationships of superordinate and quality (adjective).

While most previous semantic approaches to association have only hinted at the notion of componential analysis, there are one or two which have explicitly considered it; there are, however, no known empirical investigations. A paper by Saporta (1959) discusses the possibilities of

semantic componential analysis, and much of the following discussion is drawn from this paper.

The semantic analysis of words into morphemic components can be considered linguistically analogous to the phonological analysis of word sounds into articulatory components, and, in particular, is quite comparable to the distinctive feature analysis of Jakobson, Fant, and Halle (1952). Saporta illustrates the semantic analysis with the word BOY, the meaning of which can be described as a combination of the components (young), (male), and (human). BOY and MAN share two of these components (male) and (human), but differ on one, (young) vs. (old). BOY and GIRL also share two components and contrast on a third. On the other hand, BOY and WOMAN share only one component, (human), while contrasting on the other components.

Saporta suggests that the kind of analysis described above may be applied to at least some kinds of word associations, and points out that many responses in word association tests differ by one component from the stimulus word--hence called by McNeill (1966) minimal contrast responses. Association data from the Russell and Jenkins norms are cited as support. The frequencies of association to BOY were as follows: GIRL 319, MAN 104, WOMAN 2; and to MAN: WOMAN 394, BOY 44, GIRL 6. In both cases the two most frequent responses contrast on one component while the least frequent contrasts on two components. It is obvious, as Saporta recognizes, that if the componential analysis is to work, it must allow for a hierarchy of components, otherwise GIRL and MAN should occur about equally often to BOY. For these words, the primary associations contrast on the (sex) component.

The minimal contrast hypothesis. The foregoing proposed analysis can take the form of a more explicit hypothesis: word associations tend to be minimally semantically contrasting word pairs, or, more generally, the

probability of two words being given as associates is an inverse function of the number of their contrasting features. A reasonable modification of this hypothesis would add some restriction concerning the number of shared features. For example, it seems necessary that a pair of associates share at least one feature.

An important question concerns the domain of associations to which the minimal contrast hypothesis, or componential analysis generally, can be applied. Although the answer to this question is not really available at this point--it is perhaps enough to ask whether this hypothesis can account for any association--nonetheless, there seems to be at least one constraint which can be drawn in a preliminary way. This constraint is that the minimal contrast hypothesis seems more applicable to associations belonging to the same grammatical form class (paradigmatic associates) than to associations belonging to different grammatical form classes (syntagmatic associates). This constraint has previously been suggested by McNeill (1966). There appear to be two possible related reasons for this assumption: (1) Words of the same grammatical form class are more likely to have semantic features on comparable content dimensions. (2) Words carry grammatical features as well as semantic ones, and the grammatical features probably serve a similar marking function and thus can be an important source of contrast which reduces the likelihood of their association. These considerations are reflections of the functionally blurred distinction between syntax and semantics and the problem of whether the so-called "parts of speech" represent form or content classes. (It may be noted that even the most sophisticated attempts to describe the rules of English grammar cannot make an easy distinction between features that belong to the rules of the grammar and features that belong to the meanings of words [Chomsky, 1965].)

The assumption that paradigmatic associations are the most appropriate domain for applying the minimal contrast hypothesis leads to a semantic featural analysis of the "paradigmatic shift" (McNeill, 1966). "Paradigmatic shift" is used to describe the observation that older children and adults give a much higher percentage of paradigmatic responses than do children under seven (Ervin, 1961; Entwisle, Forsythe, and Muuss, 1964; Entwisle, 1966). The semantic featural analysis of this shift, as suggested by McNeill (1966), hypothesizes that younger children give few paradigmatic responses because their knowledge of the featural properties of words is incomplete. Since children above the age of four have largely adult grammars (Brown and Bellugi, 1964; Brown and Fraser, 1964; Ervin, 1964), their tendencies toward syntagmatic responding does not seem attributable to a failure to include grammatical form class as a feature of words. It is more reasonable to assume that beyond this age most of the child's word learning involves the addition of appropriate semantic features. However, until such time as his words become more richly marked, thus leading to an increase in paradigmatic responding, he is giving syntagmatic responses which may actually represent minimal contrasts. Such associations are thought of as coming from a set of words which have only a very small number of feature specifications. With so few semantic features the set of possible minimal contrasts would include words of different form classes. These kinds of associations have been called pseudo-syntagmatic (McNeill, 1966), because they can be considered to represent minimal contrasts.

This analysis, however, is quite consistent with the attempt to restrict the domain of the minimal contrast hypothesis to paradigmatic associations, because the implication is clearly that <u>real</u> syntagmatic associa-

tions, unlike pseudo-syntagmatic, are not minimal contrasts, either for children or adults. This account is largely speculative, to be sure, but it serves to illustrate the possible explanatory power of the minimal contrast hypothesis. Moreover, there may be some indirect evidence for this account in the finding by Entwisle (1966) that most of the paradigmatic shift can be attributed to a decrease in noun-verb associations, whereas noun-adjective associations actually show a small increase. McNeill (1966) has argued that this is consistent with the semantic feature account, if it can be assumed that nouns and verbs tend to be more semantically related than nouns and adjectives, and that, for children with incomplete featural knowledge, they may serve as minimal contrasts. Whether this assumption is justified is another matter for speculation.

The main point is that regardless of whether the semantic feature analysis can account for the paradigmatic shift, some constraint can be placed on the kinds of associations which this analysis can account for. True syntagmatic responses seem not to be minimal contrasts. (This is not to say, however, that some sort of feature matching is not involved in syntagmatic responses. These could represent matching selection restrictions with semantic features [McNeill, 1965].) Other kinds of associations undoubtedly reflect simple sequence habits, and are not to be explained by semantic features. Associations such as BOY - SCOUT and LAMP - SHADE presumable are examples of this kind of association, although both stimuli also have high frequency paradigmatic responses.

Which type of response is given to a particular stimulus is to some extent a reflection of individual differences. For example, Cook, Mefferd and Wieland (1965) have shown that contrast-logical responding is a consistent characteristic of the subject. (Similar conclusions come from

Moran, Mefferd, and Kemble [1964]). Across stimuli, the probability of paradigmatic, and hence minimal contrast, responses varies with stimulus form-class and frequency of usage. It is well established that nouns evoke a higher percentage of paradigmatic responses than other form classes (Deese, 1962a; Entwisle, Forsythe, and Muuss, 1964; Fillenbaum and Jones, 1965). The frequency effect holds primiarily for adjectives, according to Deese (1962a), who found a substantial negative correlation (-.40) between frequency of stimulus usage and frequency of syntagmatic responding. In addition there was a very high positive correlation (+.89) between frequency of usage of adjective stimuli and the frequency of occurrence of contrast responses. In another study (Deese, 1964), 29% of all common English adjectives (frequency of usage > 750 per million) form pairs of contrasts which evoke each other in association (e.g., HOT - COLD, BIG - LITTLE, ABOVE - BELOW). It appears as if adjectives which occur infrequently in the language are likely to evoke as syntagmatic responses those nouns which tend to be associated with their rare occurrences. Hence, according to Deese, pairs of words such as administrative decision occur as associates, because the variety of contexts for the adjective is low, and a noun which monopolizes its context will form a syntagmatic associate. A more frequent adjective is more likely to have occurred in more varied contexts and is then more likely to be associated with other adjectives which can share these contexts. Moreover, it seems characteristic of English that nouns, even rare ones, have a more superordinated structure than do adjectives. The hierarchical arrangement of nouns presents opportunities for paradigmatic partners that are somewhat lacking in the case of adjectives, especially rare ones.

In summary then, associations which are most amenable to semantic featural analysis are associations of the same grammatical form class. And the words which are most likely to be in this category are nouns and high frequency adjectives. These kinds of words appear to be more schematic and thus more like the highly structured kin terms to which componential analysis has been successfully applied.

The Present Experiments

Several problems arise when one considers the question of an experimental test for the minimal contrast hypothesis. The major problem is how to specify semantic features for English words--for example, whether to use empirical or intuitive methods, whether to select broad samples of words or words with more schematized feature composition, etc. Moreover, the features sufficient to describe a given word depends on the set of words being considered. To draw a featural distinction between BOY and DOG we need only specify for BOY the feature (human), but to draw a distinction between BOY and MAN a different feature is required. There would be no way to satisfy ourselves that a particular experimental outcome were not the result of an incorrect feature specification. Thus, an analysis of English word associations into semantic features was considered too adventurous prior to a more straightforward test of the feasability of the minimal contrast hypothesis, if such a test could be made.

Such a test may be accomplished by resorting to artificial materials. With artificial words the experimenter can arbitrarily specify the "semantic" features, require the subject to attain a mastery of the "meaning" of the artificial word, and then study the role of the imposed features on associations comprised of these artificial words. This was the basic experimental procedure employed in the first three experiments. With the

explicit acknowledgement that the way in which subjects handle artificial materials is not necessarily the way in which they handle English material, it nevertheless seemed reasonable that the minimal contrast hypothesis could be given at least a sufficiency test. That is, under highly constrained laboratory conditions using artificial words, it could be tested whether minimal contrast was sufficient to predict associations. In this situation there would not be the usual rival sources of association which are present in real language--for example, no S-R contiguity established through syntactical arrangement in sentences. Hence, the minimal contrast hypothesis could not be tested vis a vis some other hypothesis of association. Rather, it could be tested as a sufficient explanation in the absence of other possible explanations.

It should be noted that in using artificial words these experiments are not ignoring the warnings that have been sounded against regarding such materials as "pristine" and without meaning (Melton, 1961; Underwood and Schulz, 1961). It is known that subjects use their experience with English to enrich the tedious poverty of the nonsense syllable, that they do not always use the entire stimulus as specified by the experimenter, and that they employ mnemonic devices which are not consistent with the experimenter's assumptions about his material. In short, with respect to these experiments, there perhaps is little reason to be confident that subjects will respond to and utilize exactly those artificial words and features which are experimentally imposed. They may substitute a different functional stimulus for the nominal artificial one, and they may supplement the imposed features with the additional features of the functional stimulus. Since we could discover no solutions to these problems, we will defer their further consideration until the discussions of the results. At

this point it is enough to note that the alternative of English words presents other serious problems, and that the effect of the problems of artificial words on an experiment depends on the experimental purpose. It is a serious problem when the interest is in aspects of the stimulus, as in the role of stimulus meaningfulness in paired-associate learning; but it is less serious when the interest is in associations, where there is no marked or systematic way in which these factors can be presumed to bias the hypothesis as long as certain controls--e.g., letter repetitions between words which could lead to associations--are instituted.

The first three experiments use artificial materials in the manner mentioned above. Experiment 1 is a straightforward test of the minimal contrast hypothesis. Experiments 2 and 3 explore some of the effects of the structure of semantic features and dimensions on associations. In Experiment 2, the interest is in the role of dimension salience. In experiment 3, the interest is in dimension redundancy. Experiment 4 is a departure into English words and an investigation of the feasibility of using componential analysis as an approach to similarity, in particular whether the shared and contrasted features can predict subjects' judgments of word similarity. A fifth experiment extends feature analysis of associations to English words, in a strictly exploratory manner.

Chapter 2

Experiment 1: The Minimal Contrast Hypothesis

The first experiment was a test of the sufficiency of the minimal contrast hypothesis under conditions in which other potential determinants of association were absent.

The experimental condition was far removed from a natural language situation, not merely because the words were artificial, but also because the acquisition of the features was artificial. It seems reasonable that in the course of language acquisition, features are learned by linguistic context and by extra-linguistic information. The feature (round) for the word BALL may be acquired through repeated observation of ball-instances, while the feature (human) for the word THINK is more likely acquired through linguistic contexts--i.e., subjects of the verb TO THINK tend to have the feature (human). Regardless of how a particular feature is learned for a particular word, there presumably is a well-developed linguistic context for its expression. A lack of such linguistic context is a major difference between the present experiments and normal language.

Design and Method of the Experiment

Data from pilot subjects clearly demonstrated limits on the learning of features to artificial words. It was found that it was impossible, with the method proposed, to learn four features for eight words, where each feature was a value on a dimension. The number of six words and three features was finally settled upon as providing the optimum design. With

fewer items, not enough data could be collected, and with more items the task was too long.

Selection of Words and Features

The artificial words were chosen from a list of materials constructed by the author for general experimental use. The use of available lists of nonsense material was rejected on several grounds. It was thought that the material should be relatively "meaningless," but pronounceable, and all of the same length and without overlapping graphemic distributions. While lists of paralogs calibrated for meaningfulness (e.g., Noble and Parker, 1960) were available, they did not meet the criteria of equal length and non-overlapping letter distributions. Thus a list of 18 two-syllable words was constructed, each having five letters without repetition. Each word was comprised of a consonant-vowel-consonant-vowel-consonant sequence. No two words had the same consonant in the same position. Eighteen consonants (all except Q and X) were each used once at each position, and six vowels were used three times each at the two vowel positions. From this list of eighteen, six were randomly selected for use in the present experiment, while the remainder were designated for use in another experiment. The six were MAPOK, NYJIB, JUWEN, KEBYR, POGUL and VIKAF.

The main consideration in the selection of the features to be imposed on the artificial words was that they be relatively unlikely to co-occur as features of real words. Thus features such as (human), (male), and (young), which would constitute the featural meaning of a real word, viz. BOY, were to be avoided. In this manner the chances that many Ss would substitute the same English words for the same artificial words were minimized.

Another consideration in selecting the features was that they be pairs of adjective contrasts, since such pairs seem to come closest to representations of semantic features and would present opportunities for real contrast. The final selection of features was made from a list of about a dozen pairs, some of which had been obtained from the semantic atlas of Jenkins, Russell, and Suci (1958). The three feature dimensions thus chosen were tangible - intangible, wet - dry, and public - private. The pairings of features with words are shown in Table 1.

Table 1

Feature Composition of the Artificial Words

Word	Dimension					
	Tangible - Intangible		Dry - Wet		Public - Private	
Vikaf	+	0	0	+	+	0
Nyjib	0	+	0	+	+	0
Kebyr	0	+	0	+	0	+
Mapok	0	+	+	0	0	+
Pogul	+	0	+	0	0	+
Juwen	+	0	+	0	+	0

Both ends (features) of each dimension occurred equally often and in combination with every other feature. If all possible feature combinations had been used, there would have been $2^3 = 8$ words. Hence, the combinations of features were not exhaustive.

Experimental Tasks

The minimal contrast hypothesis was to be tested by the associations given by Ss to the artificial stimuli. The first task, therefore, was a

modified paired-associate learning task to establish the relationship between the words and the features. The second task required Ss to name the features for each word, and was, therefore, essentially a test of the success of the learning procedure. The third and fourth tasks provided the observations on the features. The third was a task of free recall of the artificial words and the final task was the association task. In the order of free recall there was the opportunity to observe the organization of the word lists according to the common features of the words involved, and in the associations there was the test of the minimal contrast hypothesis.

Subjects

The experimental subjects were forty female undergraduates who served without pay as part of the requirement for the introductory psychology courses at the University of Michigan.

Procedure

The experiment was conducted individually in the laboratory of The Language Development Program of the University of Michigan Center for Human Growth and Development. Most Ss completed all the tasks in less than an hour.

The experiment was introduced to Ss as being "concerned with certain processes involved in learning abstract concepts in an artificial situation." Subjects were told that the concepts were artificial words and that they were to learn the "experimentally imposed attributes of each concept." The complete instructions, as read to Ss, appear in Appendix I.

Before beginning the tasks, E pronounced each of the artificial words for S, and then required S to pronounce them. The vowel sound for the

first syllable was always pronounced "long," while the vowel sound for the second syllable was "short." Stress was always given to the first syllable, e.g., MĀ´POK.

The learning task. The materials for the first task consisted of two stacks of 3" x 5" cards. One stack (anticipation cards) had "feature questions" typed on one side and "feature answers" typed on the reverse side. The second stack (study cards) had feature statements for each artificial concept.

The study cards were shown to Ss before the task began. On each of these cards appeared statements giving feature information for an artificial word. Feature information consisted of one statement for each feature. One card, for example, had the following three statements, one above the other:

A MAPOK IS PUBLIC.

A MAPOK IS TANGIBLE.

A MAPOK IS DRY.

The subject saw one study card for each concept-word. The order of dimensions on the study cards was varied, so that another card, for example, had its statement on the public - private dimension second, and still another card had this dimension statement third.

The anticipation cards had the artificial concept-word plus a feature question on one side, and the answer to the question on the reverse side. For example, the question side of the card was:

MAPOK

WET or DRY?

The answer side accordingly stated

A MAPOK IS DRY.

Since there were six artificial words with three features each, one complete trial-block consisted of 6 x 3 = 18 anticipation trials.

The subject was shown the cards by E, who sat across the table. The subject pronounced the artificial word, then responded with one of the two features presented in the question. The card was then turned over by E, and S read aloud the correct statement from the reverse side, regardless of whether his response had been correct. This procedure meant that for each block of 18 anticipation trials S pronounced every artificial word six times (3 cards x 2 readings) and was thus well practiced in saying these words before the association task. The responses of the subject were recorded by E. At the end of each complete block of anticipation trials, the anticipation cards were shuffled by E to minimize order effects in the learning. After every two blocks of anticipation trials, S was presented with the study cards and given up to thirty seconds to study the stack of six. The subject was not permitted to look at more than one card at a time. This schedule was carried out until S had gone through two complete trials without an error.

Post-learning tasks. Immediately after Ss attained the learning criteria, they were asked to give all the features for each artificial word. In this task S had to name the features without the aid of forced-choice questions. This task was seen as a test of the availability of the features. It was possible to have achieved learning criteria without having attained sufficient mastery of the featural composition of the words. Ss who failed in naming any feature for any word were shown the study cards again. If an error occurred again, S's performance on the subsequent recall and association tasks was not included in the data analysis.

In the next task, S was asked to recall the artificial words in any order. The order of emission of the words was recorded by E.

The final task was the association test. Subjects were told that they would hear, one at a time, the various artificial words and that they were to respond with another artificial word. They were told that there was no right or wrong answer and that they were to respond with the first artificial word, other than the stimulus, that "came to mind." About half the Ss were further instructed to respond as quickly as possible, while the other half were not given this instruction. Nine different stimulus orders were used in this task.

Results

The results for all tasks are based on data from thirty-three Ss. Data from seven Ss were discarded for one of two reasons: (1) failure to attain a mastery of the word-feature relationships as indicated by errors in the task of recalling features for each word, or, in one case, as indicated by a failure to reach learning criteria during the anticipation trials; (2) admission by S during post-experimental questioning that, during the association task, the responses were given from a "mental list" without respect to the stimulus word.

Learning Data

There was considerable variation in the ease with which Ss learned the features for the concept-words. Four Ss missed the anticipation items only on the first trial, and hence completed the task after three trials. At the other extreme, there were three Ss who were responding incorrectly as late as the thirteenth trial. The number of trials up to, but not in-

cluding, the criterion trial ranged from 2 to 17. The mean number of trials to reach criterion was 6.94, with a standard deviation of 3.55. The median for this variable was 6.00. These data can be summarized as showing a great variability in the rate of learning, with most Ss reaching criterion sooner than average, but with a number of Ss reaching criterion quite slowly.

The question of whether some concept-words were more easily learned than others was explored by means of a similar measure. For each concept-word the number of the trial which was the first of two consecutive errorless trials was tallied for each S, provided errors did not occur for that concept-word on more than one subsequent trial. (The chance probability of two consecutive errorless trials for a given concept word is $(.5)^6 = .012$.) When this measure was averaged across Ss, the results were as follows, arranged in ascending order: POGUL, 3.85; VIKAF, 3.94; JUWEN, 4.00; MAPOK, 4.64; KEBYR, 4.79; NYJIB, 5.12. The results of a one-way analysis of variance for non-independent samples (repeated measures) are presented in Table 2. The F ratio attained borderline significance ($p < .05$). The Newman-Keuls Comparison Test (Winer, 1962) showed that the largest differece, that between POGUL and NYJIB, barely failed to attain the .05 significance level.

It may be significant that the three most quickly learned words were all (tangible), while the three most slowly learned were all (intangible). Since (intangible) in combination with features such as (wet) or (dry) yield impossible concepts, this may suggest the Ss were attempting to learn real concepts and were hindered by such impossibilities. However, this had no apparent consequence for either the recall or association tasks, as we shall see later.

Table 2

Analysis of Variance Summary for Differences between Words in Rate of Acquisition

Source	SS	df	MS	F	P
Between Ss	915.23	32			
Within Ss	629.83	165			
Between words	46.27	5	9.25	2.53	.05
Residual	583.56	160	3.65		

Free Recall Data

Since it was necessary to recall only six items, only two Ss failed to name all the concept-words. The main interest in the recall data was in the order of emission of items, in particular whether the dimensional structure of the concept words would be reflected by the joint occurrence of items with common features. That is, we were interested in whether we would observe, with the artificial words, the type of "clustering," originally reported by Bousfield (1953) and Jenkins and Russell (1952).

Since the possible sources of clustering were the dimensions of the concept-word, the joint occurrence of the various dimensions was taken as the measure of clustering. A cluster was defined as the successive recall of two words which shared at least one feature. The dimensions which were responsible for a given cluster were recorded for each subject, irrespective of the particular words producing the cluster. For example, for a S who recalled, in the order, NYJIB, MAPOK, POGUL, KEBYR, VIKAF, and JUWEN, the dimensions were recorded as t, dp, p, d, tp, where t stands for the tangibility dimension, d, for the dry - wet dimension, and p, for the public - private dimension. The first two words shared a feature on t, the second and third words shared on both d and p, etc.

It was difficult to obtain an indication of the overall clustering tendency, because the chance probability of clustering was extremely high. With a cluster defined in the described manner, the expected number of mean clusters, based on random sequences of responses, was 4.00 (σ = .89). In other words, if S were responding randomly she would still be expected to give sequences which reflect shared features between successive responses four times out of five opportunities (The maximum number of clusters = n - 1). The mean number of observed clusters was 4.20, which does not reach statistical significance when tested as a sample mean against the expected value resulting from random responding. A mean of 4.30 is required for a two-tailed probability of less than .05, while the probability associated with the obtained mean is less than .20. Thus the evidence for a general clustering effect is inconclusive, although the extremely high chance level may have operated to suppress the effect.[1]

Negative results also were obtained when the recall data were analyzed separately in terms of the features which jointly occurred in successively recalled words. No differences were found in the frequencies of occurrence of the various features and feature combinations, all of which were equiprobable. The relevant data are shown in Table 3 with the previously introduced dimension symbols, e.g., t for (tangible).

---

[1] Another factor which may have operated to diminish clustering was the administration of the feature recall task immediately prior to free recall. The order of presentation of the artificial words by E in the former task may have influenced the subsequent free recall order of S. Since the order of presentation was random, any influence on free recall order would be in the direction of randomness, and hence would reduce clustering.

Table 3

Frequency of Joint Occurrence of Features
in Consecutively Recalled Items

Joint Feature Dimension		Frequency		Test of Differences
Double Contrast				
	t	21		
	d	21		
	p	28	(70)	$\chi^2 = 2.36$, df = 5
Single Contrast				p > .78
	dt	21		
	dp	21		
	tp	24	(66)	

The most important implication of this analysis is that, since there were no differences among features in their frequency of clustering, there was no uniform dimension hierarchy. In particular, there was no tendency toward (tangible) and (intangible) clusters, a tendency which might have been expected in view of the learning data, which suggested that (intangible) concepts were more difficult than (tangible). Futhermore, there was no minimal contrast clustering. The total number of single-contrast clusters was 66, compared to 70 double-contrast clusters, the two types having equal probabilities.

Association Data

Each of the artificial words had two other words with which it formed a minimal contrast. Thus, the chance level probability of a minimal contrast response (MCR) was .40, assuming independence. This high chance level was the consequence of imposing a symmetrical (independent) dimensional structure. This was believed to be preferable to the alternative of varying the dependence between features, i.e., allowing correlations between features.

Thus the test of the minimal contrast hypothesis had to be made against this probability of .40. When the associations for all Ss were totalled, 54 percent were found to be minimal contrasts. This figure, while perhaps not large, was very highly significant ($p < .001$) when tested by the normal approximation to the binomial. Table 4 presents the distribution of associations given by the thirty-three Ss used for this analysis.

Table 4

Distribution of Associations in Experiment 1
(Parentheses denote frequency for minimal contrast pairs)

Stimulus	Response						Percent MCR
	Juwen	Pogul	Mapok	Kebyr	Nyjib	Vikaf	
Juwen	-	(12)	3	3	6	(9)	64
Pogul	(9)	-	(9)	3	4	8	55
Mapok	5	(11)	-	(6)	6	5	52
Kebyr	7	6	(9)	-	(7)	4	48
Nyjib	4	5	6	(10)	-	(8)	55
Vikaf	(8)	3	8	5	(9)	-	52
Total Responses	33	37	35	27	32	34	
Mean Percent MCR							54*

*$p < .001$

The numbers in the cells are the frequencies with which a given stimulus evoked a given response, with the frequencies for minimal contrast pairs denoted by parentheses. Notice every stimulus evoked a minimal contrast as its modal response and every stimulus evoked minimal contrasts more frequently than expected by chance. Thus, both the primary response

for every stimulus and the "primary" stimulus for every response formed minimal contrast pairs. The proportion of MCRs ranged from .48 for KEBYR to .64 for JUWEN.

The overall result apparently reflects a good deal of variation among Ss. There were 22 Ss who gave MCRs at least half of the time, and 11 Ss who did not. That is, one-third of the Ss can be said to have responded on some basis other than minimal contrast.

The analysis of the association data assumed independence among the associations, i.e., that each response had an MCR probability of .40 independent of S's other responses. The instructions to Ss permitted independence in principle, since no restrictions were placed on repetition of responses, so that Ss would essentially sample with replacement. However, it is clear that both the occurrence of the stimulus and the response given to it could be expected to change the probabilities of subsequent associations. Thus there is little assurance that the independence assumption was met. Nonetheless, for purposes of analysis this should not present a serious problem, because the order of presentation of stimuli was varied.[2] The most important position, from the standpoint of influencing later responses, was the first, and at this position four stimuli occurred five times, one six times and one seven times. The use of several stimulus orders should have compensated for the possible lack of independence. Evidence that this is the case comes from the fact that when

---

[2] Even if sampling is done without replacement, over many different Ss and word orders the probability distribution ought to be the same as under the replacement assumption (Hays, 1963, p. 67). Furthermore, the only systematic difference that results from non-replacement is that the probability of the last stimulus in a list occurring as a response is slightly lower than the response probability for other stimuli; however, there was no difference between the observed probabilities of response as a function of whether the word occurred as the last stimulus.

the MCR measures were taken by position in the stimulus list, irrespective of the particular stimulus, the proportion of MCR was above chance for every position, including the first.

Stimulus differences and response bias. The possibility that there were differences among stimuli in their probabilities of evoking an MCR was tested by a χ^2 which used the total number of contrasts to compute an expected value for each stimulus. The resulting χ^2 was non-significant ($\chi^2 = 1.39$, df = 5, $p > .90$). Thus it is reasonable to conclude that there was no real difference among stimuli in their tendency to evoke minimal contrasts.

A second question which was tested by a χ^2 was the possibility of a response bias--i.e., during the association task, whether all concept-words were given with more-or-less equal probability. The number of times each word was given as a response can be seen from the last row of Table 4. The range was from 27 for KEBYR to 37 for POGUL. This difference, however, is non-significant ($\chi^2 = 1.76$, df = 5, $p > .80$). Thus it is reasonable to conclude that there was no response bias in the association task.

Sources of contrast in the associations. A minimal contrast could share features on any two of the three dimensions. There were four different contrast pairs for each such possible combination dt, dp, and tp--i.e., four pairs which contrasted only on the dry - wet dimension, four pairs which contrasted only on the public - private dimension, and four pairs which contrasted only on the tangible - intangible dimension.

Table 5 gives the frequencies for these minimal contrast combinations, as well as frequencies for double and triple contrast pairs. Each of the double-contrast combinations also has four pairs, while the triple-contrast has six pairs, one for each stimulus. The frequencies are entered for

pairs with shared features from the given dimensions. Thus dt represents cases of minimal contrast on public - private, and t represents cases contrasting on p and d. The "none" cases are the "complete opposites", contrasting on all features. The variables d', t', and p' are used to indicate the total number of cases in which the three respective dimensions provided a shared feature between two associates. (The three variables are not independent: d' = d + dt + dp; t' = t + dt + tp; p' = p + dp + tp.)

Table 5

Shared Features between Associates

Shared Feature Dimension		Frequency	Test of Differences
Triple Contrast	None	33[1]	
Double Contrast	d	17	
	t	27	$\chi^2 = 2.67$, df = 2
	p	19	p > .30
Single Contrast	dt	38	
	dp	37	$\chi^2 = 0.58$, df = 2
	tp	32	p > .70
Total Shared Features	d'	92	
	t'	97	
	p'	88	

[1] n = 6. For all other cases n = 4.

Neither the differences among dt, dp, and tp nor the differences among d, t, and p were significant when compared by two separate χ^2 tests, with dt + dp + tp and t + d + p used to compute the respective expected values. The χ^2 for the minimal contrast cases was associated with a probability of $p > .70$, while the χ^2 for the double contrast cases was associated with

a probability of $p > .30$. Thus there was no tendency for any one dimension to provide more sources of contrast than any other. It can also be seen from this table that, while the number of minimal contrasts far exceeds the number of double contrasts, despite equal probabilities, the number of double contrasts does not exceed the number of triple contrasts, when account is taken of their different probabilities.

Conclusions

The main result of this experiment was that the minimal contrast hypothesis was supported by the association data, which showed that the proportion of MCRs was well above that proportion which could reasonably be expected to result from a random selection of responses. The support for the minimal contrast hypothesis was thus statistical, and rests on the assumption that response probabilities were equal for all possible word-pairs. Since there was no control group, the validity of this assumption cannot be proved. Thus it is possible, for example, that POGUL and MAPOK, which were minimal contrasts, somehow "go together" and would have been selected as associates irrespective of their feature composition. Phonological patterns, graphemic similarities, resemblance to English words, etc. all could be postulated to have accounted for the distribution of associations.

However, there are reasons to doubt this possibility. It is highly unlikely that factors would have been present in exactly those words which were minimal contrasts. That is, the assignment of artificial words to feature combinations was entirely arbitrary and more-or-less random. The probability of any one particular assignment of words to feature combinations is $1/6! = .001$. Furthermore, it should be noted again that there

were no significant differences among stimuli in their proportion of MCR. Moreover, for every stimulus, there were two MCR associates and in no case did the frequency of association for any other word exceed either of the MCR words. In short, the result was so general that the odds against an uncontrolled variable accounting for the results are prohibitive. Nonetheless, Experiments 2 and 3 were designed to allow control comparisons with Experiment 1. These experiments are discussed in the next chapter.

Possible Processes Involved

If minimal contrasts of features can account for a significant number of associations, what then is the psychological process involved? One suggestion which has been made with respect to paradigmatic association to real words is that such associations may take less time to produce and hence are favored by instructions to respond quickly (McNeill, 1964). Siipola, Walker, and Kolb (1955) did find, in fact, that some kinds of paradigmatic responses took less time to produce. This account does not seem reasonable for the present data, not because there were no differences between the latencies of minimal contrast responses and latencies for other responses, but because the proportion of MCR was exactly equal for Ss who were told to respond as quickly as possible and for Ss who were not given this instruction. (However, the latencies themselves were extremely variable; for this reason, and because latencies were missing for several Ss, they were not considered reliable.) In fact, there was no reason to expect Ss who were told to respond as quickly as possible to give more MCRs than Ss told otherwise. On the contrary it was thought possible that Ss under a set for speed would be more likely to respond randomly, while Ss under no speed set could take time to match the features. This

expectation assumed that the present situation was quite opposite from the natural language situation, i.e., that with artificial words, matching features is a "task", while with real words it is a facilitative tool. It appears that this assumption may have been unwarranted. Feature matching apparently occurred, but it occurred for Ss who tried to respond quickly as well as for those who took their time, which suggests it was not a difficult process.

It is not necessary to assume that feature matching took place exclusively during the association task. Words could have been matched on features during the learning task and/or during the free recall task. Although the evidence for feature clustering during free recall is inconclusive, it is at least possible that Ss were using word-lists that had been given featural structures during learning. It would seem an efficient self-instruction mechanism for S to remind herself, during the learning trials that "VIKAF and NYJIB are the same, except that VIKAF is (tangible) and NYJIB is (intangible)." The speculation here is that minimal contrasts may have been matched during learning, rather than during associations. While this may have occurred, it may not be critical to the minimal contrast hypothesis, nor is it decidable on the basis of available evidence. It may not be critical for the minimal contrast hypothesis, if the hypothesis does not assume that the moment of association is necessarily the first "realization" of feature contrast, but rather that the association may reflect a previously functioning contrast.

However, if minimal contrast associations are established during learning there are two interpretations which bear differently on inferences about the associative process. One is that S becomes "aware" that two words are minimal contrasts in the sense that he is "aware" that BOY and

GIRL are minimal contrasts. In this case the nature of a minimal contrast association is within the domain of what is meant by the minimal contrast hypothesis. The second interpretation is that S actively rehearses minimal contrast pairs to facilitate learning and that subsequent minimal contrast associations are simply the verbalization of covertly practiced pairs. In this case the nature of the associative process probably lies outside what is implied by the minimal contrast hypothesis.

There is some evidence from post-experimental questioning of Ss that they were aware of feature contrasts between pairs of words during learning. About one-third of the Ss said that they used the features somehow in either learning or associating the words, but only about half of these claimed to have matched features during learning. The fact that the association task reflected minimal contrasts to a larger extent than did the free recall task may indicate that at least some Ss were actively matching features during association for the first time, rather than giving previously formed pairs. All of this is speculative, however.

Another possible process which apparently cannot be ruled out is single-feature mediation. This hypothesis would claim that Ss were using single features as association "mediators" in a manner schematized in Figure 1. The stimulus, in this case NYJIB, evokes one of its three features,

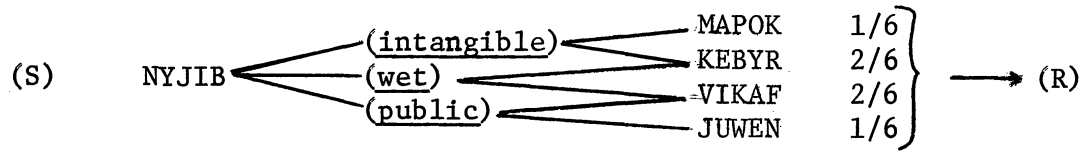

Fig. 1

Representation of S-R process under single-feature mediation model.

but not all, as a mediating response. Which feature is evoked would de-

pend on various factors, including differences among Ss, but the assumption is that over all Ss each feature has an equal probability of being evoked. This is consistent with the finding that there were no differences among dimensions in their involvement in associations. If this assumption is valid, then clearly a minimal contrast will be given as an associate more frequently than other words. Since it shares more features with the stimulus, the probability of mediation on one of its shared features is higher than for the features which the stimulus shares with another word. In the illustrated example, if either (intangible) or (wet) is the mediating feature, the minimal contrast VIKAF could be evoked. In short, only when (wet) is not a mediating feature, does the chance for a non-minimal contrast exist. The probabilities based on equi-probable feature mediation are given after the four responses. The two minimal contrast responses have a combined probability of two-thirds, compared with a probability of two-fifths under assumptions of random responding.

Notice this model does not allow the occurrence of a triple contrast, for such a word would not have a shared feature with the stimulus to function as a mediator. However, the relative proportion of triple-contrasts compared with double contrasts is too high for this aspect of the model. While both proportions are less than their expected values under assumptions of random responding the differences between the two are less than would be expected with random responding--.30 of associations were double contrasts and .16 were triple contrasts, compared to expected values of .40 and .20, respectively. The difference of .14 is less than the expected difference of .20. In other words, minimal contrasts were achieved more at the expense of double contrasts than of triple contrasts. This finding is contrary to the implications of the single-feature mediation model,

which should predict proportions of .67, .33, and 0 for single, double, and triple contrasts, respectively.

It may be more appropriate, however, to test this prediction only for Ss who responded with minimal contrasts at least half the time, so as to minimize individual differences. In other words, any process postulated to account for minimal contrast ought not to be obligated to explain the responses of Ss who did not give minimal contrasts. The differences between the 22 Ss who gave at least three MCRs and the eleven who did not, with respect to the frequencies of double and triple contrasts, is given in Table 6, along with expected values based on the ratio of total double-contrasts to total triple contrasts.

Table 6

Frequency of Double and Triple Contrasts for MCR Ss and for Non-MCR Ss. (Expected values in parentheses)

Subject Group	N	Double Contrasts	Triple Contrasts	Totals	χ^2	p
MCR	22	31 (26.58)	10 (14.42)	41	3.81	.05
Non-MCR	11	28 (32.42)	22 (17.58)	50		
Totals	33	59	32	91		

There were fewer triple contrasts given by MCR Ss compared with non-MCR Ss. (The obtained χ^2 of 3.81 is on the borderline of .05 significance for df = 1.) This result is consistent with single-feature mediation model. However, it is also consistent with other explanations, including feature matching. Furthermore, the relative frequencies of single, double, and triple contrasts even for the MCR Ss do not correspond well to the frequencies predicted by single-feature mediation. This comparison is

given in Table 7, and although a χ^2 test cannot be made on these frequencies because of the expected value of zero, the difference between the expected and observed number of triple contrasts is too large.

Table 7

Expected Distribution of Frequencies Under Single-Feature Mediation Model Compared with Observed Distribution for MCR Ss.

	Number of Contrasts		
	1	2	3
Theoretical	88	44	0
Observed	91	33	10

The foregoing analysis and discussion cannot be said to have adequately tested any hypothesis about the process involved in giving MCRs. In particular, the predictions of the single-feature mediation hypothesis, although more specific, are not very different from, for example, a matching hypothesis. In fact, the probabilities illustrated in the single-feature mediation scheme would apply exactly to a feature matching model, assuming matches were made on a single feature. The predictions made under the assumption of feature contrasting (e.g., wet - dry), rather than feature matching, would also be exactly the same. If the assumption were double-feature matching, then the predictions for the frequencies of double and triple contrasts are unclear. Presumably such contrasts would be the result of error.

In summary then, Experiment 1 has given a sufficiency test to the minimal contrast hypothesis; but it has not been particularly informative on how Ss use the features of the artificial words. The next two experiments were designed to test some aspects of this general question, as well

as to provide control conditions for the interpretation of the first experiment.

Chapter 3

Experiments 2 and 3: Dimension Salience and Redundancy

The experiments to be reported in this chapter are investigations of the possible effects that manipulations on the dimension structure of the artificial words have on associations. The manipulations were attempts to alter the symmetrical structure of the feature dimensions imposed by the first experiment. The second experiment altered this structure through what is called here "dimension salience," while the third experiment altered the structure by means of "dimension redundancy." In addition, these experiments afforded the opportunity for control-group observations with respect to the outcome of Experiment 1.

Experiment 2

In the second experiment, interest was in the effect of dimension salience, the term salience being defined in a specific operational manner: A dimension is salient in inverse degree to the proportion of words from a given set which are marked on that dimension. For example, if every word from a set is marked as (animate) or (inanimate), then animateness is a dimension of low salience. Another dimension, for example, hot - cold, which does not mark every word from the set would have relatively high salience. The salience of a dimension might also be expressed by its information value, in the sense of information theory. Thus, if half the words in a given set are marked on a given dimension, then the presence

or absence of this dimension can be thought of as carrying one bit of information. If all words are marked on the dimension, then it carries no information.

The relevance of dimension salience for minimal contrast is that when some, but not all, words from a set are marked on a certain dimension--i.e., when that dimension is salient--there is an additional source of contrast beyond that derived from feature markings, namely contrast between words that are marked on this dimension and words that are not marked. It is possible for this to be an important source of contrast. Consider the following set of words: STONE, FIRE, ICE, CAT. For this set of words, as for any set, the dimension of animateness is of low salience. Two subsets are formed by the feature contrast of (animate) vs. (inanimate): STONE, FIRE, ICE vs. CAT[1]. However, the more salient dimension of hot - cold--salient, because it marks only two of the four words--can further distinguish among the (inanimate) subset. FIRE and ICE form a subset distinguished by this dimension, and hence are potentially more similar than FIRE and STONE, even though FIRE and ICE have an additional feature contrast, (hot) vs. (cold). The argument here is that the sharing of a dimension in some cases may be a more significant source of similarity than minimal contrast of features. This is simply a restatement of the assumption that contrasts are formed at two levels, at the level of the feature and at the level of the dimension.

---

[1] This example illustrates that features also vary in salience. (Animate) in this example is more salient than (inanimate).

Design and Method of the Experiment

The concept of salience was operationalized by the addition of a fourth dimension to the dimensions used in Experiment 1. The assignment of features to words was then made so that all words were marked on two of the dimensions, four words on a third dimension, and only two words on the fourth dimension. The fourth dimension was thus the most salient. The assignment of features to words can be seen in Table 8.

Table 8

Feature Composition of the Artificial Words in Experiment 2.

Word	Dimension							
	Tan. - Int.		Dry - Wet		Pub. - Priv.		New - Old	
Vikaf	0	+	+	0	0	+		
Nyjib	0	+	0	+	+	0		
Kebyr	+	0	+	0			+	0
Mapok	0	+	0	+			0	+
Pogul	+	0	0	+	0	+		
Juwen	+	0	+	0	+	0		

The artificial words were the same used in Experiment 1, a condition employed to enable comparisons to be made. Also the first three dimensions were the same, with the additional dimension of new - old defined as highest in salience. The balance of the feature pairings has also been maintained; i.e., for each dimension, there were an equal number of artificial words having each feature on the dimension.

Minimal contrast was present in this experiment in a different sense from Experiment 1. Even for words which were marked on common dimensions exclusively, there were no pairs sharing two features and contrasting on

one. However, there was minimal contrast in the sense of the broad definition. As can be seen from Table 8, MAPOK and NYJIB, as well as KEBYR and JUWEN, were minimal contrasts in that they had the fewest number of contrasting features: zero. However, MAPOK and NYJIB did have a dimension contrast, as did KEBYR and JUWEN--i.e., NYJIB, but not MAPOK, was marked on public - private, while MAPOK, but not NYJIB, was marked on new - old. Thus, when "minimal contrast" is referred to with respect to Experiment 2, its slightly different meaning should be kept in mind.

Given this dimension structure, the main comparisons were the associations involving MAPOK and KEBYR. The hypothesis was that these two words would be given as associates to each other, because they were marked on the most salient dimension (new - old), although they are not minimal contrasts--in fact, they shared no features. The comparison thus involved the relative frequencies of their mutual association compared with the frequencies expected by chance. The prediction was that introduction of a salient dimension would alter the feature matching of the words, and result in little minimal contrast responding. That is, it was expected that KEBYR would evoke MAPOK more often than JUWEN and that MAPOK would evoke KEBYR more often than NYJIB.

Subjects and procedure. Twenty-seven female undergraduates at the University of Michigan participated in the experiment. Because three Ss failed to show satisfactory acquisition of the features, the total N for all data analysis was 24.

The procedure was identical to that employed in Experiment 1, except that no speed-instructions were given to any Ss. The same tasks used in Experiment 1 were employed in Experiment 2.

For the association task, eight orders were employed, two each containing one of the four test words in the first position. Unlike Experi-

ment 1, the main interest here was in the relative frequencies of association involving only four of the words, two of which were related by minimal contrast and two by sharing the salient dimension. Hence, since the "purest" observation was afforded by the first position, these four words--MAPOK, KEBYR, JUWEN, and NYJIB--occupied the first position an equal number of times. The two orders following each of these words differed on which of the two "interesting" possible associates preceded the other in the list. For example, KEBYR headed two lists, one with MAPOK before JUWEN, and one with JUWEN before MAPOK. (KEBYR and MAPOK shared the salient dimension; KEBYR and JUWEN were minimal contrasts in the special sense previously described.) There were thus three Ss for each list, and six for each of the first-position occurrences of the four key words. This extra precaution taken with respect to order was only because the number of observations bearing on the hypothesis were relatively small, compared to the first experiment (4 x 24 = 96, compared to 6 x 33 = 198).

Results

Learning data. The number of trials up to, but not including, the criterion trial was again taken as the measure of acquisition difficulty. The mean for this measure was 8.00, with S.D. = 3.21. The median was 7.50. All three of these statistics are higher than the corresponding values from Experiment 1. However, comparisons among the experiments are deferred until after each has been reported separately.

An analysis of variance, repeated measurements model, was again used to test differences in rate of acquisition among artificial words. Table 9 gives the summary of this analysis.

Table 9

Summary of Analysis of Variance for Differences between Words in Rate of Acquisition - Experiment 2.

Source	SS	df	MS	F	P
Between Ss	595.94	23			
Within Ss	925.50	120			
Between Words	113.81	5	22.76	3.22	.01
Residual	811.69	115	7.06		

Since the F ratio was significant, the Newman-Keuls unplanned comparison test (Winer, 1962) was applied to the means for individual words. These values were, in ascending order, MAPOK 3.58, VIKAF 4.08, KEBYR 5.00, JUWEN 5.08, POGUL 5.25, and NYJIB, 6.37. Two of the differences were statistically significant, NYJIB and MAPOK ($p < .01$) and NYJIB and VIKAF ($p < .05$); no other differences were significant. In view of the results of Experiment 1 in which the (tangible) words were the most readily learned, it may be interesting to note that the two words most quickly learned in Experiment 2 were both (intangible).

Free recall data. Again the interest in recall was whether the order of recall would reflect feature clustering, as measured by the joint occurrence of words with shared features. The expected number of joint occurrences was calculated for each dimension, based on the number of possible feature sharings for each dimension summed across all words. Table 10, with the same dimension symbols as Experiment 1, shows the expected and observed frequencies of each type of cluster. Unlike Experiment 1, the expected values are different for various dimensions and dimension combinations.

Table 10

Observed and Expected Frequencies of Joint Occurrence Features in Consecutively Recalled Words - Experiment 2.

Joint Feature Dimension		Frequency Observed	Expected by Chance	
(Min. contrast)	t	36	32	Difference
	p	20	16	
	d	25	32	not
	td	10	16	
	none[1]	29	24	Significant
	(total)	(120)	(120)	
	Kebyr-Mapok	12	8	Not
	Mapok-Kebyr			Significant

[1] Includes joint occurrences of Kebyr and Mapok.

It is clear from observation that there was no clustering, in general, nor were there joint occurrences of minimal contrast pairs. The latter can be seen by observing the frequencies for td. The last row represents the frequencies of MAPOK-KEBYR and KEBYR-MAPOK joint occurrences, tallied separately because they represented the pairs marked on the salient dimension. The frequency of their joint occurrence is only slightly above chance, and is not significant when a χ^2 is calculated with all other frequencies collapsed or even when the comparison is made between this frequency and the frequency of "none," with KEBYR and MAPOK excluded. Thus there was (1) no clustering of the words marked on the salient dimension; (2) no minimal-contrast clustering; and (3) no general feature clustering.

Association data. The distribution of associations is given in Table 11. None of the MCR percentages exceeded the chance level of $p = .20$.

Table 11

Distribution of Associations in Experiment 2
(Parentheses denote frequency for minimal contrast pair, asterisk for pair from salient dimension.)

Stimulus	Response						Percent MCR
	Juwen	Pogul	Mapok	Kebyr	Nyjib	Vikaf	
Juwen	-	6	4	(2)	7	5	.08
Pogul	8	-	8	1	2	5	
Mapok	3	6	-	9*	(4)	2	.17
Kebyr	(4)	2	7*	-	7	4	.17
Nyjib	4	4	(4)	6	-	6	.17
Vikaf	3	6	5	5	5	-	
Total Responses	22	24	28	23	25	22	

* combined percent = 33 ($p < .025$)

The proportion of KEBYR-MAPOK and MAPOK-KEBYR responses was, however, significantly above chance ($p < .025$). One-third of the total responses given to these two words were mutual evocations. Thus, the hypothesis that salience would result in the association of words from the salient dimension is confirmed. This finding also holds when observations are restricted to the first word in the list; although this represents only eight observations per word, it does indicate that the result cannot be attributed to an order effect. Again, the results are statistical, and the possibility that the association of KEBYR-MAPOK, particularly, was due to some other factor cannot be ruled out prior to consideration of the third experiment. (The comparison cannot be made with Experiment 1 because KEBYR and MAPOK were minimal contrasts in that experiment.)

The question of response bias was again routinely checked. The totals are shown in the last row of Table 11. The χ^2 based on an expected value of 24, is non-significant ($\chi^2 = 1.08$, df = 5, $p > .96$). Thus all words can be said to have had the same probability of emission.

Conclusions

The most significant finding of Experiment 2 was that the introduction of a salient dimension resulted in a significant tendency for the two words marked on the salient dimension to evoke each other as associates, at the expense of minimal contrast (in the special sense of Experiment 2).

However, since the minimal contrast pairs in Experiment 2 apparently had no contrasts at all, whereas in Experiment 1 each pair shared two features and contrasted on one, the question may be raised as to whether at least one feature contrast is necessary for above-chance probabilities of association. If contrast were necessary, this would imply that words with identical markings, synonyms, would not be associated in this type of experiment. Although this possibility was not tested, it seems rather unlikely. In Experiment 2, while the minimal contrast pairs did not have a feature contrast, they did have dimension contrasts--which is exactly the notion of dimension salience. This is presumably quite different from sharing all dimensions and all features. The explanation suggested here is that the dimension itself is a potential source of contrast, and especially, psychological, if it is a salient dimension.

Experiment 3

The third experiment added a condition of dimension redundancy to the experimental paradigm. The notion of redundancy, as it is used here, derives from its familiar use in information theory, although it is not formalized here within the information-theory framework. In the domain of concepts and word meanings, there may occur redundancies with respect to dimensions and with respect to features across dimensions. In other terms, to the extent that there is a correlation between the features or the dimensions, there is redundancy with respect to features or dimensions; a perfect correlation is complete redundancy; when the dimensions or the features are independent there is no redundancy.

Correlations are probably the general case and not the exception for words and concepts. But the implications of such redundancy for the structure of similarity and word associations is not obvious. In some cases, the correlation of features yields a new feature; thus the correlation of (tall) and (heavy) yields (big). For featural considerations, one may question whether such resultant dimensions, such as bigness, ought to be regarded as psychologically unitary. The answer probably depends again on the contrast class being constructed.[2] It is exactly the point of componential analysis that such features at times are functionally divisible into smaller components.

In the first two experiments redundancy was kept at a minimum, given the restrictions on the number of concepts which could be employed. In

---

[2] It would also depend upon the age of the subject. Presumably, younger children do not distinguish, for example, between types of bigness.

Experiment 1, the dimensions were completely independent. However, there were imperfect correlations between the features, due to the fact that all feature combinations were not exhausted. Thus, if a concept-word were (public), it was (wet) with a probability of 2/3 and (dry) with a probability of 1/3. The important points about this arrangement are that the redundancy was only partial and that it operated equally for all features; it therefore can be assumed that no "collapsing" was possible and each feature was unique. In Experiment 2, the only correlation involved dimensions, rather than features, and this simply involved the operation of salience--the probability of public - private was zero, given new - old, and vice-versa.

The point of departure for Experiment 3 was the introduction of complete dimension redundancy. Because another purpose of the third experiment was to afford control observations with respect to associations, and because feature redundancy would lead to the kind of uncertainty with respect to contrasts which was previously described (e.g., whether two redundant features would function as one), dimension redundancy rather than feature redundancy was employed. In other words, an arrangement was needed in which minimal contrasts were possible in the same sense as in Experiment 1, a condition which could not be met with complete feature redundancy. This was accomplished by complete dimension redundancy and an imperfect feature correlation which necessarily resulted from the design.

Design and Method of the Experiment

Redundancy was introduced by adding a fifth dimension to the four used in Experiment 2. All words were then marked on one dimension, tangible - intangible, with all (tangible) words being marked on two of the other four dimensions, and all (intangible) words, on the remaining two.

The resulting feature assignment, shown in Table 12, produced three words that were (tangible), (public) or (private), and (new) or (old) and three that were (intangible), (dry) or (wet), and (slow) or (fast). In this arrangement, there was not only dimension redundancy; there was, for the first time a lack of symmetry with respect to the features: two words were (public), one was (private); two were (wet), one was (dry), etc. for the four redundant dimensions. Notice that this lack of symmetry, which was a necessary consequence of the symmetry on the primary dimension of tangible - intangible, produces an imperfect feature correlation. This can be expressed in terms of conditional possibilities: given that a word is (private), it must be (old) (as well as, of course, tangible); but given that a word is (old), it may be either (public) or (private).

Table 12

Feature Composition of the Artificial Words in Experiment 3.

Word	Dimension									
	Tan. - Int.		Dry - Wet		Pub. - Pri.		New - Old		Slow - Fast	
Mapok	+	0			+	0	0	+		
Juwen	+	0			+	0	+	0		
Nyjib	+	0			0	+	0	+		
Pogul	0	+	0	+					+	0
Kebyr	0	+	0	+					0	+
Vikaf	0	+	+	0					0	+

Likewise, given that a word is (dry) it must be (fast), but given that it is (fast) it may be either (wet) or (dry). Thus, although the arrangement was asymmetrical, the redundancy was only partial. The only complete redundancy involved the tangible - intangible dimension. Given that a word

is (public) or (private) or (new) or (old) it must be (tangible), and vice-versa.

Although what to expect from this design was not certain, it was thought that clustering in free-recall, which had so far been negligible or absent, might occur in strength. That is, that the introduction of redundancy might compel a dichotomous organization of the material which would be reflected in order of free recall--that JUWEN, MAPOK, and NYJIB would form a set of (tangible) words and POGUL, KEBYR, and VIKAF, a set of (intangible) words.

With respect to associations, four of the words had one minimal contrast associate (JUWEN, NYJIB, POGUL, VIKAF) while two of the words had two minimal contrast associates (MAPOK and KEBYR). In no case did a word have as a minimal contrast associate in Experiment 3 a word which was one of its minimal contrasts in Experiment 1. This was to enable control comparisons between the two experiments. It was expected that minimal contrast responding would again be observed, and that the main effect of redundancy would be to reduce the frequency of associates without common features. That is, associations should be words from the same set, where the sets are the two dichotomized on the basis of tangible - intangible. It was predicted that words which were (tangible) would not be given to words which were (intangible).

Subjects and procedure. A new group of twenty-seven female undergraduates at the University of Michigan participated in Experiment 3. Analyses of results was again based on N = 24 since three Ss failed to show satisfactory acquisition of the features.

The procedure was identical to that of the previous experiments. For the association task, twelve orders were used, with each stimulus occurring first for two orders.

Results

Learning data. The mean number of trials before criterion in Experiment 3 was 5.79, with S.D. = 1.76. These figures were smaller than for the first two experiments. No S required more than ten trials to reach criterion, and only three Ss required more than eight. The preliminary comparisons thus suggested that acquisition of the features was quicker and less variable in Experiment 3. However, detailed comparisons are deferred until the next section.

As to differences among words in the rate of acquisition, the F ratio based on a repeated measurement analysis of variance model was again too large to accept the hypothesis that the words were of equal difficulty. The analysis is summarized in Table 13.

Table 13

Summary of Analysis of Variance for Differences between Words in Rate of Acquisition - Experiment 3.

Source	SS	df	MS	F	P
Between Ss	233.66	23			
Within Ss	385.50	120			
Between Words	128.70	5	25.74	11.54	.001
Residual	256.80	115	2.23		

The mean criterial trial for individual words was as follows: POGUL 2.67, JUWEN 2.96, VIKAF 3.75, KEBYR 4.00, MAPOK 4.08, and NYJIB 4.42, in ascending order. The Newman-Keuls unplanned comparisons test was applied to these means. The difference between POGUL and NYJIB was highly significant ($p < .01$). All the words, except JUWEN, were significantly different from POGUL ($p < .05$) and the difference between JUWEN and NYJIB was also

significant ($p < .05$). The differences apparently were not related to whether a word was (tangible) or (intangible).

Free recall data. It was expected that the recall of words would reflect two groupings, words which were (tangible) and words which were (intangible). The frequencies of these clusters and the frequency of "non-clusters" (joint occurrence of a (tangible) word and an (intangible) word) are compared to their expected values in Table 14. The value of χ^2 associated with the differences between the expected and the observed frequencies for clusters and "non-clusters" is highly significant ($p < .001$). Thus it appears that there was a tendency for groups of (tangible) words and (intangible) words to co-occur in recall.

Table 14

Expected and Observed Frequencies for "Clusters" and "Non-clusters" in Recall - Experiment 3.

Type[1]	Frequency Observed	Frequency Expected	χ^2	P
Cluster	67	48	12.53	.001
Non-cluster	53	72		

[1] A "cluster" here is the successive recall of words with the same feature from the tangible - intangible dimension; a "non-cluster" is the successive recall of words marked with opposing features from this dimension. The expected frequency of non-clusters is greater than for clusters, because for any given word there were three words marked with the opposite feature and two words marked with the same feature.

When the clusters were analyzed in more detail, there were no significant differences in the frequencies of the various types of clusters. That is, the frequencies of t, i, to, tp, iw, and if clusters were not significantly different when tested by a χ^2 for df = 5, ($p < .79$). (The

symbols t, i, etc. represent the shared features between jointly occurring words.) This indicates that clustering was due only to the tangible - intangible dimension, and that finer organization, such as minimal contrast, did not occur. This latter is inferred from the fact that the frequencies for t and i (double contrasts) were not different from the frequencies for tp, to, in and if (single contrasts), when chance is taken into account.

Association data. The distribution of associations can be seen in Table 15.

Table 15

Distribution of Associations in Experiment 3
(Parentheses denote frequency for minimal contrast pair.)

Stimulus	Response						Percent MCR
	Mapok	Juwen	Nyjib	Pogul	Kebyr	Vikaf	
Mapok	-	(8)	(6)	2	5	3	58 [40]
Juwen	(6)	-	6	6	4	2	25 [20]
Nyjib	(11)	1	-	5	3	4	46 [20]
Pogul	3	4	2	-	(10)	5	42 [20]
Kebyr	4	3	1	(5)	-	(11)	67 [40]
Vikaf	1	3	5	9	(6)	-	25 [20]
Total Responses	25	19	20	27	28	25	
Mean Percent MCR							44 [27]*

* $p < .057$. Chance level in brackets.

MAPOK and KEBYR had two MCRs each, while all other words had one. The chance probability of an MCR across all words was approximately $p = .27$, with $\sigma = .09$. The observed proportion of MCR was .44. For $N = 24$, this difference attains borderline statistical significance ($p < .057$).

However, confidence in the result may be increased by the results of the test for differences among stimuli in the frequencies of MCRs given. The χ^2 associated with the deviations of particular words from the expected values calculated from the total number of MCRs was non-significant (χ^2 = 2.87, df = 5, $p > .75$). Thus there were no differences among words in their tendencies to evoke MCRs. The check for response bias was also again non-significant (χ^2 = 2.83, df = 5, $p > .75$).

It was expected that associations would tend to be pairs of words from the same set--i.e., the (tangible) set or the (intangible) set. As can be seen in Table 16, this expectation was supported by the association results. Table 16 shows the mean frequency of two types of response to two types of stimuli: (tangible) and (intangible) responses to (tangible) and (intangible) stimuli. It is clear that whether a response was (tangible) or (intangible) largely depended upon whether the stimulus was (tangible) or (intangible). A χ^2 contingency test on the total frequencies was highly significant (χ^2 = 13.40, $p < .001$). The resulting contingency correlation coefficient (Siegel, 1956) was .31.

Table 16

Contingency Table Showing Relative Frequencies[1] of (tangible) and (intangible) Responses to (tangible) and (intangible) stimuli. (Expected values in parentheses.)

Stimulus Feature	Response Feature		Totals	χ^2	p
	(tangible)	(intangible)			
(tangible)	38 (27)	23 (34)	61	13.40	.001
(intangible)	18 (29)	46 (35)	64		
Totals	56	69	125		

[1] Cell frequencies have been corrected for unequal ns by discarding numbers from upper right and lower left cells whose means equalled the mean of the entire cell.

Conclusions

The manipulation of dimension redundancy appears to have affected all three tasks, and at least the recall and association tasks. It seems reasonable to conclude that the introduction of redundancy produced a pronounced grouping on the basis of the primary dimension, tangible - intangible. This was reflected in feature clustering during free recall, and in the groupings of the associations into (tangible) and (intangible) sets. The fact that (tangible) and (intangible) words were infrequently associated with each other may suggest that associates tend to have contrasts on subordinate dimensions, rather than superordinate dimensions. This is in accord with the case of English associations; i.e., associates to (Human) words are other (Human) words, not (Animal), (Vegetable), etc., words.

Comparisons Among the Three Experiments

Learning

Experiment 1 employed three dimensions, Experiment 2 used four dimensions, and Experiment 3 used five. The dimensions were balanced with respect to features in Experiment 1 and Experiment 2; Experiment 3 had four dimensions with unequal feature representation. Experiment 2 had a salient dimension, Experiment 3 had four redundant dimensions. With such variation, it was possible that the rate of learning the features to the words would be a function of the experimental design. A summary of the measures on the learning tasks of the three experiments is given in Table 17. The measure is the number of the trials up to but not including the criterion trial. Whether means or medians are considered, Experiment 2 required most trials for learning. If means are considered, the order of difficulty

is Experiment 2 > 1 > 3. The variability, as well as the mean, is lowest for the third experiment.

Table 17

Summary Comparison of Overall Acquisition Rates for Experiments 1, 2, and 3.

Experiment	N	Mean	S.D.	Median
1	33	6.94	3.55	6.00
2	24	8.00	3.21	7.50
3	24	5.79	1.76	6.00

These differences were initially tested by a one-way analysis of variance, with experiments as the treatment factor. The resultant F ratio of 3.10 was in the area of marginal significance ($F = 3.11$ was required for $p < .05$). Here, the variation among words was included in the error term. Therefore, it was decided to make a further test which would allow a test of differences among words at the same time.

Learning differences among words. For this comparison, the measure was the number of that trial which was the first of two consecutively correct for a particular word. Thus there were six observations for each S. With experiments as the other factor, a two-way analysis of variance model for repeated measures on one factor was used to test the effect on this measure of words and experiments. Table 18 is a summary of this analysis.

Table 18

Two-way Analysis of Variance Summary for Acquisition Rate: Experiments x Words.

Source	SS	df	MS	F	p
Between Ss					
Between Experiments	124.92	2	62.46	(< 1)	
Ss within Experiments	10513.07	78	134.78		
Within Ss					
Between Words	118.58	5	23.72	4.79	.01
Experiments x Words	105.38	10	10.54	2.13	.05
Words x Ss within Exps.	1932.35	390	4.95		

With the variance partitioned in this way, leaving words x Ss within experiments as the error term, the F ratio for the experiments factor is less than 1. The main effect due to words and the interaction of experiments x words are both significant, however. Thus, when the measure is taken on individual words, the differences among experiments are not significant. It must be pointed out, however, that this is not simply a case of isolating the error term, because the measures used for this analysis were necessarily different from those used in the one-way analysis. Nevertheless, the most tenable conclusion from both analyses is that particular words were learned with different difficulty across experiments but that certain combinations of experiments and words were more difficult than others.

The overall differences among words were tested by the Newman-Keuls test for unplanned comparisons. Table 19 shows the means for the individual words and the significance of the difference between any pair of words.

Table 19

Matrix of Significant Differences between Acquisition Rates of Individual Words Averaged over Three Experiments.

Word		Pogul	Vikaf	Juwen	Mapok	Kebyr	Nyjib
	Mean	3.92	3.92	4.01	4.10	4.60	5.30
Pogul	3.92	--				**	**
Vikaf	3.92		--			**	**
Juwen	4.01			--		**	**
Mapok	4.10				--	**	**
Kebyr	4.60					--	**
Nyjib	5.30						--

** $p < .01$. All other differences non significant. Tests were made from right-to-left above the diagonal after Newman and Keuls.

As the table indicates, NYJIB and KEBYR were learned significantly ($p < .01$) more slowly than any other word, and NYJIB was learned more slowly than KEBYR ($p < .01$). This finding does not seem attributable to the features of the words, for these varied with experiments; e.g., NYJIB was first (intangible), then (tangible), then (intangible) again, etc. Moreover, when words were ranked according to rate of acquisition in each experiment, there was no significant difference between the mean rank for (tangible) words (3.44) and the mean rank for (intangible) words (3.55).

There are two other possible explanations, however. First, only these two words have a Y vowel. It could be argued that Y is a relatively rare vowel in English, and could have led to some difficulty in pronouncing, and hence learning, the word. Secondly, it could be argued that these two words were the most difficult to learn because they were the least "meaningful"--in any of a number of senses. That is, they led to fewer English associates, were less amenable to imagery, etc. This argument would strongly suggest that Ss responded to the words by making them "meaningful."

Free Recall

Statistically significant feature clustering was found only in Experiment 3. This suggests that the redundancy of the dimensions was responsible for the clustering, and that when all words are marked on all dimensions, as in Experiment 1, the order of recall is more or less random; however, when complete redundancy is introduced, it is possible for words to be categorized into two distinctive sets, with words from the same set having an increased probability of joint recall. At least for artificial words, it seems that clustering in free recall is a less sensitive measure of feature similarity than is differential associative frequency.

Associations

In Experiments 1 and 3, associations tended to be minimal contrasts. In Experiment 2, words from the salient dimension tended to be associated. In addition to investigating redundancy, Experiment 3 also affords the opportunity to make control comparisons with Experiment 1, and, therefore, to make more (or less) confident inferences about minimal contrasts.

Experiments 1 and 3. These two experiments may be compared because they employed the same definition of minimal contrast. The comparison was made possible by requiring that the word or words which would be MCRs to a given stimulus in Experiment 3, be different words than the MCRs for the same stimulus in Experiment 1. For example, in Experiment 3 the MCR for POGUL was KEBYR, whereas in Experiment 1, JUWEN and MAPOK, but not KEBYR had been MCRs for POGUL. In no case did a stimulus have the same word as an MCR in both experiments. In this way, it was possible to compare the probabilities of a particular association under two conditions:

One in which the association constituted an MCR and once when it did not. Table 20 shows these comparisons.

Table 20

Comparison of Associative Probabilities for Pairs Words in Experiments 1 and 3. (Parentheses denote minimal contrast condition)

Response	Stimulus								
	Juwen			Pogul			Mapok		
	Exp. 1	Exp. 3	d[1]	Exp. 1	Exp. 3	d	Exp. 1	Exp. 3	d
Juwen				(.27)	.17	.10	.15	(.33)	.18
Pogul	(.36)	.25	.11				(.33)	.08	.25
Mapok	.09	(.25)	.14	(.27)	.12	.15			
Kebyr	.09	.17		.03	(.42)	.39	(.18)	.21	-.03
Nyjib	.18	.25		.12	.08		.18	(.25)	.07
Vikaf	(.27)	.08	.21	.24	.21		.15	.12	
	Kebyr			Nyjib			Vikaf		
	Exp. 1	Exp. 3	d	Exp. 1	Exp. 3	d	Exp. 1	Exp. 3	d
Juwen	.21	.12		.12	.04		(.24)	.12	.12
Pogul	.18	(.21)	.03	.15	.21		.09	.37	
Mapok	(.27)	.17	.03	.18	(.46)	.28	.24	.04	
Kebyr				(.30)	.12	.18	.15	(.25)	.10
Nyjib	(.21)	.04	.17				(.27)	.21	.06
Vikaf	.12	(.46)	.34	(.24)	.17	.07			

[1] d = [P(MCR)] - [P(Non-MCR)], irrespective of experiment; i.e., it is the number in parentheses minus the number not in parentheses. Nineteen of these twenty differences are in the predicted direction ($p < .001$).

Each part of the table shows the associative probabilities (based, of course, on observed proportions) for each response to a given stimulus in the two experiments. Parentheses denote that the given probability was for a minimal contrast; the appropriate comparison is to be made with the adjacent column. The final column in each case gives the difference between the associative probabilities under the two conditions; i.e., d = [P(MCR)] - [P(Non-MCR)].

Two of the words had two MCRs each in Experiment 1 and Experiment 3, while the remaining words had two MCRs in Experiment 1, but only one in Experiment 3, for a total of twenty comparisons. For nineteen of the twenty comparisons of associative probabilities the larger probability was for the MCR condition. The only difference not in the predicted direction was MAPOK - KEBYR, which was a minimal contrast pair in Experiment 1. (KEBYR - MAPOK was, however, in the predicted direction.) This reversal is not due to any high probability under the control condition--this is .21, exactly at chance level. This is important because it will be recalled that KEBYR and MAPOK were the salient pair in Experiment 2. That salience, and not chance, accounts for their association in Experiment 2 can be seen from the following distribution of their combined associative probabilities across conditions: Minimal contrast, .23; control, .19; salience, .33. In other words, they were associated at chance level when they were not minimal contrasts, only slightly above chance when they were, and at substantially above chance when they were from a dimension of high salience.

The result of these comparisons appears to give substantial confirmation to the minimal contrast hypothesis. The comparisons are summarized in Table 21. The entries in the table are summed proportions, multiplied by 100, of the associative frequencies between the given stimulus and the three or four responses which were MCRs for that stimulus in one experiment but not in the other, under the two conditions: when they <u>were</u> MCRs and when they were <u>not</u> MCRs.

Table 21

Summed Associative Frequencies for Words Which Were MCRs in Either Experiment 1 or 3 but Not Both.

Stimulus	N[1]	Type of Response		χ^2	p
		MCR	Non-MCR		
Juwen	3	89	42	16.86	.001
Pogul	3	96	32	32.00	.001
Mapok	4	110	62	6.70	.01
Kebyr	4	115	51	25.67	.001
Nyjib	3	101	47	19.70	.001
Vikaf	3	77	48	6.73	.01

[1] N is the number of words which served as MCRs in one experiment and non-MCR in the other.

The χ^2 values computed on the assumption of equal expected frequencies under the two conditions are highly significant for every word.

Conclusions

It appears that there can be no very conclusive statement about the process responsible for minimal contrast associations. Nevertheless, after three experiments, some tentative comments are in order. The first is that strict feature matching from a "mental list" does not seem adequate as a completely general explanation. That is, the hypothesis that Ss during the association task literally matched features in a "template" fashion is not tenable in view of the results of Experiment 2, in which there was not minimal contrast responding. Feature matching could be an adequate explanation only for Experiments 1 and 3. The same is true for most hypothetical processes, including the single feature mediation model previously discussed. It may be most reasonable to argue that a particular hypothetical process can account only for results in Experiments 1 and 3, and that the introduction of salience leads to a different process.

One process which might uniformly apply to all three experiments is word matching (or contrasting). The suggestion here would be that Ss had "paired" words prior to association, probably during learning. Thus in a condition of non-salience, Ss tend to match words which are alike on two features and contrast on one. In Experiment 2, MAPOK and KEBYR were paired together because they "stood out" as the only words marked on the salient dimension. This could account, in a rather general way, for the results of all three experiments.

Relevant to a consideration of how the features were used is the relative frequency of association of double-contrasts and triple-contrasts. This was discussed in Chapter 2, where it was pointed out that under single feature mediation or single-feature matching, there should be few or no associations without at least one common feature. The evidence on this (Tables 6 and 7) was inconclusive. It was found that Ss who gave minimal contrasts also gave more double-contrasts, relative to triple-contrasts, than did Ss who did not give minimal contrasts. On the other hand, the absolute number of triple-contrasts, while not large, was greater than zero even for MCR Ss. The same result was obtained when Ss in Experiment 3 were divided into two groups--those giving MCRs to more than half the stimuli (MCR Ss) and those giving MCRs to fewer than half the stimuli (Non-MCR Ss). The comparison is given in Table 22 (page 69).

The contingency table shows that the responses of MCR Ss tended to be words sharing one feature (double-contrast) more often than the responses of non-MCR Ss. The contingency correlation coefficient for this relationship was .37.

Table 22

Frequencies of Association Sharing One Feature and No Features for MCR Ss Compared with Non-MCR Ss - Experiment 3. (Expected values in parentheses.)

Subject Group	N	No. of Shared Features between Associations		Totals	χ^2	p
		0	1			
MCR	12	17(9)	10(18)	27	12.66	.001
Non-MCR	12	11(19)	43(35)	54		
Totals	24	28	53	81		

As in Experiment 1, this contingency suggests, not only that there are individual differences in association selection, but also that MCR Ss used the features in a general way, and even when the response was not an MCR it tended to share one feature with the stimulus. For Experiment 3, this has a more specific interpretation: The only single-feature shared between two words which were not minimal contrasts was (tangible) or (intangible), the features of the primary dimension. This is interpreted as further evidence for the grouping of the words into sets of (tangible) and (intangible).

The general relationship between the number of shared features between two words and the probability of their association is shown in Figure 2. The probabilities are based on the observed proportions corrected for unequal chance levels. (That is, chance levels have been made equal so that if associative probability were unrelated to number of shared features, the curves would be parallel to the x axis.)

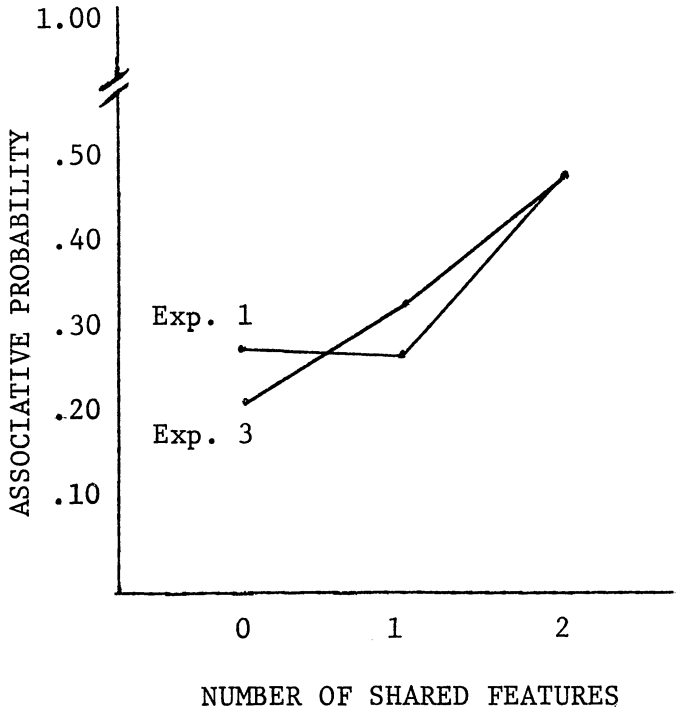

Fig. 2. Associative probability as a function of the number of shared features between associates for Experiments 1 and 3. (corrected for unequal chance levels)

The difference between Experiments 1 and 3 lies strictly in the relative heights of the curves for zero and one shared features. Minimal contrast in Experiment 3 was attained wholly at the expense of "maximum" contrasts--words with no shared features; in Experiment 1, MCRs were attained equally at the expense of single and double contrasts. This is entirely consistent with the hypothesized effect of the tangible - intangible division.

For Experiment 3, it is also consistent with the single-feature mediation hypothesis, if it is now assumed that the single-feature is from the tangible - intangible dimension. If Ss were using, for example, (tangible) as a mediator, in the manner discussed in Chapter 2, the probabilities of minimal contrast and double contrast would be similar to the observed prob-

abilities given in Figure 2. More exactly, mediation on this one feature should still produce twice as many MCRs as responses sharing only the (tangible) feature. (The set of (tangible) words included four MCRs and two words sharing only on (tangible).) Actually, the observed MCR probability was considerably less than double the observed probability for responses sharing a single feature.

Furthermore, for the twelve MCR Ss, the total number of responses which shared no feature with the stimulus was only ten. Although this number should be zero under the mediation assumption, the chance probability of such a response was rather high (three such responses for each stimulus). Since this figure is for Ss who gave MCRs to at least half of the stimuli, the number of such responses expected by chance can be calculated only for the responses which were not MCRs. For these Ss, there were 27 responses which were not MCRs. The probability, given responses which are not MCRs, of a response with no shared feature with the stimulus is 5/6. The expected number of such responses is 5/6 x 27 = 22.5. Thus the distributions of expected and observed responses of one versus no shared features are greatly divergent, as can be seen from Table 23.

Table 23

Comparison of Expected and Observed Distribution Of Non-MCR Responses for MCR Ss in Experiment 3.

Distribution	No. of Shared Features between Associates	
	0	1
Expected by Chance	22.5	4.5
Observed	10	17

It must be emphasized that although the distributions in Table 23 are highly divergent, the expected values are those calculated on the basis of chance, not on those expected under the hypothesis of single-feature mediation.[3] (The expected distribution for the latter would be 27 and zero.) Also, the number of Ss is not large. Nonetheless, while single-feature mediation cannot be substantiated as the process involved, neither can it be ruled out. Furthermore, the assumptions of feature matching would be met equally by these outcomes.

Summary

The results of the three experiments appear to point to the following conclusions: (1) Associations given to artificial words tend to be other artificial words which contrast with the stimulus on a minimal number of features, other factors being equal. (2) When feature dimensions are not equally represented among the artificial words--i.e., when one dimension is salient--minimal contrast responding does not occur for words from the salient dimension; instead the "salient", and contrasting, words tend to be associated with each other. (3) When feature markings on only one dimension serve to distinguish completely between two sets of words--i.e., when other dimensions are redundant--associations tend to be minimal contrasts. (4) The order of free recall does not reflect an organization of words on the basis of features, except when there are redundant dimensions. In this case there is clustering only on the basis of the primary dimension, i.e.,

---

[3] Another caution required here is that the observed distribution of Table 23 is exactly that distribution for MCR Ss tested in Table 22 against the distribution of Non-MCR Ss. Thus the interpretation is simply that the distribution for MCR Ss differs from the distributions of (1) Non-MCR Ss and (2) Chance--these are distinct but not independent inferences.

there is simply a dichotomy. This suggests that clustering in free recall may be a less sensitive measure of feature similarity than is differential associative frequency. (5) The manner in which Ss utilize the features to produce associations is a matter for speculation. The following hypotheses cannot be ruled out on the basis of the data: (i) Features are actively matched or contrasted by S during association; (ii) Words are "paired" prior to association, presumably during learning; (iii) Ss use a single feature of the stimulus to mediate the response.

Chapter 4

Experiment 4: Similarity Judgments

The experiments reported thus far have been concerned exclusively with artificial verbal materials. It is to be emphasized that work with artificial words can provide only a meager beginning for an investigation of componential analysis. The next step in this line of study is an attempt to apply componential analysis, in an exploratory way, to English words. Specifically, in Experiment 4 the interest was in judgments of meaning similarity for English words which had been componentially analyzed.

Meaning similarity ought to be the natural domain of componential analysis. If meanings of words can be described in terms of features, then the sharing of common features between two words should be in some correspondence to what is ordinarily meant by "meaning similarity."[1] To put it briefly, if, for a set of words, all semantic features are specified, the pair of words most similar in meaning ought to be that pair which share the greatest number of semantic features. We can be certain, however, that this is too simple. For example, we can suppose that semantic features are differentially significant for the meaning of given words; i.e., that there must be weightings applied to features if similarity is to be predicted. However, as a beginning, this very important

---

[1] What is ordinarily meant by "meaning similarity" is, of course, not obvious. Philosophers have long labored on definitions of meaning and meaning similarity. In this case, we must dodge the meaning of "meaning" and rely instead on the judgments of the subjects.

consideration has been ignored and the simpler prediction made: Two words which share a greater number of features should be judged more similar than words sharing a lesser number of features.

In extending componential analysis to English words, we are faced with the problems mentioned in Chapter 1, the main one of which is the criteria to be employed in the selection and marking of features. (One criterion not employed here is that semantic features be strictly linguistic. This criterion, which is applied by Katz and Fodor, excludes the type of information coded in the present features.) Even when sematic features can be selected with some confidence there are additional problems, such as how to evaluate the relative significance of different features. This difficulty, as well as others, will be seen more clearly as the method and results of Experiment 4 are reported.

Method

Selection of Words and Features

The words used in this experiment were selected from an unpublished list of 70 emotion words, 40 nouns and 30 adjectives, which were made available by Professor C. E. Osgood. The words had been analyzed into thirteen feature dimensions by Osgood and his colleagues at the University of Illinois Institute of Communications Research. (For a paper describing a featural analysis of interpersonal verbs, see Osgood, 1966.) For the present experiment, 38 of the nouns and 23 of the adjectives were used with Osgood's features unaltered. It should be emphasized that the features, which are the result of sophisticated intuition, are not necessarily final. Further efforts may result in the addition or elimination of dimensions, or even re-coding.

Feature Coding. Each word was marked on each of the bi-polar feature dimensions. The feature codes were (+), (0), and (-). The (+) denotes the presence of one of the two features from a given dimension, (-) denotes the presence of the opposite feature, and (0) denotes that neither feature is consistently present or that, in some sense, the dimension does not apply. For example, the word SUDDEN was marked (0) on the dimension of pleasant - unpleasant, (+) on active - passive, and (-) on controlled - uncontrolled. The (+) code always denotes the first-named feature, the (-) code, the second. The features refer to the content of the emotion (for nouns) or the content of the emotion description (for adjectives). Below is a brief designation of each feature dimension, with clarifying examples:

1. pleasant - unpleasant. CONTENTMENT (+); DESPAIR (-)

2. active - passive. VIOLENT (+); DULL (-)

3. controlled - uncontrolled. FIRM (+); SUDDEN (-)

4. external - internal. This is roughly parrallel to the privateness of the emotion. PITY (+); AMAZEMENT (+); ANXIETY (-); COMPLACENCY (-)

5. cognitive - visceral. DISTRUST (+); LOATHING (-)

6. ego - alter directed. This could also be described as reactive (+) versus active (-) or perhaps respondent (+) versus operant (-). PAIN (+); HORROR (+); AWE (+); EXPECTANCY (-); LOATHING (-) DETERMINATION (-)

7. future - past. HOPEFUL (+); APPRECIATIVE (-) i.e., one is HOPEFUL that something will (or will not) happen, but APPRECIATIVE about something that did (or did not) happen.

8. superordinate - subordinate. This refers to the status of the person within the interpersonal relationship implied by the emotion: FIRM (+); HUMBLE (-)

9. social - nonsocial. SHAME (+); GUILT (-). This is related to, but different from, external - internal; e.g., HORROR is external but nonsocial.

10. terminable - interminable. This refers to the duration of the emotion. GLEE (+); BOREDOM (-)

11. striving - nonstriving. DETERMINATION (+); SORROW (-)

12. third-person dimension. Evaluation of the emotion requires a third-person (+), or does not require a third-person (-). SINCERE (+); HOPEFUL (-)

13. intensive - nonintensive. DESPICABLE (+); CONSIDERATE (-)

The complete feature codings for all words used in the experiment are given in Table 24.

Table 24

Feature Codings on 13 Dimensions

Word	Dimension*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
(Adjectives)													
firm	0	0	+	0	+	-	0	+	0	-	+	0	0
hopeful	+	0	0	0	0	-	+	0	-	-	+	-	0
acute	0	+	-	0	0	0	0	0	0	+	0	-	+
eager	+	+	0	0	0	-	+	0	0	0	+	0	0
sudden	0	+	-	0	0	0	0	0	0	+	0	0	0
dull	-	-	0	-	0	0	0	0	-	-	-	0	-
despicable	0	0	0	0	0	-	0	0	+	0	0	+	+
emphatic	0	+	+	+	0	-	0	+	0	+	+	0	+
considerate	0	0	0	+	+	-	0	0	+	0	0	+	-
desperate	-	+	-	0	0	0	+	-	0	0	+	0	+
hot	0	+	-	0	-	-	0	0	0	+	0	-	+
proud	+	0	+	0	0	-	0	+	+	0	0	0	0
deliberate	0	0	+	0	+	-	+	+	0	-	+	-	0
guilty	-	0	0	0	0	+	0	-	+	0	0	0	0
cynical	0	0	+	0	+	-	0	+	+	0	-	0	-
vague	0	0	-	-	0	0	0	0	-	-	0	-	-

Table 24 (continued)

Word	Dimension*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
(Adjectives)													
selfish	0	0	0	0	0	-	0	0	+	0	0	+	0
humble	0	0	0	0	0	+	0	-	+	0	-	0	0
violent	0	+	-	0	-	0	0	0	0	+	0	0	+
appropriate	0	0	0	+	0	0	0	0	+	0	0	+	0
appreciative	+	0	0	+	0	+	-	0	+	0	-	0	0
uncertain	-	0	-	-	+	0	+	-	-	+	0	-	-
satisfying	+	0	0	0	0	0	-	0	0	0	0	-	0
(Nouns)													
pain	-	0	-	0	-	+	0	0	-	0	0	-	0
glee	+	0	0	0	+	0	0	0	0	+	0	0	-
sorrow	-	0	-	0	0	0	-	0	0	0	-	0	0
disgust	-	0	0	+	0	-	-	+	0	+	0	0	0
excitement	0	+	-	0	0	0	0	0	0	+	+	0	0
dismay	-	0	-	+	+	+	0	0	-	+	-	0	-
horror	-	+	-	+	-	+	0	0	-	0	0	0	+
expectancy	0	0	0	0	+	-	+	0	-	0	0	-	-
amazement	0	+	-	+	0	+	0	0	0	0	-	0	0
shame	-	0	-	0	-	0	-	-	+	-	0	0	0
contentment	+	-	0	-	0	0	0	0	-	-	-	0	-
despair	-	0	-	-	0	0	+	-	-	-	-	0	0
sulkiness	0	0	0	-	0	0	0	0	+	-	+	0	0
distrust	0	0	0	0	+	-	+	0	+	0	0	0	0
surprise	0	+	-	+	0	+	0	0	0	+	0	0	0
loathing	-	+	-	+	-	-	0	+	+	-	+	0	+

Table 24 (continued)

Word	Dimension*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
(Nouns)													
complacency	+	-	0	-	0	0	0	+	0	-	-	+	-
pity	0	0	0	+	0	-	-	+	+	0	-	0	0
stubbornness	0	0	+	0	0	-	+	0	+	-	+	0	0
joy	+	+	-	0	-	0	0	0	0	0	0	0	+
suspicion	0	0	0	0	+	-	-	0	+	0	0	0	0
bewilderment	-	0	-	0	+	+	0	-	-	+	0	0	0
boredom	-	-	0	0	0	0	0	0	0	-	-	-	-
awe	0	-	-	+	0	+	0	-	-	-	-	0	0
anxiety	-	+	-	-	-	0	+	0	-	0	0	0	0
sadness	-	-	0	-	-	0	0	0	-	-	-	0	-
determination	0	0	+	0	0	-	+	0	0	0	+	0	0
contempt	0	0	+	0	0	-	0	0	+	0	0	0	0
interest	0	0	0	+	+	-	0	0	0	+	0	0	0
embarrassment	-	0	-	+	0	0	0	-	+	0	0	0	0
rage	0	+	-	0	-	-	0	0	0	0	0	0	+
adoration	+	-	0	0	0	0	0	-	+	-	-	0	0
puzzlement	0	0	0	0	+	+	0	0	-	+	+	-	-
dread	-	+	-	0	0	0	+	-	0	-	0	0	+
annoyance	-	0	0	+	0	+	0	0	0	+	0	0	-
worry	-	0	0	0	+	0	+	0	0	0	0	0	0
scorn	0	0	+	+	0	-	0	+	+	0	0	0	+
fear	-	+	-	0	-	0	+	0	0	0	0	0	0

* The dimensions are labelled by numbers. Content descriptions corresponding to each numbered dimension appear in the foregoing text.

Experimental Design and Procedure

All except nine words from Osgood's list were used in the present experiment. Eight words were discarded for failing to have non-zero feature markings on at least three of the thirteen dimensions, and one was discarded because it had identical feature markings with another word on the list.[2] Thus, 61 words were used in the experiment.

The format of the judgment task required Ss to make a judgment of relative similarity between two comparison words and a standard word. The standard word appeared in the left-hand column of the page followed by the two comparison words. For example, one judgment was the following:

EMBARRASSMENT: ___CONTEMPT ___BEWILDERMENT

Ss had to check one of the two comparison words as more similar to EMBARRASSMENT.

Each word served as a standard under two conditions: (1) The comparisons contrasted on an equal number of features with the standard, but varied in the number of features shared with the standard. This was called the F+ condition. (2) The comparisons shared an equal number of features with the standard, but varied on the number of features contrasted with the standard. This was called the F- condition. The unit of the sharing or the contrast was always the feature, never the dimension. In other words, only if two words had (+)'s or both had (-)'s for a given dimension was there a feature sharing. A contrast occurred, by definition, only if one word was marked (+) and the other (-). If one of the words was marked (0), that dimension was disregarded.

---

[2]The discarded word was MUTUAL which had the identical feature markings given to APPROPRIATE. The obvious differences in meaning between these two words is a strong indication that the meanings of at least some words cannot be adequately specified in these 13 dimensions.

The example previously given can clarify this aspect of the design:

EMBARRASSMENT: CONTEMPT (1,1) BEWILDERMENT (3,1)

The parentheses after the comparisons have two numbers: The first is the number of features shared with the standard, and the second is the number of features contrasted with the standard. The example is from the F+ condition. The comparisons have an equal number of contrasts with the standard (1), but an unequal number of sharings. Thus, the prediction in this case was that BEWILDERMENT would be judged more similar to EMBARRASSMENT more often than would CONTEMPT.

While the prediction for the F+ condition was straightforward, what to expect from the F- condition was uncertain. It was not obvious whether a larger number of contrasts should lead to more or fewer similarity judgments. One hypothesis was that contrasting on a dimension would produce more similarity than not sharing the dimension at all, when the number of sharings is equal for the two comparisons.[3] On the other hand, it was possible that Ss would judge the word with more contrasts as closer to "an opposite" of the standard, and hence eliminate it as a similar. The following is an example of an F- judgment:

EMBARRASSMENT: PAIN (2,1) ANXIETY (2,2)

The two comparisons now share an equal number of features with the standard and contrast on an unequal number.

Another example can be used to show the variability between comparisons:

SUSPICION: EXCITEMENT (0,0) PITY (3,0)

This is an F+ condition, with one comparison having no dimensions in common

---

[3] The distinction between shared dimensions and shared features was made in Chapter 3. On a given dimension, if one word is (+) and the other (-), the two share the dimension, but not the feature. If one word is (0), they do not share the dimension.

with the standard; i.e., EXCITEMENT and SUSPICION have no contrasts and no sharings. Another F+ judgment was

CONTENTMENT: COMPLACENCY (6,0) EXPECTANCY (2,0)

Here one comparison has four more features shared with the standard than does the other comparison.

For both the F+ and F- conditions, the difference between the two comparisons in the number of shared or contrasted features (one or the other was always constant) ranged from 1 to 4. This variability, at least for the F+ condition, enabled the prediction to be made that the greater the difference in the number of shared features between two comparisons, the greater the probability that the comparison with more shared features would be judged more similar. Again, for the F- condition, what to expect was uncertain. Other variables which were to be tested as predictions of similarity will be described later.

Judgment sets. There were four complete sets of judgments, two with nouns and two with adjectives. Sets A and B were nouns and Sets C and D were adjectives. When a noun was a standard under the F+ condition in Set A, it was F- in Set B, and vice-versa. When an adjective was a standard under the F+ condition in Set C, it was F- in Set D, and vice versa.[4] In order for every word to occur as a standard, it was necessary for every word to occur twice as a comparison. Every noun was a comparison for two

---

[4] There was one exception to this symmetry. After many fruitless attempts, it finally proved impossible to maintain perfect symmetry and at the same time meet the requirement that each word serve as a comparison an equal number of times. Since the latter seemed the more important restriction it was necessary for one word, GUILTY, to serve under the F- condition in both sets. Hence, there were 22 adjectives for the F+ judgments and 24 for the F-.

standards in Set A and a comparison for two different standards in Set B. Every adjective was a comparison for two standards in Set C and a comparison for two different standards in Set D. No word was used as a comparison for a standard which had itself been a comparison when the given word was standard. In other words, every similarity judgment was unique.

The orders of the judgments were randomized with the restriction that several judgments intervene between occurrences of the same word. Usually there were four or more intervening judgments.

Every S received three of the four sets. The comments from pilot Ss indicated that the task became somewhat repetitious when all four sets were given, and data collection was too inefficient if only two were given. With each S receiving three tasks, an N of 30 could be obtained with 40 Ss. Half the Ss, those with two sets of nouns and one set of adjectives, thus made 99 judgments, while the other half, those with one set of nouns and two sets of adjectives, made 84 judgments. For all Ss, the order of noun versus adjective sets was alternated to reduce the possibility of excessive repetition. Finally, across Ss, every set occurred an equal number of times as the first, second, and third set in the task.

Appendix II presents the four sets of words presented to Ss.

Subjects and procedure. Forty female Ss, almost all of whom were obtained from the University of Michigan introductory psychology pool, made the similarity judgments. Data were collected in two group sessions. The instructions asked Ss to mark the word which was "more similar in meaning" to the standard. They were urged to make careful judgments even when neither word appeared very similar to the standard. Appendix III gives the verbatim instructions.

Main Results

On each of the four sets of judgments there was a control judgment inserted for the purpose of defining the quality of the judgments of a particular S. These control judgments were designed to allow elimination of Ss who (1) were atypical in their judgments or (2) did not attend to the task. Following are these control judgments:

(Set A)	ODOR:	___SMELL	___ANGER
(Set B)	DOUBT:	___COURTESY	___SKEPTICISM
(Set C)	IMAGINATIVE:	___CREATIVE	___COLD
(Set D)	SICK:	___ILL	___SINCERE

Each S encountered three of these items, and any S who did not give the obvious judgment on all of the three was eliminated from the data analysis. Two Ss were eliminated by this test, and two additional Ss were obtained to keep the total N at 40 and the N for each judgment at 30.

The measure of similarity used was simply the percentage of Ss judging a particular comparison word as more similar than the alternative comparison word to a given standard. This measure is called Relative Similarity (RS). The results are presented below separately for the two conditions, first for F+, then for F-. (The RS scores for each judgment item can be seen in Appendix II.)

The F+ Condition

The main result under the F+ condition was that the RS measure tended to be higher for comparison words which shared the greater number of features with the standard. Across all judgments, the comparison with the greater number of shared features with the standard (Fg word) was judged

more similar than the word with the lesser number of shared features (F1 word) in 39 cases. The F1 word was judged more similar in 18 cases, and in two cases there were ties--i.e., half the Ss chose the Fg word and half chose the F1 word. The total number of cases here is 59 instead of 60, because judgments for one standard had to be omitted from analysis due to an error in the preparation of judgment forms.

When the judgment data are considered separately for nouns and adjectives, a marked difference is observed. For adjectives, the prediction that the Fg word would have a higher RS than the corresponding F1 word failed only four times out of 22 (about 18%). For nouns, however, this prediction failed fourteen times out of 37 (about 38%).

Much the same picture is obtained when mean RS or median RS is considered. The overall mean RS for Fg words was 65.3. This mean is based on 1770 judgments (59 judgment items x 30 Ss) and is very much above the level of 50.0 expected by chance. Again, the prediction fared better for the adjectives than for the nouns. The mean RS was 61.4 for nouns and 71.8 for adjectives, a highly significant difference ($p < .001$). However, even when significance tests were made separately, the mean RS for both adjectives and nouns were highly significant ($p < .001$). Table 25 presents a summary of the mean RS data.

Table 25

Mean RS for Fg Word

	Fg - F1		
Form-Class	1 or 2	3 or 4	Overall
Nouns	60.7	62.7	61.4
Adjectives	66.7	82.9	71.8
Combined	63.0	69.7	65.3

Table 25 also provides a check on the effect of the difference in the number of features each comparison shared with the standard. It was expected that the RS for an Fg word would increase as the difference between Fg and Fl on the number of shared features increased from one to four. (This difference is denoted by Fg - Fl.) Because this difference was more often two or three than one or four, the means for one and two and for three and four have been combined to provide comparisons based on comparable ns. As can be seen from Table 25, there was a marked difference in the RS for adjectives between the two levels of Fg - Fl; for the nouns alone, there was no significant difference. When nouns and adjectives are combined the overall difference is highly significant ($p < .01$)[5].

The data presented thus far can be summarized: When the judgment involved two comparisons which differed only in the number of features shared with the standard (F+ condition), the word with the greater number of shared features (Fg) was judged as more similar to the standard significantly more often than the word with the lesser number of shared features (Fl). For adjectives, but not for nouns, a larger difference between the two comparisons in the number of features shared with the standard (Fg - Fl) produced a higher similarity score (RS) for the Fg word than did a smaller difference between the two comparisons. We shall return to discussion of these results after a consideration of the F- findings.

The F- Condition

The total number of judgments used in the F- analysis was 62--38 nouns and 24 adjectives. The main result for the condition in which the two com-

---

[5] All significance tests referred to in this section are tests of proportions where the observed proportion is assumed to be an estimate of the probability parameter, p, whose variance is $\frac{p(1-p)}{N}$.

parisons shared the same number of features with the standard, but differed in the number of contrasts, was that the comparison with the lesser number of contrasts (Cl) tended to be chosen as more similar to the standard than the comparison with the greater number of contrasts (Cg). The mean RS was 38.7 and 35.6 for Cg nouns and Cg adjectives, respectively. The difference is not significant, and the combined figure is 36.9. It was not possible to check the effect of Cg - Cl, the difference between the two comparisons in the number of contrasts, because the variability of this difference for the adjectives was too small.

These data suggest the possibility that Ss were to some extent eliminating "opposites" in the F- judgments. It had been speculated that Ss might be willing to judge two words which were somehow contrasting as more similar than two words which had "nothing" in common. Although the overall data for the F- case do not support this speculation, the appropriate test is a comparison of mean RS levels between Cg words with no features shared with the standard and Cg words with at least one feature shared with the standard. This comparison is given in Table 26.

Table 26

Mean RS for Cg Word Compared for Different Levels of Feature Sharing

	Number of Shared Features Between Standard and Cg Word				Alternative Had No Shared or Contrasted Features	
	None		At Least One			
Form-Class	N	Mean RS	N	Mean RS	N	Mean RS
Nouns	16	39.2	22	38.3	6	25.0
Adjectives	12	37.8	12	33.4	6	36.4
Combined	28	38.4	34	35.1	12	30.7

The Table shows the mean RS score for the Cg comparison word under two conditions, when (1) both comparison words had no shared features with the standard, and when (2) both comparison words had at least one shared feature with the standard. If contrast were responded to as a source of similarity in the absence of shared features, then the RS measure should be higher for (1). The first two columns give this comparison for nouns and adjectives separately and for the two combined.

The differences are all neglibible. The far-right column gives supporting data on this subject. The figures are the mean RS scores for Cg words whose alternative had no shared features and no contrasts with the standard. These data imply that when Ss were faced with the alternative of judging as more similar either a word which had only contrasts with the standard or a word with nothing at all in common with the standard, they chose the latter. Thus, in general, under the F- condition Ss were eliminating "opposites."

Additional Analyses

Most of the remaining discussion will be devoted to a closer look at some of the specific judgment items and a consideration of possible predictive variables, both primarily with respect to the F+ condition.

Some Examples

A consideration of individual judgment items reveals that most of the judgments can be characterized as "non-obvious." That is, rarely, if at all, did one of the comparison words approach synonymy with respect to the standard. The nearest approximation to synonymy, in fact, occurred in one of the control judgments, ILL - SICK; the other control judg-

ments, ODOR - SMELL, DOUBT - SKEPTICISM, IMAGINATIVE - CREATIVE, also are very similar, intuitively.

However, for the test items high intuitive similarity was almost completely lacking. Consider the following examples:

(1) DISMAY: SUSPICION (24.1); ANXIETY (75.9)

(2) HOPEFUL: APPROPRIATE (82.8); DULL (17.2)

(3) CONTEMPT: DISGUST (50.0); SCORN (50.0)

These three judgments are all from the F+ condition. In each case the underlined word is the Fg word--i.e., it shares a greater number of features with the standard than does the alternative comparison word--and hence it was predicted that the underlined word would have the higher RS in each example. These examples illustrate the nonobviousness of the predictions. In judgment (1), DISMAY and ANXIETY are both (unpleasant), (uncontrolled), and (nonsocial), while DISMAY shares with SUSPICION the feature (cognitive); both comparisons have two contrasts with the standard. Although neither comparison is ordinarily substitutable in meaning for DISMAY, about three-fourths of the Ss selected the Fg word (with the greater number of shared features) as more similar.

Examples (2) and (3) illustrate failures of the prediction. A more comprehensive consideration of such failures will be made later. For present, it is sufficient to note the nonobviousness of both predictions, and, in particular the ambivalent nature of (3). Both DISGUST and SCORN are coded as sharing (alter-directed) with CONTEMPT. The difference lies in that SCORN and CONTEMPT share (controlled) and (social). Subjects, however, were evenly divided on which comparison was more similar.

Some Possible Variables

Given the main findings that the word selected as more similar tended to have more shared features or fewer contrasts, depending upon the condition, an investigation was made of possible variables which could post-predict the judgment data. Four variables, all concerned with the number of shared and/or contrasting features between standard and comparison were tested. The four are briefly described below within the conceptual framework of the F+ condition. They are, of course, applicable to the F- condition as well.

(1) Fg - Fl. This was a priori the variable expected to result in best prediction. This would predict that the probability of a comparison word's being judged similar is directly related to the difference between the number of features it shares with the standard and the number of features the alternative word shares with the standard.

(2) Fg - Fl / Fg + Fl. This variable adds proportionality to the difference variable. It is analogous to the Weber-type discriminability function familiar to psychophysics. It is derived from the notion that discriminable differences in the number of shared features are the important variable, and that these depend inversely on the combined number of features which the two comparisons share with the standard. For example, it would predict that a judgment whose two comparisons were (3,1)[6] vs. (1,1) would produce a greater RS for the first comparison than would a judgment whose two comparisons were (5,3) vs. (3,3), although Fg - Fl = 2 for both cases. The value on variable (2) would be $\frac{3 - 1}{3 + 1}$ = .50 in the first case and $\frac{5 - 3}{5 + 3}$ = .25 in the second case.

---

[6] The parentheses, it will be recalled, include the dimension commonality between standard and comparison. The first number is the number of shared features and the second is the number of contrasted features.

(3) Fg + Cg / Fg + Cg + Fl + Cl. This variable assumes that what is significant is the ratio of total shared dimensions of one comparison to the total shared dimensions of both comparisons. It does not distinguish between a shared and a contrasted feature. Two examples:

[a], (3,2) vs. (1,2)

[b], (3,0) vs. (1,0).

For [a] the first comparison would have a value on this variable of $\frac{3 + 2}{3 + 2 + 1 + 2} = .625$; for [b], the first comparison would have a value $\frac{3 + 0}{3 + 0 + 1 + 0} = .750$. The two judgments would have identical values on variables (1) and (2) but different values on (3).

(4) Fg. The fourth variable was simply the number of features shared between the standard and the comparison with the greater number of features.

It should be noted that with all these variables, the F's become C's when we are dealing with the F- condition. For example, (1) becomes Cg - Cl, (2) becomes $\frac{Cg - Cl}{Cg + Cl}$, and (4) becomes simply Cg. The notation for variable (3) does not change.

For both the F+ and the F- conditions, and for the two combined, correlations were computed between each of the four predictor variables and RS, the dependent variable. Under the F+ condition, values for each variable were obtained for the Fg words, while under the F- condition, values were obtained for the Cg words. Product-moment correlations were employed, but with some reservation because of the relatively small number of words and because of the fairly restricted range on variable (1). However, since the range was thought to be satisfactory for the other variables, including RS, and since this post-hoc phase of the research was primarily exploratory, it was decided that product-moment correlations could give adequate information for our limited purposes.

Of the four variables, only (2) showed any predictive correlation pattern with RS, and even for (2) the pattern was not uniform. Variable (1) correlated 0 with RS for nouns and adjectives, and for F+ and F-. Variable (3) correlate negatively for nouns for both F+ and F-, but 0 for adjectives F+ and F-. Variable (4) showed consistent 0 correlations except for the D+ combined condition. D+ combined the values from the F+ and F- condition and is thus a measure of dimension commonality, irrespective of whether the commonality is the result of shared or contrasted features. Table 27 gives a summary of the correlations.

Table 27

Product-Moment Correlations
between Four Predictive Variables and RS.

Form-Class	Condition[1]	Variable			
		(1)	(2)	(3)	(4)
Nouns	F+	NS[2]	NS	-.85	NS
	F-	NS	NS	-.22	.24
	D+	NS	NS	-.45	.40
Adjectives	F+	NS	.41	NS	NS
	F-	NS	-.41	NS	NS
	D+	NS	NS	NS	.44
Combined	F+	NS	.22	-.26	NS
	F-	NS	-.22	NS	NS
	D+	NS	NS	-.18	.40

[1] The D+ condition simply ignores the distinction between shared and contrasted features and, in effect, is a combination of the F+ and F- conditions.

[2] All correlations, except NS (non-significant), exceed at least the .05 significance level.

The patterns of correlations shown in Table 27 indicate that variable (2) was the best predictor of RS. Although it failed to produce a significant correlation for the nouns, it produced significant and interpretable correlations for adjectives and for all words when form class was disregarded. The interpretability of the correlation is to be emphasized, because while variables (3) and (4) produced significant correlations, they were difficult to interpret. For example, variable (3) produced negative correlations for the nouns under both F+ and F- conditions; opposite-signed correlations would have been more consistent with the observed effect of these two conditions on the similarity judgments. The large negative F+ correlation would imply that the smaller the discriminable difference between the number of dimensions each comparison shared with the stimulus, the more likely that the Fg word would be judged more similar. Furthermore, variable (3) did not show a correlation for the adjectives, despite the fact that the adjectives produced significantly higher RS for the Fg words than did the nouns. Variable (4) must be judged unsuccessful for a similar reason. Although it produced consistent correlations for the D+ combined conditions, this is not interpretable in the absence of correlations for the F+ and F- conditions, except as a statistical artifact resulting from the combination of two zero correlations.

Variable (2) thus was the most valid predictor of RS. Its failure to show correlations for the nouns is not unreasonable in view of the observed differences between nouns and adjectives. That correlations were positive for the F+ condition and negative for the F- condition is exactly what is to be expected: the larger the ratio of Fg - Fl to Fg + Fl, the greater the RS of Fg; the larger the ratio of Cg - Cl to Cg + Cl, the smaller the RS of Cg.

At least for adjectives, the correlations imply that Ss were to some extent sensitive to differences between the comparison words with respect to features both shared and contrasted with the standard. They further imply that Ss were not using total dimension commonality as judgment information, but rather that they selected as similar words with more shared features and fewer contrasting features. Finally, however, the relatively small ns which provided the data for these correlations should caution against over-interpretation. They are to be regarded as exploratory.

Where Prediction Failed: Suggestions for Modification

The similarity judgments showed considerable uniformity. Although there were differences between nouns and adjectives, there was surprising agreement between the F+ and F- conditions, as can be seen from Table 28.

Table 28

Deviations of Median RS Scores from 50

Form-Class	Condition	
	F+ (Fg Word)	F- (Cg word)
Nouns	+14.3	-15.0
Adjectives	+26.7	-27.0
Combined	+20.0	-20.0

The numbers in Table 28 are the deviations of the median RS scores from .50. Significant here, in addition to the difference between nouns and adjectives, is the near identity of the F+ and F- deviation, for both nouns and adjectives. Loosely speaking this means that a comparison word was equally likely to be judged more similar whether it had a maximum num-

ber of shared features or a minimum number of contrasted features, compared with the alternative comparison. The correlations and the data in Table 28 strongly support the notion that subjects utilized both sharings and contrasts as reliable sources of information for the similarity judgments.

The uniformities in the data, however, should not obscure the irregularities. In particular, for the F+ condition there were 18 cases, or about 30%, in which the F1 word obtained a higher RS than the Fg word. These are failures of prediction which may be accounted for in numerous ways, three of which were previously acknowledged as potential sources of difficulty. Failures in prediction reflect a need at least for (1) a weighting technique applied to the dimension, (2) additional dimension to describe more adequately the meanings of some of the words, and/or (3) a re-coding of the features. The remainder of this section will be a brief exploration into the significance of the first two considerations.

An important dimension. There is no reason to expect that the 13 dimensions identified for the words in this experiment should be equally significant as semantic components. Neither is it likely that the relative significance of the dimensions should be the same for each word. Furthermore, a comprehensive semantic account of the words would recognize that the significance of a dimension, even for a given word, can change with both linguistic and non-linguistic contexts. Although a comprehensive semantic account is beyond the present purpose, we can consider the question of overall dimension significance as it relates to the judgments in this experiment, at least by means of illustration.

Given the feature which marked all the words in this experiment (emotional), it was not difficult to arrive at at least one dimension, which

appeared intuitively to be more powerful than the others. Pleasant - unpleasant was a natural candidate, although one may suppose others are of above-average significance. The significance of the pleasant - unpleasant dimension was tested by consideration of all judgments for which the standard was marked either (+) or (-) on this dimension. From these were selected only those judgments in which: (1) for the F+ condition the Fg word was marked on the dimension, but with a contrasting feature with respect to the F1 word and (2) for the F- condition the Cg word was marked on the dimension, but with a contrasting feature with respect to the C1 word. There were 29 such judgments. The rationale of this analysis was that if pleasant - unpleasant were no more significant than other dimensions, then, for the 29 judgments in which only one of the comparisons had the same marking as the standard, the comparison chosen as similar should be the Fg word (for the F+ condition) and C1 word (for the F- condition), irrespective of the distribution of the markings for pleasant - unpleasant. Table 29 shows the relevant data (using medians because of small ns).

Table 29

RS Comparisons for Words Sharing with Words Contrasting on Pleasant-Unpleasant with Standard

Comparison	Shared or Contrasted with Standard on Pleasant-Unpleasant	N	Number of Cases with RS > 50	Median RS of Given Comparison	Median RS for all Comparisons
(F+) Fg	Shared	9	7	75.9	70.0
	Contrasted	3	0	20.0	
(F-) Cg	Shared	5	5	76.7	30.0
	Contrasted	12	1	23.3	
(Combined) g word[1]	Shared	14	12	76.3	50.0
	Contrasted	15	1	23.2	

[1] Comparison with greater number of common dimensions with standard, irrespective of whether feature is shared or contrasted.

It is obvious from the comparisons in Table 29 that whether a comparison word shared or contrasted with the standard on the pleasant - unpleasant dimension made a difference in whether it was judged similar. In the F+ condition, the median RS for all Fg words was 70.0; but for none of the three cases in which the standard and the Fg word contrasted on pleasant - unpleasant did the Fg word have an RS of greater than 50. For the F- condition, the median RS for all Cg words was 30.0; but for all five of the cases in which the Cg word and the standard shared a feature from pleasant - unpleasant, the RS for the Cg word was greater than 50, with the median RS being 76.7. An overall comparison can be made from the last two rows of the table. One of the interesting outcomes of the overall similarity data was that the median RS of the g-word--i.e., that comparison with either the greater number of shared features or the greater number of contrasted features--was 50, across the two conditions. (This of course reflects the great dependence of the judgment on whether the dimension commonality was due to a shared or contrasted feature.) However, the last two rows of the table show great differences in the RS of the g-word, depending on whether the g-word had a shared or contrasted feature with the standard on the pleasant - unpleasant dimension.

We can thus conclude that for this particular set of words, as a whole, the dimensions were of unequal significance, and particularly that pleasant - unpleasant was quite important. We cannot, however, draw further conclusions about the dimensionality of the words without much additional analysis. A factor analysis would not have been appropriate for this purpose because the number of cases for each dimension was too few, partly because of the number of zero codings. For the present, it must suffice merely to point out the need for some sort of weighting technique, and to recognize

that even with such a technique for a set of words, there would always be smaller subsets of words which could not be described adequately with the dimensionality derived from the larger set.

An additional dimension. A consideration of the thirteen dimensions led to the conclusion that there were other dimensions which could possibly add further semantic discrimination within the set of words. In particular, it was thought that there was lacking an evaluative dimension, which has been found to account for so much of the connotative meaning measured by the semantic differential (Osgood, Suci, and Tannenbaum, 1957). On the assumption that the addition of an evaluative (good - bad) dimension might add further discriminability, each of the 61 words used in the experiment was rated on a good - bad dimension by three judges including the author.[7] The criterion for coding a word as (+), (-), or (0), was whether the word had a connotative content which American middle-class society would judge as good, bad, or indifferent, in a moral sense. There was unanimity among the judges in most cases, and majority judgment was used in those cases in which there was not unanimity.

Attention again is directed to the 30% of the judgment items which were not in the predicted direction. These 18 judgment items were selected for further analysis, to determine whether the evaluative dimension could account for some of the unsuccessful predictions. Table 30 lists these 18 judgment items, along with the markings of the words on the evaluative dimension and the RS scores for each comparison word. Since all items represent unsuccessful predictions, RS score is higher for the F1 word in each case.

---

[7] The assistance of the other two judges, William L. Gekoski and Phillip R. Kingsley, is happily acknowledged.

Table 30

Eighteen Judgment Items Not in Predicted Direction
Coded on Evaluative Dimension
(RS of Each Comparison Underlined)

Standard		Comparisons		Consistency with Evaluative Hypothesis	
PAIN (-)	SHAME (-)	73.3,	AWE (0)	26.7	[+]
JOY (+)	RAGE (-)	3.3,	EXCITEMENT (+)	96.7	[++]
PAIN (-)	DESPAIR (-)	56.7,	HORROR (-)	43.3	[0]
FEAR (-)	AWE (0)	76.7,	SHAME (-)	23.3	[-]
ANXIETY (-)	DETERMINATION (+)	80.0,	JOY (+)	20.0	[0]
DISGUST (-)	RAGE (-)	66.7,	PITY (+)	33.3	[++]
GLEE (+)	ANNOYANCE (-)	53.6,	DISMAY (-)	46.4	[0]
LOATHING (-)	SUSPICION (-)	76.7,	EMBARRASSMENT (-)	23.3	[0]
SADNESS (-)	SORROW (-)	62.1,	DESPAIR (-)	37.9	[0]
SUSPICION (0)	EXCITEMENT (+)	65.5,	PITY (+)	34.5	[0]
SCORN (-)	STUBBORNNESS (-)	53.3,	PITY (+)	46.7	[++]
PITY (+)	FEAR (-)	51.7,	DISTRUST (-)	48.3	[0]
DISTRUST (-)	ANNOYANCE (-)	58.6,	EXPECTANCY (0)	41.4	[+]
ANNOYANCE (-)	SUSPICION (-)	51.7,	BOREDOM (-)	48.3	[0]
CYNICAL (-)	DESPICABLE (-)	69.0,	DELIBERATE (0)	31.0	[+]
VAGUE (-)	SELFISH (-)	64.3,	ACUTE (0)	35.7	[+]
HUMBLE (+)	APPROPRIATE (+)	65.5,	GUILTY (-)	34.5	[++]
HOPEFUL (+)	APPROPRIATE (+)	82.8,	DULL (-)	17.2	[++]

The right-hand column of Table 30 has the symbols [+], [0], and [-] to denote levels of consistency with the hypothesis that the evaluative dimension can account for some of the judgments. Cases in which the standard was marked with the same feature from good - bad as the F1 comparison were categorized at one of three levels of consistency: If the Fg comparison had the same feature from good - bad as the F1 comparison, this was

considered neither consistent or inconsistent [0]; if the Fg comparison were marked (0), then the agreement between F1 and the standard was considered consistent [+]; if the Fg and F1 comparisons had opposite features, then the agreement between F1 and the standard was highly consistent [++]. Likewise, if the F1 word had a contrasting evaluative feature relative to the standard, the case was called inconsistent [-] if the Fg word was (0) on the evaluative dimension, and highly inconsistent [--] if the Fg word shared an evaluative feature with the standard.

The results of the comparisons shown in Table 30 give strong evidence that an evaluative factor was operating. The following statements summarize the results of the consistency analysis. Of the 18 judgments:

(1) None is highly inconsistent with the evaluative hypothesis.

(2) Only one is inconsistent.

(3) Nine are at least consistent, with five highly consistent.

In other words, in 9 out of the 10 cases which can be predicted on the basis of an evaluative dimension, the predictions are correct.

There was, however, the possibility that we had not identified a new dimension, but merely given a double weighting to one already coded. Specifically, there was the possibility that the good - bad dimension is completely redundant with respect to the pleasant - unpleasant dimension. Since the latter had already been shown to have been an important one, it could not be ruled-out as the significant dimension for these 18 words. To check this possibility it was necessary merely to put the codings for pleasant - unpleasant to the same test applied to good - bad. When this was done, only five cases were found in which the standard had a non-zero code on pleasant - unpleasant which agreed with one, but not both, of the comparisons. Of these, three were consistent and two were inconsistent with

the hypothesis that pleasant - unpleasant can account for the judgments. If to these cases are added (1) those in which the standard contrasted with one, but not both, comparisons and (2) those cases in which the standard was (0), but only one comparison was (0), then only four of nine judgments were consistent with the pleasant - unpleasant hypothesis.

It is thus clear that the good - bad dimension adds an important dimension that is not redundant with respect to pleasant - unpleasant. An example from the list of 18 can illustrate the point:

DISGUST: RAGE (66.7) PITY (33.3)

PITY was the Fg word but RAGE was judged more similar. The pleasant - unpleasant dimension does not help here. Both comparisons are (0). But on good - bad RAGE is (-), and PITY (+). Since the standard, DISGUST, is (-), there are two changes in the original feature relationship: The F1 word has an additional shared feature with the standard and the Fg word has an additional contrasting feature with the standard. Similar changes occur for other words from the list of 18. In some cases, a shared feature is added to the F1 word, in others a contrasted feature is added to the Fg word, and in some cases both occurs. In only one case, FEAR: AWE (76.7) SHAME (23.3), would the evaluative dimension add a shared or contrasted feature to the "wrong" comparison. In this case it adds another shared feature to the Fg word, SHAME and FEAR both marked (-), AWE (0). Apparently, for these words the evaluative dimension is less significant than some other dimension.

It is unlikely that other added dimensions would prove as powerful as the evaluative dimension. The point is that other dimensions of meaning can be added to increase the discriminatory precision of the semantic components. It is probable that for a given subset of words, additional dim-

ensions are needed to increase discrimination among the subset, but which would not appreciably increase discriminability for the larger set of words as a whole. The larger the set of words, the greater the number of features needed to achieve adequate componential description.

Summary

This chapter has described an experiment in similarity judgments. Ss judged one of two comparison words as more similar to a standard word. About half the judgments were made under the F+ condition, in which one comparison (the Fg word) shared more features with the standard than did the other comparison (the Fl word), with the number of contrasts equal. The remaining half of the judgments were made under the F- condition, in which one comparison (the Cg word) contrasted with the standard on more features than did the other comparison (the Cl word), with the number of shared features equal.

The results were that the Relative Similarity (RS) measure averaged across judgment items and Ss was equally high for Fg words and Cl words, equally low for Fl words and Cg words. Comparisons with a greater number of shared features or a lesser number of contrasted features were selected as more similar equally often. The conclusion, supported by correlations as well as by the main result, is that Ss were differentially sensitive to shared vs. contrasted features, and did not use total dimension commonality as a source of similarity. It is also probable, based on the results of the F- condition, that Ss tended to "eliminate opposites" in making judgments for cases in which neither comparison seemed similar in meaning.

The system of feature codings employed in the experiment would provide more semantic information, and hence more accurate predictions of similar-

ity, if dimensions were weighted for their semantic significance. For example, pleasant - unpleasant was of above-average importance for this particular group of words. Finally, the addition of an evaluative dimension improved the semantic description of these words by accounting for half the cases in which the similarity judgments were not in the predicted direction.

Chapter 5

Extensions and Speculations

The questions of whether semantic featural analysis can be applied to experimental descriptions of meaning similarity and to studies of word association have been answered, at least tentatively, in the affirmative. Artificial word associations and similarity judgments both have proved to be partially predictable on the basis of semantic features. More specific questions about cognitive aspects of semantic features cannot be answered from the results of these experiments, however. We cannot describe, for example, the psychological nature of semantic features in the associative process. Accordingly, in this final chapter we shall consider some speculations about the possible functions of semantic features and their application to other problems. We shall also consider the relation between the two areas investigated, associations and meaning similarity, and summarily report the results of a fifth experiment.

Meaning Similarity and Word Associations

The question of the relationship between meaning similarity and word associations has been with us since Aristotle's three laws of association: contiguity, similarity, and contrast. Despite modifications by later philosophers, Hume, for example, and both Mills, the three laws can still be found today, even in modern psychology. Contiguity in particular has had a rather easy history of survival owing to its compatibility with the

frequency principle. Indeed, contiguity has usually been regarded since the philosophies of Hume and Berkeley as the fundamental law of association, to which the others are reducable.

Nonetheless, similarity and contrast have also survived. With modern psychology, association itself became the "fundamental law", thus enabling similarity and contrast to assume descriptive, rather than causal, status. The classification of association responses by Woodrow and Lowell (1916) showed similarity to be one of the three major types of adult associations (contrast and coordination being the other two), and it has continued to be regarded as important.

For example, a study by Riegel, Riegel, Quarterman and Smith (1967) is representative of studies using the restricted association tasks of Riegel (1965) as associative approaches to semantics. The relevant finding of this study from the present point of view is that the responses given by college Ss under instructions to give Similars accounted for about 40% of the responses given in free association. As a group, Coordinates, Similars, and Contrasts, while overlapping a great deal with each other, accounted for most of the free associations. This was contrary to the results for younger children, whose free associations overlapped primarily with Parts, Functions, Qualities and Superordinates. This finding is consistent with previous results regarding age differences in paradigmatic responding. For older children and adults word associations tend to be same form-class, semantically similar words. The implication may be that free associations are reflections of semantic structures, and it is the structures themselves which change with linguistic development.

Other studies, representing various approaches, have assumed that word associations are determined by, or at least reflect, semantic similarity

(e.g., Pollio, 1963; Geer and Mallenauer, 1964). Numerous other experiments--e.g., Kjeldergaard (1962), Siipola et al (1955), Moral, Mefferd, and Kimble (1964)--have shown that semantic "opposites" as well as similars account for many associations. (In our view, "opposites" are frequently minimal contrasts, hence "similars".) Wynne, Gerjuoy and Schiffman (1965) have interpreted such findings as suggesting that a principle of least conceptual effort is involved in free word associations, and that contrast responding may represent a least conceptual effort. With a slight extension, this represents a significant statement about association: Similarity (or contrast) is a determinant of association under certain task conditions, namely those which are conducive to least conceptual effort on the part of S. (The influence of task instructions, such as speed set, on associations support this contention.) Words which contrast minimally may generally meet the requirement of least effort.

Although it apparently can be assumed that semantic similarity plays a role in word association, whether similarity itself can in every case be explained by more "basic" principles, such as contiguity and frequency, is another question, beyond the present discussion.

We have suggested that semantic features may be useful in describing word association and in describing meaning similarity, and this we do without proposing a marriage of similarity with association. From our point of view, the most important fact about word associations is that they are multiply-determined. Semantic features have an obvious application to meaning and hence to meaning similarity. Generally speaking they have the same application to association--that is, we can describe the words involved componentially. But the words will not always be minimal contrasts; some will be syntagmatic; some will be sequential completion; etc.

We probably would be correct in proposing that associations always share at least one semantic feature, even if the associations are not paradigmatic. For syntagmatic associations such as EAT - FOOD there is semantic similarity involving the selection restrictions of the verb, which include the semantic markers of the noun. (We would not be likely to observe EAT - INTUITION which lacks this matching of selection restriction and semantic marker.) It is beside the point to wonder whether EAT - FOOD is observed because the two words have a history of sentential contiguity. Their sentential contiguity is made possible because the features of the noun satisfy the selection restrictions of the verb, a point to which we shall return later.

Experiment 5

As an exploratory step toward determining whether English word associations would behave in the same manner as the artificial associations of the first three experiments, a fifth experiment was undertaken. The purpose was simply to collect free word associations to words for which semantic features already were available, that is, associations to words from Osgood's list. The question of interest was whether there would be any tendency for the associations to be minimal contrasts with respect to the dimensions identified by Osgood. The results have been analyzed rather coarsely and are reported here in summary form.

Method. Forty of the 70 emotional words from Osgood's list were used as test stimuli, along with 60 filler words, making a total of 100 stimuli. The filler words were relatively non-emotional words matched with the test words for form class and frequency. Most of the words had a frequency of fewer than 50 per million in Thorndike-Lorge (1944) general word count.

The stimuli were given in a booklet with five words per page. The order of the pages was randomized, with the constraint that every page be first at least once. Both the liberal use of filler words and the randomization procedure were employed to reduce context bias, which might otherwise have increased the probability of emotion-word responses. Thirty new Ss were used for this experiment.

Results. The main results for the 40 emotional words can be outlined as follows:

(1) About 20% of the responses (tokens) to the 40 stimuli could be found on the original list of 70 emotional words. This included a few words which could be added to the list by form-class extrapolation--for example, ANGRY was observed as an associate, and although it was not on the list of 70, ANGER was; thus ANGRY was assumed to have the same feature codings as an adjective as ANGER did as a noun.

(2) The primary, secondary, and tertiary responses, in terms of frequency, were recorded for further consideration. Subsequent statements about results refer only to these three most frequent (PST) responses, whose total frequency across the 40 stimuli accounted for slightly fewer than half of the total token responses.

(3) For the PST words only, about 27% of the tokens appeared on the list. For the 73% which did not, a compatibility check with the 14 dimensions (Osgood's 13 plus evaluative) was made. It was found that 88% of these responses could be reasonably coded as emotion words on these dimensions.[1] Thus approximately 91% (27% + 88% [73%]) of the PST responses can

---

[1] Although this was a subjective task, most cases seemed beyond debate. Examples of words considered codable were HAPPINESS, INNOCENT, STRONG, HAUGHTY, and HATE. Examples of words considered not codable were MULE, PARTY, MOVIE, and STORM.

be considered potentially codable on the 14 dimensions, and hence are potential minimal contrasts.

(3) Out of the approximately 40 x 3 = 120 PST responses, 32 were words from the list of 70. The feature codings for these 32 words were compared with the feature codings of their respective stimuli as a test for minimal contrast. Of these, 20 could be classified as minimal contrasts. A response was classified as minimal contrast if it had 0 or 1 contrast with the stimulus and shared at least one feature. The 32 cases which could be tested for minimal contrast are shown in Table 31 (see page 110).

A few observations about the results in Table 31 are in order. First, HOT - COLD could be excluded. While they probably are minimal contrasts, the dimensions relevant to the physical senses of HOT and COLD may more readily account for their association than the dimensions for their emotional senses.

Secondly, notice that some stimuli, namely PITY, DREAD, ANXIETY, SADNESS and DISMAY appear twice. For example, PITY evoked SORROW as its primary response and SHAME as its secondary response. The first pair represents a minimal contrast, the second does not. Only DREAD had minimal contrast responses in both cases (FEAR and WORRY). This serves to illustrate that, for a nonexhaustive feature catalogue, a minimal contrast hypothesis does not predict one response only, so long as the prediction is based on the minimal number of contrasted features and not on the maximum number of shared features.

This again raises the point that we have not been able to decide whether the maximum number of shared features or the minimal number of contrasted features is the more important variable. We learned from Experiment 4 that

Table 31

32 Associations Involving Words from Osgood's List

Stimulus	Observed Response	Example of Minimal Contrast if different from Observed Response
Humble	Proud [+9; -6,8]	Appreciative [+0,6,9,11]
Pity	Sorrow [+7,11; -0]	-
Disgust	Anger [+0,4,7]	-
Hot	Cold	-
Dread	Fear [+0,1,2,3,7]	-
Rage	Anger [+0]	-
Determination	Stubbornness [+3,6,7,11; -0]	-
Anxiety	Worry [+0,1,7; -5]	Fear [+0,1,2,3,5,7]
Cynical	Scornful [+0,3,6,8,9; -13]	-
Sadness	Sorrow [+0,1,11]	-
Contempt	Scorn [+0,3,6,9]	-
Glee	Joy [+0,1; -5,13]	Interest [+0,5,10]
Disgust	Contempt [+0,6]	-
Eager	Anxious [+2,7; -1]	Hopeful [+1,6,7,11]
Acute	Pain [+3,12]	-
Pity	Shame [+7,9; -0,8]	Disgust [+4,6,7,8; -0]
Excitement	Joy [+0,2,3]	-
Dread	Worry [+0,1,7]	-
Dismay	Worry [+0,1,5]	-
Awe	Amazement [+3,4,6,11; -2]	-
Anxiety	Fear [+0,1,2,3,5,7]	-
Loathing	Disgust [+0,1,4,6,7; -10]	Rage [+0,2,3,5,6,13]
Violent	Angry [+0]	-
Sadness	Joy [+5; -0,1,2,13]	Sorrow [+0,1,11]
Vague	Uncertain [+0,3,4,9,12,13; -10]	-
Contentment	Joy [+0,1; -2,13]	Complacency [+1,2,4,10,11,13; -0]
Boredom	Dull [+0,1,2,10,11,13]	-
Joy	Sorrow [+3; -0,1]	Excitement [+0,2,5]
Suspicion	Fear [+0; -5,7]	Distrust [+0,5,6,9; -7]
Scorn	Anger [+0,4]	-
Dismay	Sad [+0,1,9,11,13;-4,5,10]	Puzzlement [+5,6,9,10,13; -11]
Proud	Humble [+9; -6,8]	Cynical [+3,6,8,9]

Note: The numbers in brackets indicate the dimensions of the shared (+) and contrasted (-) features between stimulus and response. (See the key on p. 76.) Dimension [0] is the evaluative dimension. The first 9 associations in the list represent primaries, the remainder are secondary and tertiary.

they are equally important for meaning similarity when manipulated independently. In the artificial association experiments, the two variables were confounded. It is reasonable to argue that there is an interaction between the number of shared and contrasted features, or at least that one contrast is necessary for association. This would follow from the assumption that only words of identical meaning have no contrasts. Where cases of no contrast exist it is a deficiency in feature assignment.

In any case this question remains beyond the present results. Experiment 5 cannot bring much evidence to bear on the minimal contrast hypothesis, although we may assume that the result would be improved by coding of additional words. We shall also assume that minimal contrast is no simple thing. For example, features probably should be weighted for the prediction of English associations, as well as for meaning similarity. In Experiment 5, the evaluative dimension seems to have been very important, as an inspection of Table 31 suggests. We would also speculate that the word in isolation has certain discriminably salient features, but that the word in the context of a given sentence may have different features which are salient. Or more generally, the feature salience of a word is a function of context. With the acknowledgement of these kinds of complications, we will leave the topic of association and turn to speculations about other aspects of semantic features.

Further Aspects of Semantic Features

We shall conclude by exploring a few selected points of contact which can be made between semantic features and other topics. First we shall consider some syntactic and cognitive aspects and then turn to a brief discussion of additional extensions.

Syntactic Aspects

In the present experiments, we have dealt exclusively with words in isolation. However, a word within a syntactic structure is also to be defined by features, both syntactic and semantic. Indeed, the formulation of semantic markers by Katz and Fodor (1963) appears to serve the need to include certain kinds of information as features of words, rather than features of a grammar. We will briefly consider the relationship between syntactic structures and semantic features and the psychological implication of such a relationship.

The formal generative grammars of Chomsky (1957, 1965) provide comprehensive treatments of syntactic features. In the more recent version of the transformational generative grammar (Chomsky, 1965), the question of whether certain syntactical-semantic features are best considered as rules of the grammar or as features of words in the lexicon is left open. The possibility that such features may be included in the lexicon is what has made it possible to consider a general semantic theory, as Katz and Fodor (1963) and Katz and Postal (1965) have done, as a complementary and non-redundant component of the complete description of a language. In other words, the Katz and Fodor (1963) assumption that "Linguistic Description Minus Grammar Equals Semantics" (p. 483), which essentially defines the domain of semantics as anything which cannot be accounted for in the grammar, seems to be correct only if the semantic markers which compose the heart of their theory can be assumed to more appropriately lie outside the grammatical description.

More important is that, regardless of where these syntactic-semantic features are placed, they play an important part in syntactic structures. If we assume for a moment that they belong in the grammar after all, we can

take advantage of Chomsky's distinction between two types of syntactical restriction rules. The first is the strict subcategorization rules which assign major grammatical categories to symbols. For example, the major grammatical markers such as [noun] and [verb] are sub-categories and their assignment to symbols are by strict-subcategorization rules. The second type of rule is the selection restriction rules which subcategorize lexical categories by syntactical context-sensitive features. An example of a selection restriction rule would be the constraint on the verb of a noun-subject which has a certain syntactic feature, for example [+Abstract].[2] If the subject is [+Abstract], a selection rule would restrict the choice of the verb by excluding such possibilities as FRIGHTEN or THINK which are [-Abstract]. In addition, THINK is also [+Human] and hence could not be selected as the verb for a [-Human] subject.

The selectional qualities of these features is not altered by considering them semantic rather than syntactic, and this is exactly the point. A sentence such as Hope thinks eternal is equally anomalous regardless of whether the features which make it so are in the lexicon or the grammar. If such features can properly be considered semantic, then we have need for the projection rules of Katz and Fodor (1963) to interpret, to paraphrase, and to detect ambiguity and anomaly.

Much of the point here is that there is a difference between the type of features illustrated above and the type of features which we have employed in our experiments. This difference seems to be exactly the difference between what may be termed "pragmatic" features and syntactic-

---

[2] Brackets are used to distinguish syntactic features from semantic features which are noted by parentheses. The distinction may be unnecessary, since the point here is that the same features may be included in the lexicon.

semantic features. The nature of this difference can be seen by considering the kinds of anomalies which result from violations of the two types of feature requirements. Hope thinks eternal is anomalous because, as we have seen, there are selection restrictions placed on the verb, one of which is that it occur with a (+Human) subject. Likewise, The boy frightens sincerity is anomalous because a selection restriction on the verb frighten includes that it take objects which are (+Animate). The detection of such anomalies rest on the operation of linguistic rules, whether considered strictly grammatical or as involving semantic projection rules. Consider now a phrase such as green crows. We would hesitate to call this phrase anomalous in the previous sense, but we would recognize an irregularity. Whereas, upon hearing The boy frightens sincerity we assume we are witnessing either nonsense or a linguistic example (which depends upon our training), we do not so dismiss He likes green crows. Instead we may ask questions, such as "Are there birds which are crows and also green?" (Notice we do not ask, "Is there an abstract quality which is sincerity and which can be frightened?") Also we might ask, "How did the crows become green?" but not "How did sincerity become frightened?" (not even, "How did the boy do it?").

The nature of these different types of anomaly is that the semantic markers and projection rules of Katz and Fodor detect anomaly which results from violations of linguistic rules, while pragmatic features detect anomaly which results from "violations" of empirical knowledge. However, both types of features exist within syntactic structures and both have syntactic consequences in the broad sense; i.e., we are unlikely to hear either (1) hope thinks or (2) green crows, although for different reasons. In the case of (1), the selection restriction of think, viz., (+Human), is

not included in the semantic markers for hope, viz., (-Human). In the case of (2), empirical knowledge about the color of crows, coded as the semantic feature (+Black), excludes the adjective green. The anomaly of (2), unlike that of (1) can be removed by alterations in the environment or by new knowledge--for example, by a dash of green paint or the discovery of a new species. Constraints of the green crow type can be considered aspects of empirical knowledge, while those of the hope thinks type can be considered as aspects of linguistic competence (i.e., the result of the operation of an implicit linguistic rule, either of grammar or of semantics).

Additional Extensions

It may be possible to extend the application of semantic features to other cognitive areas. To briefly mention a few of these, there are at least two broad areas of contact. (1) The experimental concerns of verbal learning--learning, transfer, and retention. (2) The general area of sentence usage and comprehension--the communication of meanings. Under (1), there are possibilities that semantic features may be useful in identifying variables such as similarity and codability. For example, to the extent that common semantic features can identify similar meanings, they ought also to predict such things as facilitation in paired associate learning, generalization effects and transfer of training, interference effects, and other effects that have been shown to be a function of semantic or secondary similarity.

In addition, common semantic features may be a factor in coding and organization. For example, the well-established observation that words recalled from an unstructured free-recall list tend to be recalled in orders which reflect an organization of items under common categories may be ex-

pressed in terms of semantic features, at least in some cases. In the case of the artificial words in our experiments, it appears that free recall did not reflect feature organization to the extent that associations did. However, the semantic organization which reportedly occurs for real words may be accounted for by common semantic features.[3]

Also, contextual organization such as that demonstrated by Miller and Selfridge (1950) or Marks and Miller (1964), may be described in terms of common semantic features. The syntactically structured but semantically anomalous strings of Marks and Miller (1964), which are more poorly recalled than meaningful sentences, may be said to differ from meaningful sentences only in the lack of shared features between successive elements in a sentence. These represent only a few examples of instances in which semantic features make at least conceptual contact with variables in verbal learning. It must be frankly acknowledged, however, that methodological contact is more difficult in the absence of a satisfactory procedure for feature coding.

Under (2), the communication of meanings, there may be a general approach to language based partly on the notions of selection restrictions, discussed previously. We could suggest, for example, measures of information transmission based on feature commonalities between sentence elements. The transitional probabilities of sentence elements, the use of which has been derived from notions of information theory, to measure the information value of sentences (e.g. Miller, 1951), may be analyzed in terms of shared features. In fact a semantic featural analysis may be able to mea-

---

[3] There is some controversy over whether this organization is to be accounted for solely on the basis of inter-item associative strength. Cofer (1965) has concluded that the processes involved are not yet fully understood, but that categorization is a factor beyond mere association.

sure concepts like information and redundancy in a valuably different way from transitional probabilities.

For example, a phrase such as dirty dirt would be measured as having high information value by the traditional measure, because it would probably never be observed; the transitional probability of the first word to the second is close to zero. But a proper featural measure would show the phrase dirty dirt to be virtually redundant; the adjective has no semantic feature that the noun lacks, hence there is no information. Only an adjective which has some feature which the noun lacks can be a true "modifier" for that noun. One may even suppose that there is some optimum number of shared features to produce enough information and enough redundancy for efficient communication.

Such an approach may also allow for descriptions of meaning change and metaphor. We might assume, for example, that the word LEG at one time contained semantic markers something like the following: (animate) → (vertical appendage) → (support) → < used in ambulation >[4]. Thus we had the leg of a dog. In this account we assume metaphor is created when a match between semantic marker and selection restriction is made after deleting one of the markers, in this case (animate). This leaves the selection restriction to be matched with (vertical appendage) → (support), thus creating metaphors such as leg of a table. Since repeated use of a metaphor diminishes its creative, and hence metaphoric, quality, we are left with a new meaning: Now (animate) is not at the first node, but rather

---

[4] Acute-angle brackets enclose semantic distinguishers which comprised the unique sense-characterization of a word in the original Katz and Fodor (1963) treatment. Bollinger (1965) has made some general criticisms of the Katz and Fodor theory, including the contention that the distinction between markers and distinguishers is arbitrary. In the most recent version of the theory, Katz (1966), the distinguishers have disappeared.

has been "displaced" after (vertical appendage). Another marker, (inanimate) now heads a new path for the previously metaphoric reading which has now become commonplace and part of the regular lexical entry. This illustration, which deviates only slightly from a strict Katz and Fodor account, nevertheless demonstrates the potential of this general type of formalized semantic featural approach to questions of metaphor and meaning change.

Summary

In this concluding chapter we have (1) discussed the relationship between association and meaning similarity, and (2) pointed to selected areas to which semantic featural analyses may be extended. We have taken the position that both meaning similarity and association can be studied through semantic featural analyses, but that associations are not accounted for by similarity. We have assumed that associations are multiply determined, and that under certain conditions, possibly those which lead the subject to minimize conceptual effort, they may be partially accounted for by minimal contrast. The results of a fifth experiment which extended the featural analysis of associations to English words were reported as exploratory.

It was again pointed out that cognitive processes involved in the association of artificial words could not be inferred on the basis of any of the present experiments. It is even impossible to select between minimal number of contrasting features and maximal number of shared features as the appropriate description of the data. The discussion was extended to the relationship between semantic and syntactic features. A distinction was made between features which are aspects of linguistic competence

(syntactic-semantic features) and features which are aspects of empirical knowledge (pragmatic features). Also, we suggested that, at least conceptually, semantic featural analysis could be extended to problems of verbal learning, transfer, and retention. Finally, we discussed the possibility that such an analysis may be a fruitful approach to such general linguistic topics as information, meaning change, and metaphor.

Appendix 1

Instructions to Subjects

for First Three Experiments

(1. Word-feature acquisition)

This experiment is concerned with certain processes involved in learning abstract concepts in an artificial situation. The concepts are artificial words, and your task is to learn the experimentally imposed attributes of each concept.

Before we begin, let's go over the names of the concepts so that you will know how to pronounce them. (Pronounce words for S, then have S pronounce them.)

The name of the concept appears on one side of one of these cards (show card). In addition to the name of the concept, there is, on each card, two attributes, one of which applies to the particular concept. When you see the card you are to say the name of the concept aloud, and then guess which of the attributes belongs to the concept. You will then be shown, on the reverse side of the card, a statement giving the correct answer; read this aloud also. To clarify the task, here is an example: (Example given of GOJEY, regular or irregular?)

The first time, however, instead of guessing which attributes belong to each concept you will see, on one card, the concept with all its attributes. This is a study trial. You are to read aloud the information on the card. Other study trials will occur at other points during the task.

As we go through the cards you will see that all the concepts have several attributes and that there is no logical way of knowing which concepts have which attributes. For this reason it is a fairly tedious task

and will probably require a good deal of concentration. We will go through the stack of cards as many times as is necessary for you to learn all the concepts.

Are there any questions?

(2. Feature Recall)

Now, when I say a concept-word, you are to say all its attributes--that is, all those which you have just learned (for example, "wet", "dry", etc.). Each word has three attributes. Ready?

(3. Concept Recall)

Now, I want you to recall out loud all the six concepts that you have been learning in any order at all.

(4. Association)

This is the final task. I am going to say a concept-name, and you are to give the first other artificial concept-name that comes to mind. (You must respond as quickly as possible.)[1] There is no right or wrong answer, but your response must be one of the artificial concept-words that you have just learned.

[1]This instruction given to half of the Ss in Experiment 1 only. All other Ss were told "You are not asked to respond quickly. You may take your time. The times recorded from the clock are just for the record."

Appendix II

Four Sets of Similarity Judgments

As Presented to Ss in Experiment 4

With Percentage of Ss Selecting Each Comparison

Nouns: Set A

joy:	96.7	excitement [+2,3]*	3.3	rage [+2,3,5,13]
surprise:	3.3	stubbornness [-3,6,10]	96.7	suspicion [-6]
sadness:	26.7	determination [-11]	73.3	scorn [-4,9,13]
bewilderment:	26.7	loathing [+1,3; -5,6,8,9,10]	73.3	interest [+5,10; -6]
amazement:	20.0	pain [+3,6]	80.0	dismay [+3,4,6,11]
adoration:	40.0	anxiety [-1,2,9]	60.0	expectancy [-9]
horror:	96.7	puzzlement [+6,9; -5,13]	3.3	contentment [+9; -1,2,4,13]
excitement:	96.7	surprise [+2,3,10]	3.3	disgust [+10]
determination:	63.3	glee [none]	36.7	sorrow [-3,7,11]
embarrassment:	20.0	contempt [+9; -3]	80.0	bewilderment [+1,3,8; -9]
shame:	40.0	pity [+7,9; -8]	60.0	loathing [+1,3,5,9,10; -8]
suspicion:	30.0	dread [-7]	70.0	anxiety [-5,9]
interest:	13.3	complacency [-4,10]	86.7	adoration [-10]
stubbornness:	73.3	fear [+7; -3]	26.6	puzzlement [+11; -6,9,10]
boredom:	90.0	sadness [+1,2,10,11,13]	10.0	embarrassment [+1]
contentment:	70.0	sorrow [+11; -1]	30.0	annoyance [+13; -1,4,10]
contempt:	50.0	disgust [+6]	50.0	scorn [+3,6,9]
sulkiness:	40.0	dismay [-4,9,10,11]	60.0	worry [none]
pain:	43.3	horror [+1,3,5,6,9]	56.7	despair [+1,3,9]

(Nouns: Set A continued)

dread:	90.0	distrust [+7]*	10.0	determination [+7; -3]
anxiety:	76.7	bewilderment [+1,3,9; -5]	23.3	rage [+2,3,5]
fear:	23.3	shame [+1,3,5; -7]	76.7	awe [+3; -2]
scorn:	46.7	pity [+4,6,8,9]	53.3	stubbornness [+3,6,9]
loathing:	20.0	excitement [+2,3,11; -10]	80.0	sadness [+1,5,10; -2,4,9,11,13]
dismay:	23.3	contentment [+9,11,13; -1,4,10]	76.7	surprise [+3,6,10]
odor:	100.0	smell	0.0	anger
annoyance:	51.7	suspicion [-6]	48.3	boredom [+1,13; -10]
disgust:	27.6	glee [+10; -1]	72.4	complacency [+8; -1,4,10]
puzzlement:	72.4	interest [+5,10; -6]	27.6	embarrassment [-9]
rage:	24.1	amazement [+2,3; -6]	75.9	horror [+2,3,5,13; -6]
pity:	51.7	fear [-7]	48.3	distrust [+6,9; -7]
despair:	75.9	sulkiness [+4,10; -9,11]	24.1	contempt [-3,9]
awe:	86.2	adoration [+2,8,10,11; -9]	13.8	joy [+3; -2]
distrust:	41.4	expectancy [+5,6,7; -9]	58.6	annoyance [-6]
complacency:	96.6	boredom [+2,10,11,13; -1,12]	3.4	pain [-1,12]
worry:	0.0	joy [-1,5]	100.0	sulkiness [none]
glee:	86.2	awe [-10]	13.8	dread [-1,10,13]
expectancy:	10.3	shame [-5,7,9]	89.7	amazement [-6]
sorrow:	13.8	worry [+1; -7]	86.2	despair [+1,3,11; -7]

Nouns: Set B

shame:	70.0	disgust [+1,7; -8,10]*	30.0	sulkiness [+9,10]
distrust:	56.7	horror [-5,6,9]	43.3	amazement [+5,6,7; -9]
boredom:	80.0	interest [-10]	20.0	excitement [-2,10,11]
joy:	23.3	sadness [+5; -1,2,13]	76.7	adoration [+1; -2]
determination:	100.0	stubbornness [+3,6,7,11]	0.0	contempt [+3,6]
annoyance:	26.7	complacency [+13; -1,4,10]	73.3	fear [+1]
rage:	36.7	puzzlement [-5,6,13]	63.3	worry [-5]
pain:	26.7	awe [+3,6,9]	73.3	shame [+1,3,5; -9]
anxiety:	20.0	joy [+2,3,5; -1]	80.0	determination [+7; -3]
bewilderment:	30.0	surprise [+3,6,10]	70.0	dismay [+1,3,5,6,9,10]
amazement:	83.3	glee [none]	16.7	contempt [-3,6]
despair:	66.7	fear [+1,3,7]	33.3	loathing [+1,3,10; -4,8,9,11]
worry:	6.7	boredom [+1]	93.3	expectancy [+5,7]
adoration:	67.9	dread [+8.10; -1,2]	32.1	annoyance [-1,10]
pity:	60.0	determination [+6; -7,11]	40.0	rage [+6]
horror:	60.0	shame [+1,3,5; -9]	40.0	sadness [+1,5,9; -2,4,13]
complacency:	42.9	embarrassment [-1,4,8]	57.1	distrust [none]
excitement:	90.0	adoration [-2,10,11]	10.0	scorn [-3]
fear:	83.3	sulkiness [none]	16.7	glee [-1,5]

(Nouns: Set B continued)

contempt:	50.0	bewilderment [-3,6,9]*	50.0	boredom [none]
dread:	46.7	pain [+1,3]	53.3	despair [+1,3,7,8,10]
awe:	93.3	interest [+4; -6,10]	6.7	stubbornness [+10; -3,6,9,11]
disgust:	33.3	pity [+4,6,7,8]	66.7	rage [+6]
surprise:	16.7	sorrow [+3]	83.3	horror [+2,3,4,6]
glee:	53.6	annoyance [+10,13; -1]	46.4	dismay [+5,10,13; -1]
doubt:	0.0	courtesy	100.0	skepticism
loathing:	23.3	embarrassment [+1,3,4,9; -8]	76.7	suspicion [+6,9; -5]
sulkiness:	37.0	amazement [-4,11]	63.0	contentment [+4,10; -9,11]
scorn:	24.1	joy [+13; -3]	75.9	awe [+4; -3,6,8,9]
sadness:	62.1	sorrow [+1,11]	37.9	despair [+1,4,9,10,11]
expectancy:	37.9	puzzlement [+5,9,12,13; -6]	62.1	dread [+7; -13]
stubbornness:	72.4	distrust [+6,7,9]	27.6	worry [+7]
suspicion:	65.5	excitement [none]	34.5	pity [+6,7,9]
interest:	64.3	disgust [+4,6,10]	35.7	scorn [+4,6]
embarrassment:	34.5	pain [+1,3; -9]	65.5	anxiety [+1,3; -4,9]
puzzlement:	6.9	contentment [+9,13; -10,11]	93.1	surprise [+6,10]
sorrow:	31.0	loathing [+1,3; -11]	69.0	bewilderment [+1,3]
dismay:	24.1	suspicion [+5; -6,9]	75.9	anxiety [+1,3,9; -4,5]
contentment:	10.3	expectancy [+9,13]	89.7	complacency [+1,2,4,10,11,13]

Adjectives: Set C

hot:	96.7	eager [+2,6]*	3.3	vague [+3,12; -10,13]
appropriate:	0.0	desperate [none]	100.0	considerate [+4,9,12]
cynical:	93.3	selfish [+6,9]	6.7	humble [+9,11; -6,8]
sudden:	100.0	violent [+2,3,10]	0.0	despicable [none]
firm:	73.3	proud [+3,6,8]	26.7	satisfying [none]
emphatic:	6.7	considerate [+4,6; -13]	93.3	hot [+2,6,10,13; -3]
uncertain:	10.0	acute [+3,10,12; -13]	90.0	hopeful [+7,9,12; -1,10]
desperate:	23.3	deliberate [+7,11; -3,8]	76.7	guilty [+1,8]
satisfying:	93.3	appreciative [+1,7]	6.7	proud [+1]
considerate:	50.0	sudden [none]	50.0	cynical [+5,6,9,13]
deliberate:	33.3	appropriate [-12]	66.7	emphatic [+3,6,8,11; -10]
imaginative:	100.0	creative	0.0	cold
dull:	16.7	cynical [+11,13; -9]	83.3	uncertain [+1,4,9,13; -10]
despicable:	72.4	vague [-9,12,13]	27.6	satisfying [-12]
guilty:	69.0	acute [none]	31.0	eager [-1,6]
humble:	13.3	sudden [none]	86.7	deliberate [-6,11]
appreciative:	30.0	uncertain [-1,4,7,9]	70.0	firm [-6,11]
eager:	86.7	hopeful [+1,6,7,11]	13.3	firm [+6,11]
hopeful:	82.8	appropriate [-9,12]	17.2	dull [+9,10; -1,11]

(Adjectives: Set C continued)

selfish:	90.0 despicable [+6,9,12]*	10.0 violent [none]
vague:	14.3 emphatic [-3,4,10,13]	85.7 appreciative [-4,9]
proud:	76.7 hot [+6; -3]	23.3 guilty [+9; -1,6,8]
violent:	0.0 humble [none]	100.0 desperate [+2,3,13]
acute:	10.0 dull [-2,10,13]	90.0 selfish [-12]

Adjectives: Set D

violent:	66.7	firm [-3,5,10]*
cynical:	69.0	despicable [+6,9; -13]
vague:	64.3	selfish [-9,12]
satisfying:	10.0	desperate [-1,7]
hot:	80.0	violent [+2,3,5,10,13]
deliberate:	90.0	proud [+3,6,8]
hopeful:	13.8	acute [+12; - 10]
appropriate:	10.0	vague [-4,9,12]
acute:	100.0	emphatic [+2,10,13; -3]
dull:	76.7	satisfying [-1]
eager:	76.7	deliberate [+6,7,11]
humble:	65.5	appropriate [+9]
uncertain:	69.0	hot [+3,10,12; -5,13]
sick:	100.0	ill
despicable:	6.9	hopeful [+6; -9,12]
guilty:	43.3	sudden [none]
appreciative:	76.7	humble [+6,9,11]
proud:	46.7	considerate [+6,9]
selfish:	83.3	eager [+6]
	33.3	proud [-3]
	31.0	deliberate [+3,5,6,8; -11]
	35.7	acute [+3,12; -10,13]
	90.0	humble [none]
	20.0	sudden [+2,3,10]
	10.0	uncertain [+5,7,12; -3,8,10]
	86.2	appreciative [+1; -6,7,9,11]
	90.0	firm [none]
	0.0	despicable [+13; -12]
	23.3	violent [-2,10,13]
	23.3	considerate [+6]
	34.5	guilty [+6,8,9]
	31.0	satisfying [+12; -1,7]
	0.0	sincere
	93.1	desperate [+13]
	56.7	emphatic [-6,8]
	23.3	appropriate [+4,9]
	53.3	cynical [+3,6,8,9]
	16.7	hopeful [+6; -9,12]

(Adjectives: Set D continued)

sudden:	75.9 appreciative [none]*	24.1 dull [-2,10]
considerate:	58.6 vague [+13; -4,9,12]	41.4 guilty [+9; -6]
desperate:	70.0 uncertain [+1,3,7,8]	30.0 selfish [none]
emphatic:	20.0 cynical [+3,6,8; -11,13]	80.0 eager [+2,6,11]
firm:	26.7 dull [+10; -11]	73.3 hot [+6; -3,5,10]

* Enclosed in brackets are the dimensions of shared features [+] and contrasted features [-] between standard and comparison. In Set A, judgments to HORROR were excluded from analysis, since an error resulted in comparisons which varied both in the number of shared features and in the number of contrasted features.

Appendix III

Instructions to Subjects for Experiment 4

I N S T R U C T I O N S

This study is part of a project in which we are attempting to measure meaning similarity. You can be of great assistance in this project by making judgments of meaning similarity between pairs of words. Your judgments will be the only basis from which we can measure the similarity of these words, so we ask you to cooperate with us in this task by making careful and considered judgments. If, after beginning the judgment task, you feel that it is too demanding on your patience, you may be excused from completing the experiment. However, if you find word meaning interesting, you should enjoy making the judgments.

The judgment task is set up as follows: The standard word appears in the extreme left-hand column of the page, followed by a colon. There are then two other words, each preceded by a blank. You are to mark an X on the line next to the word which you think is more similar in meaning to the standard (the word at the extreme left). An illustration is given below:

blue: __X__ green ____ large

Here the standard word is "blue," and your judgment is to select either "green" or "large" as being more similar to "blue." If your judgment were "green," you would place an X in the manner shown in the illustration. In most cases, neither choice will seem to be very similar to the standard word, but you are urged to make careful judgments, nonetheless. Even in cases where neither choice seems at all similar, there may be a basis for selection. Consider the following verbs:

to love: ____ to run ____ to ask

Although neither choice is at all similar to the verb "to love," the verb

"to ask" is similar at least to the extent that the action implied by both verbs requires a second person in addition to the actor, and the choice could be made on this basis, or on a different one altogether. The main point is that even when both choices seem equally unrelated to the word on the left, you should still make the effort to decide which one has more in common.

There are four sets of words, each labelled according to the part of speech of the words involved. All words within a set belong to the same part of speech. Furthermore, all the words in each set have a mild emotional content.

You are now ready to begin. The most important thing in making the judgments is to know the meaning of each word and to use that meaning to determine which word is more similar to the standard word. If there is a word whose meaning you are not sure of, mark a circle around that word and omit judgment of that item.

REFERENCES

Berlyne, D. E. Mediating responses: A note on Fodor's criticisms. J. verb. Learn. verb. Behav., 1966, 5, 408-411.

Bollinger, D. The atomization of meaning. Language, 1965, 41, 555-573.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. J. gen. Psychol., 1953, 49, 229-240.

Bousfield, W. A. The problem of meaning in verbal learning. In Verbal learning and verbal behavior, C. N. Cofer (Ed.), New York: McGraw-Hill, 1961.

Brown, R. and Bellugi, U. Three processes in the child's acquisition of syntax. In New directions in the study of language, Eric Lenneberg (Ed.), Cambridge, Mass.: MIT Press, 1964.

Brown, R. and Fraser, C. The acquisition of syntax. In The acquisition of language, Ursula Bellugi and Roger Brown (Eds.), Monogr. Soc. Res. Child Develpm., 1964, 29, 43-79.

Bruner, J. S., Goodnow, J. J. and Austin, G. A. A study of thinking. New York: Wiley, 1956.

Chomsky, N. Syntactic Structures. The Hague: Mouton & Co., 1957.

Chomsky, N. Review of Skinner's Verbal behavior. Language, 1959, 35, 26-58.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press, 1965.

Cofer, C. N. (Ed.) Verbal learning and verbal behavior. New York: McGraw-Hill, 1961.

Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychol., 1965, 20, 261-272.

Conklin, H. C. Hanunoo color categories. Southwest J. Anthro., 1955, 11, 339-344.

Cook, T. H., Mefferd, R. B. Jr., and Wieland, B. A. Effects of idiodynamic set and stimulus characteristics on numeral homonym associates. Psychol. Rep., 1965, 17(2), 583-589.

Deese, J. Form class and the determinants of association. J. verb. Learn. verb. Behav., 1962a, 1, 79-84.

Deese, J. On the structure of associative meaning. Psychol. Rev., 1962b, 69, 161-175.

Deese, J. The associative structure of some common English adjectives. J. verb. Learn. verb. Behav., 1964, 3, 347-354.

Entwisle, D. R. Word associations of young children. Baltimore: The Johns Hopkins Press, 1966.

Entwisle, Doris R., Forsyth, D. F. and Muuss, R. The syntactic-paradigmatic shift in children's word associations. J. verb. Learn. verb. Behav., 1964, 3, 19-29.

Ervin, S. M. Changes with age in the verbal determinants of word-association. Amer. J. Psychol., 1961, 74, 361-372.

Ervin, S. M. Imitation and structural change in children's language. In New directions in the study of language. Eric Lenneberg (Ed.), Cambridge, Mass.: MIT Press, 1964.

Fillenbaum, S. and Jones, L. V. Grammatical contingencies in word association. J. verb. Learn. verb. Behav., 1965, 4(3), 248-255.

Fodor, J. A. Could meaning be an r_m? J. verb. Learn. verb. Behav., 1965, 4, 73-81.

Fodor, J. A. More about mediators: a reply to Berlyne and Osgood. J. verb. Learn. verb. Behav., 1966, 5, 412-415.

Geer, J. H. and Mollenauer, S. O. Meaning class and affective content in word association. Psychol. Rep., 1964, 15(3), 900.

Goodenough, W. Componential analysis and the study of meaning. Language, 1956, 32, 195-216.

Greenberg, J. H. Universals of language. Cambridge, Mass.: MIT Press, 1963.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart, and Winston, 1963.

Jakobson, R., Fant, C. G. M. and Halle, M. Preliminaries to speech analysis. Tech. Rep. No. 13. Acoustics Lab. MIT. January, 1952.

Jenkins, J. J. and Palermo, D. S. Coordinate responses to the stimulus words of the Palermo-Jenkins word association list. Penn. State Univ., Research Bull. No. 54, 1964.

Jenkins, J. J. and Russell, W. A. Association clustering during recall. J. abnorm. soc. Psychol., 1952, 47, 818-821.

Jenkins, J. J., Russell, W. A. and Suci, G. J. An atlas of semantic profiles for 360 words. Amer. J. Psychol., 1958, 71, 688-699.

Katz, J. J. The philosophy of language. New York: Harper & Row, 1966.

Katz, J. J. and Fodor, J. A. The structure of a semantic theory. Language, 1963, 39, 170-210. Reprinted in J. A. Fodor and J. J. Katz, (Eds.), The structure of language: readings in the philosophy of language. N.J.: Prentice-Hall, Inc., 1964. p. 479-518.

Katz, J. J. and Postal, P. M. An integrated theory of linguistic descriptions. Cambridge, Mass.: MIT Press, 1964.

Kjeldergaard, P. M. Commonality scores under instructions to give opposites. Psychol. Rep., 1962, 11, 219-220.

Lounsbury, F. G. A semantic analysis of Pawnee kinship usage. Language, 1956, 32, 158-194.

Marks, L. E. and Miller, G. A. The role of semantic and syntactic constraints in the memorization of English sentences. J. verb. Learn. verb. Behav., 1964, 3, 1-5.

Marshall, G. R. and Cofer, C. N. Associative indices as measures of word relatedness: A summary and comparison of ten methods. J. verb. Learn. verb. Behav., 1963, 1, 408-421.

Melton, A. W. Comments on Professor Postman's paper. In Verbal learning and verbal behavior, C. N. Cofer, (Ed.), New York: McGraw-Hill, 1961.

McNeill, D. Development of the semantic system. Dittoed manuscript, Center for Cognitive Studies, Cambridge, Mass., 1964. To appear as part of a chapter by G. A. Miller and David McNeill in the Handbook of Social Psychology.

McNeill, D. Is child language semantically consistent? Dittoed manuscript, Center for Cognitive Studies, Cambridge, Mass., 1965.

McNeill, D. A study of word association. J. verb. Learn. verb. Behav., 1966, 5, 548-557.

Miller, G. A. Language and communication. New York: McGraw-Hill, 1951.

Miller, G. A. and Selfridge, J. A. Verbal context and the recall of meaningful material. Am. J. Psychol., 1953, 63, 176-185.

Moran, L. J., Mefferd, R. B., Jr. and Kimble, J. P., Jr. Idiodynamic sets in word association. Psychol. Monogr., 1964, 78, No. 579.

Noble, C. E. and Parker, G. V. C. The Montana scale of meaningfulness (m). Psychol. Rep., 1960, 7, 325-331.

Osgood, C. E. Meaning cannot be r_m? J. verb. Learn. verb. Behav., 1966, 5, 402-407.

Osgood, C. E. Speculation on the structure of interpersonal intentions. Tech. Rep. No. 39. Institute of Communications Research, Urbana, Ill., 1966.

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. The measurement of meaning. Urbana, Ill.: University of Illinois Press, 1957.

Perfetti, C. A. A study of denotative similarity with restricted word associations. J. verb. Learn. verb. Behav., in press.

Pollio, H. R. Word association as a function of conditioned meaning. J. exp. Psychol., 1963, 66(5), 454-460.

Pollio, H. R. Composition of associative clusters. J. exp. Psychol., 1964, 67(3), 199-208.

Riegel, K. F. The Michigan restricted association norms. Report No. 3, NIMH Grant No. MH 07619-01A1. University of Michigan, Department of Psychology, 1965.

Riegel, K. F., Riegel, R. M., Quarterman, C. J. and Smith, A. E. Developmental differences in word meaning and semantic structure. (to be published in Human Development, 1967).

Romney, A. K. and D'Andrade, G. G. Cognitive aspects of English kin terms. Amer. Anthro. (spec. pub.), 1964, 66, no. 3, 146-170.

Saporta, S. Linguistic structure as a factor and as a measure in word association. In Associative processes in verbal behavior; a report of the Minnesota Conference, J. J. Jenkins (Ed.), Minneapolis: University of Minnesota, Department of Psychology, 1959.

Siegel, S. Nonparametric statistics. New York: McGraw-Hill, 1956.

Siipola, E., Walker, W. N. and Kolb, D. Task attitudes in word association, projective and nonprojective. J. Person., 1955, 23, 441-459.

Thorndike, E. L. and Lorge, T. The teacher's word book of 30,000 words. New York: Columbia, 1944.

Underwood, B. J. and Richardson, J. Some verbal materials for the study of concept formation. Psychol. Bull., 1956, 53, 84-95.

Underwood, B. J. and Schulz, R. W. Meaningfulness and verbal learning. Chicago: J. B. Lippincott Co., 1960.

Wallace, A. F. C. and Atkins, J. The meaning of kinship terms. Amer. Anthro., 1960, 62, 58-79.

Wallach, M. A. On psychological similarity. Psychol. Rev., 1958, 65, 103-116.

Weinreich, U. On the semantic structure of language. In Universals of language, J. H. Greenberg, (Ed.), Cambridge, Mass.: MIT Press, 1963.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Woodrow, H. and Lowell, F. Children's association frequency tables. Psychol. Monogr., 1916, 22, No. 5 (whole No. 97).

Wynne, R. D., Gerjuoy, H. and Schiffman, H. Association test antonym-response set. J. verb. Learn. verb. Behav., 1965, 4(5), 354-359.

UNIVERSITY OF MICHIGAN
3 9015 03125 9883