



City Research Online

City St George's, University of London

Citation: Coope, E. C. (1974). The characteristics and functions of semantic features in human memory. (Unpublished Doctoral thesis, The City University)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/38188/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

THE CHARACTERISTICS AND FUNCTIONS OF SEMANTIC FEATURES
IN HUMAN MEMORY

Emma Clare Coope

THE CITY UNIVERSITY LIBRARY
ST. JOHN STREET, LONDON, E.C.1.

Thesis submitted for the degree of Doctor of Philosophy

Department of Social Science and Humanities
The City University, St. John Street,
London, EC1 4PB

March 1974

ABSTRACT

The work reported in this thesis was concerned with the specification of the organisation of semantic information for which no a priori hierarchy was apparent. An overall objective was to test the validity of a proposed ordering for a set of semantic distinctions. If confirmation were achieved, then the organisation of other parts of the internal lexicon might be similarly deduced.

The first of a set of four studies tested the hypothesis that three semantic distinctions might be organised according to the age at which they were first acquired. The hypothesis was only partially supported by the results. One confounding variable was found to be the operation of specific word effects. Individual differences in the salience of the distinctions was also a possible confounding factor.

The second study tested the use of Repertory Grid Technique as a means of obtaining subject-generated semantic distinctions of assured salience. Examination of the completed grids and of cluster analysis of the data suggested that the technique's novel use in a non-clinical setting was viable.

In the third set of studies, the notion that the grids provided information about basic semantic memory structure was tested in two verbal learning experiments and received strong support from the first of these.

The final study compared the salience of self-generated semantic distinctions (obtained through grid completion) with that of the three distinctions used in the first study. Self-generated distinctions were found to be of greater salience in determining episodic memory performance

which indicated that individual differences in salience might have confounded the validation of the proposed ordering of the distinctions.

The present studies provided some support for the proposed ordering for the set of distinctions, but it was concluded that it did not adequately account for the complexities of adult semantic memory structure. Some proposals for future research in which the structure could be elucidated further were considered.

CONTENTS

	<u>Page</u>
Abstract	2
Acknowledgements	4
List of Tables	8
List of Figures	10
List of Tables in the Appendix	11
 <u>Chapter 1: Introduction</u>	 12
1.1 General Introduction	13
1.1.a The theoretical approach	13
1.1.b Introduction to the remainder of Chapter 1	17
1.2 Semantic Memory: Structural Aspects	18
1.2.a The basic elements of meaning	18
1.2.b Supportive evidence for the componential analysis approach to semantic memory structure	 23
1.2.c Techniques for investigating semantic memory structure	 26
1.2.d Individual differences in semantic memory structure	34
1.2.e Imagery and the study of semantic memory structure	35
1.2.f Summary of approach to semantic memory structure	37
1.3 The Organisation of Non-Hierarchical Semantic Distinctions in Memory	38
1.3.a Introduction	38
1.3.b The theoretical background	39
1.3.c Summary of approach	44

1.4 Semantic Memory: Functional Aspects	45
1.4.a Introduction	45
1.4.b Episodic versus semantic memory	46
1.4.c Semantic memory in verbal learning tasks	48
1.4.d Individual differences in semantic memory function	56
1.4.e Summary of approach to semantic memory function	58
1.5 Introduction to the Experimental Chapters	60
 <u>Chapter 2: Semantic Memory Organisation and the Age-of-</u>	
<u>Acquisition Index</u>	62
2.1 Introduction	63
2.2 <u>Experiment 1.</u> The assignation of three semantic distinctions to low frequency words	64
2.3 <u>Experiment 2.</u> Differences in retrieval time for three semantic distinctions in a sorting experiment	72
2.4 <u>Experiment 3.</u> Differences in the salience of three semantic distinctions for individual words	83
2.5 General Discussion	93
 <u>Chapter 3: Repertory Grid Technique: Individual Differences</u>	
<u>in Semantic Memory Structure</u>	98
3.1 Introduction	99
3.2 Personal Construct Theory	100
3.3 Repertory Grid Technique	102
3.4 The Elicitation of Semantic Distinctions using Repertory Grid Technique	106
3.5 <u>Experiment 4.</u> The generation of constructs for a set of common concrete nouns	113
3.6 General Discussion	120

<u>Chapter 4: Semantic Memory in Verbal Learning Tasks</u>	122
4.1 Introduction	123
4.2 <u>Experiment 5</u> . Paired-associate learning of semantically similar versus unrelated word pairs	125
4.3 <u>Experiment 6</u> . The free recall learning of semantically similar and dissimilar sets of words	130
4.4 General Discussion	139
 <u>Chapter 5: The Importance of the Subjective Element in the Salience of Semantic Distinctions</u>	 143
5.1 Introduction	144
5.2 <u>Experiment 7</u> . The amount of recall as a function of processing words in terms of phonological, general semantic or self-generated semantic distinctions	147
5.3 General Discussion	161
 <u>Chapter 6: Concluding Discussion</u>	 164
6.1 Introduction	165
6.2 The Main Relevant Findings and Interpretations	165
6.3 General Discussion of the Organisation of Non-Hierarchical Semantic Distinctions	172
6.4 Lines of Further Development	176
6.5 Summary of Principal Findings and their Implications	181
 References	 184
Appendix	192

LIST OF TABLES

	<u>Page</u>
<u>Experiment 1</u>	
Table 2.1 The 40 low frequency words broken down into categories according to the three distinctions	65
Table 2.2 The percentage of words for which the semantic distinctions were assigned correctly expressed in terms of whether the words received an A, B or C rating	68
Table 2.3 The percentage of words for which the semantic distinctions were assigned correctly expressed in terms of whether the definitions given for the words were Correct, $\frac{1}{2}$ Correct or Wrong.	68
<u>Experiment 2</u>	
Table 2.4 The 16 concrete nouns categorised according to the three semantic distinctions	73
Table 2.5 The mean sorting time in seconds for the three distinctions	76
Table 2.6 The number of sorting errors for each distinction	78
<u>Experiment 3</u>	
Table 2.7 The eight stimulus words differentiated according to the three semantic distinctions	86
Table 2.8 The mean rankings of the three distinctions for the eight words	89
Table 2.9 Kendall coefficients of concordance, W, showing the degree of agreement between the 43 subjects on the salience rankings of the three distinctions for the eight words	90

Table 2.10 The summed rankings of the salience of the three distinctions for the eight words	91
---	----

Experiment 4

Table 3.1 A section of the dendrogram for subject X showing the clustering for 17 of the 30 elements	109
Table 3.2 A section of the dendrogram for subject X showing the clustering of 10 of X's 24 constructs	110
Table 3.3 The complete dendrogram showing the clustering produced by analysis of the element x element similarity matrix summed across subjects	111
Table 3.4 A section of a typical dendrogram showing the clustering for 20 of the subject's 40 elements	118

Experiment 7

Table 5.1 The 24 words used as elements, displayed in terms of the three distinctions used by Group G	150
Table 5.2 The bipolar distinctions used by the three groups	152
Table 5.3 The mean recall scores for Groups P, G and S	153
Table 5.4 A content analysis of the three constructs chosen for each subject	154

LIST OF FIGURES

PageExperiment 5

- Figure 4.1 Paired-associate learning curves for the High and
Low stimulus lists 128

Experiment 6

- Figure 4.2 The free recall learning curves for the High and
Low stimulus sets 134
- Figure 4.3 The increase in ITR scores over trials for the
two conditions 136

Experiment 7

- Figure 5.1 The percentage of subjects in Group G whose self-
generated distinctions were similar to the general
distinctions C/NC, S/L and N/M, as judged by
a committee of three judges 157
- Figure 5.2 The percentage of subjects in Group S whose
self-generated distinctions were similar to the
general distinctions C/NC, S/L and N/M, as judged
by a committee of three judges 157

LIST OF TABLES IN APPENDIX

Experiment 2

Table A1	The analysis of variance on the individual mean sorting times for the three semantic distinctions	193
----------	---	-----

Experiment 5

Table A2	The mean recall scores for the High and Low lists of word pairs	194
Table A3	The analysis of variance on the recall scores for the High and Low lists of word pairs	195

Experiment 6

Table A4	The mean number of words recalled correctly for the Low and High conditions	196
Table A5	The analysis of variance on the recall scores for the Low and High conditions	197
Table A6	The mean ITR scores for the Low and High conditions, expressed as a percentage of the maximum possible score	198
Table A7	The analysis of variance on the ITR scores for the Low and High conditions	199

Experiment 7

Table A8	The analysis of variance on the recall scores of Groups P, G and S	200
----------	---	-----

Chapter 1

INTRODUCTION

Chapter 1

Introduction

1.1 General Introduction

1.1.a The theoretical approach

We do not record our experience of the world directly, but select, categorise, label and elaborate it before we store it in memory. Language plays a double role in this process. It is used to characterise our experience, the linguistically coded version being stored together with any iconic representation of the raw experience itself. In addition to this, the symbols of language are themselves stored in memory, in a form easily retrievable for use in producing and interpreting speech.

The term 'semantic memory' is used to refer to the memory necessary for the use of language. Tulving (1972, p. 386) defines it as: '. . . a mental thesaurus, organised knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts and relations.'

The representation of a word in the internal lexicon includes a variety of information: its sound and structure, its syntactic properties, its meaning and its emotional and factual connotations. The list is long and the vocabulary of an educated adult language user comprises many thousands of words. Consideration of the efficiency with which information can be retrieved from this representational system, such as the speed with which a spoken word can activate the appropriate meaning in a listener's mind, indicates

that it must be highly organised.

The elucidation of the organisation of this system is a major concern of current semantic memory research and is the central objective of the thesis.

The investigation of how the meanings of words are organised in memory raises an initial question regarding assumptions as to the nature of the basic units in the system. The word might be the basic unit of meaning, and the system consists of a network of inter-word connections, with a word's specific meaning being determined by its position in the associative network. Alternatively, the basic units in the system might be more elemental conceptual components than words. The meaning of a word would then be seen not as an unanalysable whole, but as being built up from a number of conceptual meaning elements. It is a central assumption of the thesis that the meaning of words is built up from more basic units of meaning. An account of the organisation of semantic memory is thus seen as involving (i) specification of what the basic units of meaning are and (ii) specification of how they are structured. The thesis focuses on the second of these aspects.

It is the objective of the thesis to test a proposed system of organisation for a small part of the internal lexicon. This organisation was proposed by Johnson-Laird and Shapiro (unpub.) for a set of three semantic distinctions. These semantic distinctions were viewed as representing three elemental meaning contrasts. They were the contrast between whether or not a substance could be eaten or drunk, which was termed 'Consumable as opposed to Nonconsumable' (C/NC); the contrast between whether or not something was solid or liquid termed 'Solid as opposed to Liquid' (S/L) and the contrast between

whether or not a substance occurred naturally or required the intervention of man in its production, which was termed 'Natural as opposed to Manmade' (N/M). By reference to certain theoretical and empirical considerations Johnson-Laird and Shapiro derived a proposal for how these three semantic distinctions might be organised in the internal lexicon. The particular ordering proposed and the theoretical background to it, together with the specific research proposals to which it leads are presented in Section 1.3. The concept of a semantic distinction and its place in theories of semantic memory structure is discussed in Section 1.2.a. A major dimension on which the proposed ordering of the three distinctions was based was that of their supposed importance, salience or utility.

Experiments 1, 2 and 3, reported in Chapter 2 test the proposed ordering for the three semantic distinctions. Their results provide some support for the proposed ordering. However, the overall pattern of results was not completely convincing, and it was suggested that individual differences in the salience of the three distinctions were confounding the emergence of the proposed ordering.

In order to test this possibility a method of obtaining semantic distinctions of assured salience for individual subjects was sought, and Repertory Grid Technique was used for this purpose. The self-generated distinctions produced by this technique could then be compared in salience with the three distinctions for which the ordering was proposed. Experiment 4 in Chapter 3 reports the preliminary test of the use of the technique. Inspection of the data produced suggested that meaningful semantic distinctions were being generated and so empirical evidence of their representing basic elements of the semantic memory system was sought.

As a test of the validity of the grid data its predictive power for the outcome of two verbal learning tasks was examined. The rationale for this method of testing the validity of the grid data is given in Section 1.4. The results of the first verbal learning study lent strong support to the contention that the semantic distinctions generated by grid technique represented basic elements of semantic memory structure. The findings of both verbal learning studies, Experiments 5 and 6, are reported in Chapter 4.

In Experiment 7 the salience of three self-generated distinctions produced by grid completion was compared with the salience of the three distinctions for which the ordering had been proposed. An empirical criterion of salience was derived by reference to the 'levels of processing' framework for memory research. The logic of this derivation is discussed in Section 1.4.d. According to the criterion used, the self-generated distinctions were found to be the more salient. However an explanation of their superior salience in terms of their generally more salient phrasing, rather than their being intrinsically more salient for the individual subjects who generated them, was also possible. An experiment in which the possibility of the superior salience of the distinctions arising from their phrasing could be tested is proposed in Section 6.4 of the Concluding Discussion.

A secondary objective of the thesis was to develop the use of Repertory Grid Technique as a general tool for semantic memory research. As its novel use in the field of semantic memory research was being investigated for a specific purpose, i.e. the generation of individually salient semantic distinctions for comparison with the three distinctions for which the ordering had been proposed, it seemed

of heuristic value to also consider its more general potential for use in the field. Its possible uses as a means of investigating semantic memory structure are discussed in Section 1.2.c, in combination with an account of two other techniques, which are already in use in this field.

1.1.b Introduction to the remainder of Chapter 1

The remainder of the chapter comprises four sections, of which the last, Section 1.5, introduces the experimental chapters which follow.

Of the other sections, Section 1.2 considers the structural aspects of semantic memory. The use of the concept of features as the basic structural elements is discussed, some techniques for isolating semantic features are presented, individual differences in semantic memory structure are considered and the importance of imagery in the study of semantic memory structure is outlined.

Section 1.3 considers the organisation of non-hierarchical semantic information. This is the specific focus of the thesis in its concern with semantic memory structure. The proposed ordering for the three semantic distinctions C/NC, S/L and N/M is presented and the theoretical and empirical bases for this proposal are discussed. Reference is made to Chapters in which the particular experiments carried out to test the proposed ordering are reported.

Section 1.4 considers the functional aspects of semantic memory. In order to test whether individual differences in the salience of the semantic distinctions for which the ordering was proposed were confounding the emergence of this ordering, Repertory Grid Technique was introduced as a means of obtaining individual subjects' semantic distinctions. The use of the technique involved

testing whether or not the semantic information generated by the technique does represent basic aspects of semantic memory organisation. To do this two verbal learning studies are used, the logic for this deriving from the general assumption that semantic memory is involved to a considerable extent in nearly all such tasks. The functioning of semantic memory in these tasks is therefore discussed. Individual differences in semantic memory function are also discussed. This arises in connection with the testing of the relative salience of the subject-generated and the C/NC, S/L and N/M distinctions, which involves an assumption about the formation of memory traces of words. This particular assumption about semantic memory function is considered in terms of the 'levels of processing' approach to memory research (Craik, 1972; Craik & Lockhart, 1972). The basic tenet of this approach is that the strength or persistence of the memory trace is a function of the 'depth' to which the stimulus is processed or analysed. Craik and Lockhart define greater salience in semantic processing as greater depth of processing. It is assumed that this results in superior memory performance in episodic memory tasks.

1.2 Semantic Memory: Structural Aspects

1.2.a The basic elements of meaning

The body of experimental results has failed to establish conclusively whether words or some other elements should be taken as primitive. The conception of words as the basic structural units is part of the approach to memory structure in which the structure is viewed as a semantic network of associations between words, a word's meaning being defined by its position in the network. The approach used in the thesis is that word meaning is built up from more elemental

components. The distinction between these two approaches will be discussed in the present section.

The proposition that meaning can be analysed into conceptual elements or atoms derives from the componential analysis approach to semantic organisation of linguists such as Katz and Fodor (1963), Weinreich (1966) and Bierwisch (1969). This approach defines the structure of the vocabulary in terms of a relatively small set of very general elements of meaning and their various possible combinations. The terms 'components', 'markers' and 'features' are used by different theorists to refer to these basic conceptual elements.

For the purposes of the present discussion the term 'semantic feature' will be used to refer to a conceptual element of meaning. The existence of such features can only be inferred but examples of possible features have been proposed. Thus the set of semantic features for the word 'woman' might include such concepts as 'human', 'adult', and 'female'. In the experimental chapters the term 'semantic distinction' will be used. This differs from the term 'semantic feature' only in that it is a bipolar expression of a conceptual element. Examples of semantic distinctions are 'human-nonhuman', 'male-female' and 'concrete-abstract'.

Although the thesis focuses on the organisation of the basic conceptual elements of semantic memory i.e. semantic features or distinctions, this is not to deny that words are represented in memory along dimensions other than the conceptual one.

The essence of the attribute theory of word storage is the assumption that a word may be conceived as represented in memory as a collection of components or 'attributes', which uniquely characterise

that word and distinguish it from all others. This point of view has been reflected in several recent theoretical papers on memory structure (Bower 1967; Morton, 1970; Norman, 1969; Underwood, 1969; Wickens, 1970). The distinction between semantic features and attributes will now be considered.

Examples of the broad classes of attributes in terms of which words are interrelated and distinguished are provided by Voss (1972). He proposes four structural dimensions along which the attributes of words are stored. He stresses that the criteria employed for distinguishing between the dimensions are purely empirical and that the listing of dimensions is not intended to be exhaustive. The dimensions are formal, associative, semantic and syntactic. The formal dimension refers to the organisation of words in terms of their formal acoustic and visual similarities. The associative dimension refers to a variety of bases of associative relation between words, indeed the random occurrence of any two words under a given set of circumstances may result in their becoming associated. The semantic dimension refers to the relations of category membership and other conceptual relations that are implied by the meaning of a single word. These relations are termed semantic features in the thesis. The syntactic dimension refers to the organisation of words in terms of their grammatical characteristics.

These examples of the dimensions in terms of which words may be organised should serve to enable a clear distinction to be made between a semantic feature and what are termed attributes by the attribute theory of word storage. The four dimensions do not represent single attributes, for in fact within each there are a large

number of different attributes by which different words may be distinguished. What is intended to be made clear is that the term semantic feature refers only to the organisation of words in terms of their semantic characteristics, while the term 'attribute' may refer to a variety of different modes of organisation, such as the four dimensions described above. If encoding may be defined as the process by which certain aspects of the stimulus are employed in establishing its internal representation, then attributes may be defined as those aspects of the stimulus which are encoded. The basic and elemental semantic attributes which are available for encoding are the semantic features.

Having distinguished semantic features from attributes, the characteristics of semantic features will now be considered in greater detail. According to the componential semanticists, semantic features are not part of the vocabulary itself, but rather are conceptual elements representing the basic dispositions of cognitive and perceptual structure. They are therefore common to all natural languages. The meaning of a word is represented not by its relations to other words, but by its particular set of semantic features. A word is semantically ambiguous if it has more than one complex of semantic features assigned to it. Semantic similarities between words are expressed in terms of shared features. Sets of lexical entries whose meanings have certain features in common form a 'semantic field' or 'domain'. Such a domain is formed by kinship terms whose elements share a feature configuration 'animate', 'human', 'relative'. The verbs of motion are another example.

In Katzian theory (e.g. Katz, 1966) (in which the conceptual

elements of meaning are termed 'markers') the 'names' of the markers are heuristically useful but the theory would not be changed if they were all replaced by numbers (under appropriate identity and distinctness conditions). Katzian accounts of meaning relationships tally with empirical observations only to the extent that the theory predicts judgements of certain types (e.g. paraphrase and ambiguity) which are themselves formally defined only by the theory. Marshall (1970) remarks that a fully adequate semantic theory should do considerably more than this. Perhaps, as Bierwisch (1967) has said, 'true candidates for the universal set of semantic markers' should be candidates that 'may be interpreted in terms of the basic dimensions of the human apperceptive system'. Unfortunately it has proved very difficult for psychologists to provide formal perceptual and cognitive interpretations for the 'names' of the semantic features that occur in current semantic theories.

In the face of such difficulties, one possible solution in conceptualising semantic memory structure is the replacement of the features by the actual words conventionally used to denote such concepts. Anglin (1970), for example, has assumed no qualitative difference in the internal representation of words and features.

Although much of the empirical data on semantic memory structure is amenable to interpretation in terms of words or semantic features as the basic structural units, there are certain areas in which the concept of semantic features has greater explanatory potential. An account of some of this research will now be presented, as supportive evidence for the use of the concept of semantic features in the thesis.

1.2.b Supportive evidence for the componential analysis approach to semantic memory structure

One model of semantic memory structure in which the basic units of meaning are words rather than semantic features is that of Collins and Quillian (1972). This work stems from a computer model and uses reaction times as the main dependent variable. The basic memory structure is viewed as a hierarchical semantic network. Words are organised in a network of different sorts of associations, including class-inclusion and part-whole relations. The model comprehends a sentence by finding a grammatically permissible path through the semantic network in a way that relates the words in the sentence. Thus 'a canary has a skin' is comprehended by finding the path: a canary is a bird, a bird is an animal, and an animal has skin. In applying the model to human memory it is assumed that it takes time to move from one level of the hierarchy to another, and since the steps are taken successively, it is predicted that comprehension time will increase linearly with number of steps.

The organisation of the network is assumed to reflect the way in which the information is first acquired. For example, that canaries have skin is rarely learned as a separate fact; rather one would learn that a canary is a bird, a bird is an animal and an animal has skin. Hence there would be no direct link between 'canary' and 'skin' in the semantic network. It should therefore take longer to retrieve the relation between 'canary' and 'skin' than between 'canary' and 'bird', which are directly linked.

In order to test this model, subjects performed a categorisation task in which they were shown sentences expressing various types of

category membership and were required to judge them true or false, and press an appropriate button as quickly as possible. As predicted by the model, decision times for positive judgements increased linearly with number of steps in the hierarchy that must be traversed. But although the closeness of the items in the network facilitated positive judgements about them, it appeared to impede negative judgements. The longer decision times in evaluating 'a canary is an ostrich' rather than 'a canary is a fish' suggested that the former decision was more difficult. To retain the network model it is necessary to assume that negative judgements are impeded by the now irrelevant links between neighbouring words.

In summary, the representation of a word in semantic memory, according to this theory, would consist of a hierarchical network of associations in which the word's meaning is expressed by its connections to other words, maximum similarity between words being expressed by a direct link between them.

The model of Collins and Quillian has difficulty explaining decision times for 'false' sentences, i.e. in explaining the fact that the closeness of the items in the network appears to impede negative judgements about them. An alternative explanation for this result can be provided by resorting to a semantic feature approach to semantic memory organisation.

Schaeffer and Wallace (1969, 1970) proposed that sets of semantic features were involved in categorisation tasks. Decisions about whether items were members of categories were seen as involving the retrieval and comparison of the items' respective sets of features. The comparison involved sampling a certain amount of information from both sets of features in order to satisfy a criterion. Similar

words by definition have more features in common than dissimilar words, so the sample for a positive judgement will be relatively small for similar words, while the sample for a negative judgement will be relatively small for dissimilar words. Thus the semantic feature approach offers a unified account of the differences within a category and of the differences between hierarchies of categories; it seems better able to handle the results of these reaction time studies than does the semantic network point of view.

In addition to the support given to the semantic feature approach by studies of the categorisation of nouns, further support is provided by analyses of the semantic classification of verbs.

An unpublished analysis by George Miller and Johnson-Laird shows that, for verbs, simple hierarchies give way to intricate cross-classifications. For example, verbs such as 'travel' 'own' and 'see' lack a causal component that is present in the meaning of 'lift', 'seize' and 'look at', and these verbs in turn lack an intentional component that is present in 'chase', 'buy' and 'look for'. Since the semantic network theory has no way of freeing meaning from the specific verbiage in which it is expressed, it seems committed to the view that meanings of verbs are acquired, in part, from explicit statements involving 'cause' and 'intention'. As Johnson-Laird (1974, p. 8) points out: 'Only a theory countenancing underlying conceptual elements seems to offer sufficient flexibility to account for the acquisition of meaning.'

In summary, it must be acknowledged that it is difficult to distinguish between a semantic network in which the basic elements are words and a semantic system composed of sets of more elemental semantic features. Empirical data are often compatible with either sort of explanation.

However, when the organisation of verbs rather than nouns is considered, the need for certain conceptual features, such as 'cause' and 'intention' seems more apparent. Unless the existence of such meaning dimensions, independent of explicit verbiage, is accepted, we are forced to the unlikely conclusion that children learn the meanings of verbs from explicit statements involving 'cause', 'intention' etc. The evidence in support of the conception of semantic memory structure in terms of semantic features is certainly not conclusive. But, in the absence of strong evidence to the contrary, the use of the concept of semantic features as 'conceptual elements of meaning' will be maintained in the present thesis.

Although the concept of semantic features is of heuristic value, it remains true that specification of the nature of the set of semantic features which is assumed to form the basic semantic structure of all natural languages is far from complete. Section 1.2.c discusses three attempts which have been made to isolate these basic components of meaning.

1.2.c Techniques for investigating semantic memory structure

The lack of systematic knowledge about the meanings of words, such as what the basic semantic features are, has severely limited the scope of research in semantic memory. As Marshall (1970, p. 193) states:

'Any serious attempt to show how semantic readings are assigned in perceptual tasks or encoded in production tasks pre-supposes the existence of 'correct' dictionary entries for a fairly extensive part of the vocabulary. The very small number of precise lexical analyses that linguists have proposed will not

support large scale experimental work.'

In brief, it is difficult to study the organisation and retrieval of information without making strong assumptions as to what information is involved.

A number of empirical techniques have been used in an attempt to reveal the components of semantic memory structure. Two techniques used by Miller will now be considered in greater detail, and the novel use of Repertory Grid Technique for semantic memory research will be discussed.

Miller's Sorting Experiments

In a series of papers Miller (1967, 1969) has described a sorting procedure for the study of semantic memory structure. Typically subjects are given a deck of cards with a word and its definition printed on each card. The subject is required to sort the words into piles on the basis of similarity of meaning. He is allowed any number of piles with any number of words per pile. For each subject an incidence matrix is constructed showing for every pair of words whether or not they were put in the same pile. The sorting data from individual judges is then pooled and subjected to hierarchical cluster analysis (Johnson, 1967).

The basic assumption in such experiments was that if all items were conceptually distinct for the judges, then in grouping them, they must necessarily ignore some of the distinguishing features. Thus an analysis of the sortings was hoped to reveal 'which conceptual features have been ignored and thus, by indirection, what the features are' (Miller, 1969).

In an experiment involving 48 common nouns, features which could be characterised as object-nonobject, living-nonliving and

human-nonhuman appeared to emerge from an intuitive analysis of the data (Miller 1967). Major features are inferred from an inspection of the tree diagrams produced by cluster analysis, the proximity of items indicating the degree of their similarity in meaning. Although this labelling of the major emergent categories is purely inferential, the advantage of the procedure is that the actual hierarchical ordering of the features in terms of generality and clarity of usage comes out directly, in terms of the numbers of native speakers agreeing on co-assigning items.

Miller (1972) reports that the reflection in the word clusters of the most common bases of semantic classification used by judges is generally in accord with intuitive sense. That is, the sorting method seems to reveal an organisation that adequately reflects our subjective impressions of semantic similarity. The results of a similar work on a set of verbs were difficult to interpret, and appeared to be very sensitive to context, leading Miller to the suggestion that the semantic analysis of grammatical predicates might be basically different from that of grammatical subjects. As Miller (1967, p.23) remarks: '... neither the object/concept distinctions appeared, nor did anything else that was recognisable.' He emphasised, however, that the method is not a discovery procedure; it is a method of verification and falsification. Given a theory that stipulates how words should be related, or given rival theories of semantic organisation, the method makes possible the reaching of some empirical judgement as to which theory best fits the empirical facts.

In order to fit the sorting data, therefore, some fairly explicit theory about the semantic components that are common to a particular

set of words is necessary. It was for this purpose that Miller evolved a second technique of relevance here, the technique of incomplete definitions.

The Method of Incomplete Definitions

Realising the immensity of the task of systematising the entire lexicon, Miller (1972) chose a group of 217 related words that seemed to illustrate several conceptual dimensions and explored the hypothesis that these dimensions comprised an organic system of concepts and conceptual relations. The category of motion verbs was chosen for the case study, and the attempt made by the method of incomplete definition, to find a set of semantic components inherent in their meaning whose psychological validity as concepts could be tested by experimentation and which might eventually be shown to include some of the primitive concepts used to encode experience generally.

An incomplete definition is a substitutable phrase that has a more general meaning than the word it replaces. Words can have similar meanings without being synonyms; the method of incomplete definitions is used to characterise that shared part of their meaning on which our judgement of similarity depends. Two words that imply the same incomplete definition can be said to be similar in meaning to that extent; they may actually be synonyms, but in most cases they will not. For example the noun dog can be incompletely defined as 'a warm-blooded animal'. An elephant is also a warm-blooded animal, so to that extent the words share semantic components and are similar in meaning. An incomplete definition is thus a part of all definitions of a class of words, and so serves to explicate the important but rather vague concept of a lexical 'domain'.

Applying this method to verbs of motion, Miller suggests the use of the phrase 'changes location' as an incomplete definition for the entire domain. In order to differentiate further among these verbs it is necessary to add further components to the definition. For example some verbs of motion have a causative component (e.g. 'He lifted it' means he caused it to change location). A directional component may be added to reflect the fact that some verbs incorporate in their meanings the direction in which the location changes (e.g. 'He lifted it' means he caused it to change location in an upward direction).

It is apparent that this componential analysis depends on a purely intuitive choice of paraphrase. The analyst has little more than his own lexical intuitions to help him decide among various alternative classifications.

However, the resulting analysis can be tested empirically by the method of sorting, described above, which gathers and summarises the lexical intuitions of many language users. The method does not provide complete support for classifications suggested by 'incomplete definition', but it does serve as a corrective for the more obvious intuitive biases of the semantic theorist. A componential analysis that agrees closely with subjects' pooled judgements of semantic similarity is a more satisfactory solution than one that does not.

Miller found that the domain of English motion verbs could be mapped satisfactorily with a dozen general components, i.e. semantic features that have applicability in other lexical domains as well. For complete definition, most motion verbs require additional specific components that have no obvious function elsewhere.

By reducing an important semantic domain to a short list of semantic elements, packaged variously in the various words, Miller has established a valuable beachhead into conceptualising the English lexical system. The challenging task of formulating the organisation of components that determines specific meaning should proceed more effectively when the nature of the combining elements is specified.

Repertory Grid Technique

Chapter 3 in the present thesis reports the novel use of Repertory Grid Technique in a non-clinical setting as a means of studying individual semantic memory organisation. Grid technique is more fully discussed in the introduction to that chapter. However, a brief discussion of the grid's potential as a heuristic tool in semantic memory research seems appropriate at this point.

The technique was developed by Kelly (1955), as an adjunct to his theory of personal constructs. The crux of the theory was the tenet that each person evolves for his convenience in understanding the world an interlinked and relatively coherent system of constructs. Each person's construct system is unique. A construct is concept expressing the way in which the individual sees some things as being similar and yet different from others. When considering three acquaintances, for example, a person might formulate a construct expressing the way two of them seem to be kind while the third seems to be cruel. He is thus using the construct Kind-Cruel. Constructs are always bi-polar, in that they express a similarity and a relevant contrast.

Repertory Grid Technique involves the elicitation of constructs for a set of objects known as elements. A grid or matrix is used so

that every element may be given a numerical score for every construct. Any entities or concepts may be used as elements, from names of close friends to names of abstract qualities like fear or jealousy. When names of concepts rather than of actual people are used, constructs are generated by finding a way in which two of the elements are 'similar in meaning' and different from the third.

Consideration of the use of the technique in the field of semantic memory analysis immediately suggests that this is a means of tapping subject-generated semantic distinctions. Subjects can be presented with a set of words as elements, chosen to represent a particular linguistic domain, and asked to supply meaning dimensions for them by thinking of constructs which expressed similarities and relevant differences among the set of words. The position of every word in terms of every construct could be ascertained, using some numerical scoring method, as in semantic differential technique, and thus the words could be grouped according to degree of similarity, in terms of the individual subject's construct system. The technique thus permits a semantic analysis of a set of words, similar to that obtained by Miller's sorting technique, but which has the added advantages of allowing for the expression of finer shades of meaning, and of subjects supplying their own labels (constructs) for the meaning dimensions they used.

The groupings of words provided by such a technique could be used for testing alternative theories of semantic organisation. Repertory grid technique should lead to the same basic features being produced as with other techniques such as Miller's if (1) those features are basic and (2) the grid is an adequate tool for tapping semantic memory. Although the emphasis in the grid's clinical usage

has been on idiographic data collection, between-subject consistencies in the study of linguistic structure should be expected. The use of the grid as a heuristic tool in the field of semantic memory research is further discussed in Section 6.4 of the Concluding Discussion, in the light of the findings of the experiments in which it was used which are reported in the thesis.

Summary of attempts to isolate semantic features

For certain restricted domains, relatively simple classification methods, such as Miller's sorting technique, or repertory grid technique, seem able to produce useful data. In essence, the subjects are required to judge 'similarity of meaning' and the resulting data are submitted to cluster analysis. The information thus revealed about the semantic features in operation makes possible evaluations of the generality of proposals about semantic structure.

However, Miller (1967) has drawn attention to the difficulties in making judgements about verbs, finding that their classification is considerably more complex than that of nouns. In this area, Miller's (1972) analysis of the domain of verbs of motion seems to offer ^apromising method of further analysis.

While these studies reveal some of the important features involved in the subjective lexicon, they do not go very far in characterising the structural organisation of such features. As Fillenbaum and Rapoport point out (1971, p. 249): 'We know very little about the general organisational properties in terms of which restricted lexical domains are tied together or connected in the subjective lexicon.'

Nevertheless, while acknowledging that the psychological problem is to specify the cognitive organisation of these semantic components, it must be recognised that the problem could scarcely be formulated

prior to a full characterisation of the semantic components involved. A rich knowledge of the subjective lexicon is absolutely essential for any semantic theory, because such information constitutes the data with which any such theory must cope. The development of techniques such as those described above, which do yield this information, is therefore a very valuable first step.

1.2.d Individual differences in semantic memory structure

The consideration of individual differences in semantic memory first becomes an issue in the thesis when the possibility is raised that the failure of the experiments reported in Chapter 2 to fully support the proposed ordering of the three semantic distinctions might in part be due to individual differences in the salience of those distinctions. The proposed ordering was based on the assumption that there would be a decrease in the importance or salience of the distinctions from a maximum in the C/NC distinction through the S/L distinction to a minimum in the N/M distinction. Large individual differences were shown in the data for Experiment 2 and so an investigation of any *masking* effects due to individual differences was proposed. If reliable differences in the salience of semantic distinctions within individual's semantic memory structures were found, then this would make the attempt to specify, in terms of salience, the ordering of such distinctions across subjects considerably more complex.

In order to obtain examples of semantic distinctions of assured salience to individual subjects, the use of Repertory Grid Technique was introduced. This technique was drawn from the field of clinical psychology and its novel use in the field of semantic memory research

was not assured of success. An account of the initial use of the technique is given in Chapter 3. It was possible that the constructs generated by subjects by grid technique would not in fact constitute examples of basic conceptual elements of meaning and so their use as 'semantic distinctions' in the same sense as the term was used in Chapter 2 might not prove justified. However, a verbal learning study reported in Chapter 4 provided support for the validity of these constructs as representing basic and fundamental elements of meaning.

Repertory Grid Technique was therefore viewed as of heuristic potential for obtaining information about semantic memory structure. The general development of this potential was beyond the scope of the present thesis and a specific use of the grid technique was proposed.

In Chapter 5, grid data were used specifically to compare the salience of self-generated distinctions with the three distinctions discussed previously (C/NC, S/L and N/M). Degree of salience was indicated by free recall performance in an incidental learning paradigm.

1.2.e Imagery and the study of semantic memory structure

The thesis aims to investigate the organisation of the semantic memory system in terms of its basic units. It adheres to a conceptualisation of these units as conceptual elements of meaning. The association of images with these unitary concepts will now be considered. This attention to the operation of imagery in connection with semantic memory is important because the outcome of the studies of semantic memory structure reported in the thesis may have been affected by the subjects' use of imagery.

It has been suggested by Bower (1971) and Paivio (1971) that semantic memory involves two parallel systems, one based on language,

the other on visual imagery. Similarly Collins and Quillian (1972) point out that the components of semantic memory need not comprise only verbal concepts. They state (p. 315): 'Concepts are built up out of sensori-motor experience as well as language use, and there is every evidence that people utilize imagery extensively.'

Paivio (1971) proposes a theoretical approach to meaning in which mental images and words are regarded as the major psychological reactions to objects and verbal stimuli. Concrete terms such as HOUSE readily evoke both images and words as associative reactions, whereas abstract words such as TRUTH more readily arouse only verbal associations. He proposes that to the extent that imaginal processes are coordinated with abstractness-concreteness, a scaling of words on that dimension would define their capacity to evoke images. Studies by Frincke (1968) and Paivio (1968) in which subjects rated a set of nouns on a number of variables such as familiarity, emotionality, generality-specificity and an imagery index showed that this index was the best predictor of verbal learning scores. It has also been shown that instructions to use visual imagery greatly facilitate learning (Bower, 1971).

The research in this area is not sufficiently developed to allow confident assessments to be made of the levels of imagery associated with the three semantic distinctions for which the proposed ordering was to be tested. Neither were precise measures possible of the differences in imagery among the concrete nouns used as stimuli. However the imagery value of the abstract N/M distinction might be expected to be lower than that of the C/NC and S/L distinctions. Of these, the purely perceptual characteristics of the S/L distinction indicate that its imagery value would be high. If certain of the words used as

stimulus materials evoked strong images this might lead to the S/L distinction being responded to as most salient for those words. Thus the operation of imagery might tend to confound the emergence of the proposed ordering of the distinctions, which was of a decrease in salience from the C/NC distinction through the S/L to the N/M distinction. If imagery was important, the effective ordering might become S/L, C/NC, N/M.

1.2.f Summary of approach to semantic memory structure

The term semantic feature was defined as the basic element of meaning in the semantic memory system and distinguished from an attribute in that it represents only one very specific type of attribute. The evolution of the proposition that meaning can be analysed into conceptual elements was then considered.

The point was made that much of the empirical data on semantic memory structure is amenable to interpretation in terms of words or semantic features as the basic structural units. However a more detailed assessment of the relative interpretive power of the two approaches suggested that that of the semantic feature approach was greater. The value of conceptualising the basic structure of semantic memory in terms of semantic features in the thesis was therefore supported.

One difficulty of the semantic feature approach is the lack of a comprehensive list of semantic features and of any simple procedure for discovering them. Two attempts by Miller (1967, 1969, 1972) to isolate the basic semantic elements were presented and the potential use of repertory grid technique as a third method of isolation was proposed.

Individual differences in semantic memory structure were then considered, as it was possible that the results of the experiments to be reported in Chapter 2 might be interpreted in part by the operation of significant subject effects. A method of assessing the importance of this factor was proposed and Experiment 7 reported in Chapter 5 was carried out for this purpose.

Finally the operation of imagery in semantic memory tasks was discussed as this was another possible confounding factor in the experimental studies reported in Chapter 2. It was concluded that for stimulus words of high imagery value, the S/L rather than the C/NC distinction might be most salient, a result counter to the theoretical predictions, but unfortunately current research on imagery was not sufficiently advanced to offer a simple control for this factor.

1.3 The Organisation of Non-Hierarchical Semantic Distinctions in Memory

1.3.a Introduction

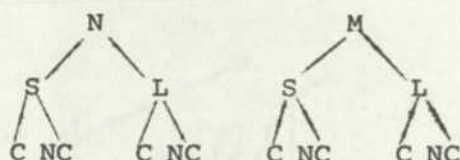
The attempt to specify the organisation of a small part of the internal lexicon, for which no a priori hierarchy was apparent, was a major objective of the thesis.

A more specific account of the theoretical background to the experiments to be reported on the organisation of non-hierarchical semantic information will now be given.

The elemental components studied were expressed as bipolar semantic distinctions such as 'male-female' and 'concrete-abstract'. The semantic contrast between two semantic features was used rather than the single features because this facilitated the selection of stimulus

words which could be analysed into mutually exclusive, but interconnected categories.

Using the semantic distinctions C/NC, S/L and N/M eight categories could thus be formed:



1.3.b The theoretical background

The hierarchical structure of the internal lexicon has been emphasised by both linguists and psychologists. Miller (1969a) has described how the features of words can often be cast into a hierarchical or nest-like relation. Class inclusion relations can be represented by tree-diagrams with the words at branch level and increasingly generic class labels or features at successively higher nodes of branch intersection. If two words share a given feature in a nest, they necessarily share all higher-order features within that nest.

However, the hierarchical relationship is not the only type of relationship that binds words. Miller has made the point that the nature of the structural organisation of semantic features may also be linear or paradigmatic. The distinction between hierarchical and paradigmatic organisation is that in paradigmatic memory-structures every entry has a value on each dimension or feature, while taxonomic hierarchies have no consistent set of features for each entry. As Fillenbaum has said (1971, p. 291): '... if one asks about the nature of the structural organisation of semantic features in a particular case, whether it is linear, hierarchic, paradigmatic or whatever, one

raises some major and essentially unsolved conceptual problems and also measurements problems with regard to the sorts of techniques that might be used to analyse findings and to evaluate the adequacy of various structural models.'

In view of the variety of types of organisation that may exist, it seems limited to theorise about semantic memory on the basis of results of studies that have emphasised hierarchies such as that of Collins and Quillian (1972). Linguists such as Chomsky (1965) have often pointed out that many classificatory dimensions do not yield hierarchies unless we are prepared to impose an arbitrary ordering upon them. The experiments reported in Chapter 2, therefore, focus on some semantic distinctions for which there is no natural a priori ordering, and attempt to elucidate their relationship to each other.

One aspect of the deficiency in our knowledge of semantic memory structure is that in the absence of a postulated hierarchy, statements about the retrieval of semantic information become difficult. In an attempt to elucidate the structural organisation of a small part of the internal lexicon, that concerned with expressing the characteristics of physical objects and substances, Johnson-Laird and Shapiro (unpublished) proposed a possible ordering of three semantic distinctions. The distinctions were that between those things which man does or does not eat or drink, expressed as (Consumable as opposed to Nonconsumable), that between those things which are solid and those which are liquid, expressed as (Solid as opposed to Liquid) and that between those things which occur naturally and those which require the intervention of man in their production, expressed as (Natural as opposed to Manmade). These distinctions are abbreviated to C/NC,

S/L and N/M in the thesis. It should be noted that the *three* distinctions are logically independent and that any physical substance must be assignable to at least one category within each bipolar distinction.

The ordering proposed for these distinctions was based on supposed differences in the age at which they were first acquired. The proposal hinged upon the tenet that survival needs would ensure that the child would acquire first information that was of primary relevance for survival. He would similarly be expected to acquire semantic information of primary relevance to his ability to cope with and adapt to his environment before he acquired information of lesser relevance. It would also follow that information of such primary importance, salience or utility would be readily accessible within the semantic memory structure, while information of secondary importance would be less readily accessible. Thus the age of acquisition of a semantic distinction should correlate with its speed of retrieval from semantic memory. Other variables which might correlate with age of acquisition were rated salience of the distinctions for sets of words and accuracy of assignation to partially known words; semantic distinctions of greatest importance for survival and environmental adaptation would be expected to be of maximum salience and to be most well learnt. On the basis of such considerations the ordering proposed by Johnson-Laird and Shapiro for the three distinctions was C/NC, S/L and N/M, where the C/NC distinction was thought to be that acquired first, and the N/M distinction to be that acquired last. This ordering will be referred to in the thesis as the age-of-acquisition index. The predictions which follow from it are that the C/NC distinction would achieve a maximum in speed

of retrieval, rated salience and accuracy of assignation and that there would be a trend of decreasing speed, salience and accuracy of assignation from it, through the S/L distinction to a minimum for the N/M distinction. Experiments 1, 2 and 3, reported in Chapter 2 are tests of these predictions.

As it has formed the central pivot of the research reported in the thesis, the bases for Johnson-Laird and Shapiro's proposed ordering will now be considered in greater detail. Support for the correlation of age-of-acquisition with utility or salience was provided by Brown's (1958) observation that vocabulary is acquired according to its usefulness. The same principle should govern the acquisition of the underlying semantic distinctions. Accordingly, Johnson-Laird and Shapiro stipulated that the C/NC distinction should be acquired before the other two distinctions because of its overriding importance for survival. Similarly, they proposed that the obvious perceptual correlates of the S/L distinction should ensure its acquisition before the more abstract N/M distinction.

Johnson-Laird and Shapiro suggested that this order was also supported by Bruner and Olver's (1963) observation that children in describing ways in which items are alike or different place emphasis first upon their functional properties and then upon their perceptual properties. A direct reference to Piaget's work on developmental stages (e.g. Piaget, 1955), in which perceptual distinctions are shown to be grasped earlier than abstract distinctions, would seem to provide strong support at least for the ordering of the (Solid-Liquid) and the (Natural-Manmade) distinctions. Also Piagetian theory lends support to the

principle of utility; concepts being learnt in order to enhance the child's ability to understand and manipulate his environment. Similarly Bruner et al. (1968) maintain that the child understands the world first through its actions (enactive representation), then through imagery (iconic representation) and finally through language (symbolic representation). Once again, the ordering of the last two distinctions seems valid, but that of the (Consumable-Nonconsumable) and the (Solid-Liquid) distinctions unclear. Both seem to be open to both enactive and iconic representation.

The derivation of the notion of ordering by utility, from Brown (1958), is also questionable. Brown (p. 20) states: 'The categorisation that is most useful for very young children may change as they grow older.' If this is also the case for the underlying semantic features, then the maximally useful features of a word may change with age. Should this be the case, then predictions of order effects based on age of acquisition in experiments using adult subjects would seem somewhat weak.

However, a recent experiment by Carroll and White (1973) provides confirmation of the predictions of retrieval latency based on age of acquisition. In a picture-naming task, using adult subjects, these researchers found that objects whose names were learned early in life were named faster. Age-of-word variables were shown to contain the significant information necessary to predict latency, the pattern of results not being affected when both word frequency and word length variables were eliminated. This finding suggests that memory trace strength is a function of age-of-acquisition. Certainly, it seems that age of acquisition should be viewed as a more relevant variable in psycholinguistic studies than it has in the past.

In the light of this finding, Johnson-Laird and Shapiro's prediction of order effects due to the age of acquisition of particular semantic distinctions seems more viable.

Experiments 1, 2 and 3 reported in Chapter 2 were therefore carried out to test this prediction. Experiment 1 tested whether the age-of-acquisition index predicted accuracy of assignment of the three distinctions for partially known words. As adult subjects were used, this entailed the assumption that in learning the meanings of words in adulthood, the same order of priority (based on importance for survival and environmental adaptation) of semantic distinction assignment would apply as in childhood. Experiment 2 tested whether the index would predict sorting times for common concrete nouns according to the three distinctions. If retrieval speed were greatest for the C/NC distinction then the latency of responses for this distinction should be at a minimum and would increase through to a maximum for the N/M distinction. Experiment 3 tested whether the index would predict salience ratings for the distinctions for a set of common concrete nouns. The prediction was that there would be a trend of decreasing salience from a maximum for the C/NC distinction through to a minimum for the N/M distinction.

1.3.c Summary of approach to the organisation of non-hierarchical semantic distinctions

The central objective of the thesis is to investigate the organisation of semantic features for which no a priori hierarchy is apparent. Most semantic memory research has focused on hierarchical semantic organisation and it seemed valuable to attempt to elucidate the interrelationship of classificatory dimensions that can not be cast

in hierarchies on an a priori basis.

A proposed ordering for the three semantic distinctions C/NC, S/L and N/M was considered. If empirical support for this ordering were provided, then the reasoning on which the proposal was based could be used to specify the organisation of further parts of the internal lexicon.

The theoretical and empirical bases of the proposal were discussed. The central tenet was that semantic distinctions were acquired according to their importance. Thus the distinctions acquired first in childhood should have priority in the semantic memory structure over those acquired later in life. When applied to the three distinctions in question the ordering C/NC, S/L to N/M in terms of decreasing importance was generated (Johnson-Laird and Shapiro, unpub). However the theoretical support for this ordering of specifically the C/NC and the S/L distinctions was thought to be weak.

Three experiments in which this ordering was tested are reported in Chapter 2 and further consideration is given in the Chapter to the validity of the proposed ordering. The remainder of the thesis focuses on the implications of the results of those experiments which did not wholly support the age-of-acquisition index. In particular, the importance of individual differences in semantic memory organisation is explored.

1.4 Semantic Memory: Functional Aspects

1.4.a Introduction

Chapter 4 in the thesis reports two verbal learning studies which were conducted as tests of the validity of information about semantic memory structure provided by repertory grid technique.

There is a great deal of evidence to suggest that the operation of semantic memory in verbal learning tasks strongly facilitates performance. Therefore prior knowledge of what the relevant semantic memory organisation was for a set of to-be-remembered (TBR) material (provided by grid technique) should allow the experimenter to facilitate or inhibit the subjects' learning performance accordingly. This would only be possible if the information about semantic memory organisation provided by repertory grid technique were valid.

In order to support this use of the verbal learning tasks some of the empirical evidence for the importance of semantic memory in these tasks will be presented. First, however, the distinction proposed by Tulving (1972) between semantic and episodic memory will be outlined, as it is most relevant to the exact role which semantic memory plays in verbal learning tasks.

1.4.b Episodic versus semantic memory

In his definition of semantic memory Kintsch (1972) proposes that it is an organised internal lexicon which can serve as a basis for processing information in a variety of memory tasks, including free recall. Tulving (1972) suggests that although they do not explicitly state it, Collins and Quillian in their work on semantic memory (Collins and Quillian, 1972) would want their 'language user' to be able to remember lists of unrelated words. This inclusion in definitions of semantic memory of specifications for its role in verbal learning tasks raises the question of exactly what is stored in a subject's memory during a typical verbal learning task. What happens when a subject encounters a familiar item e.g. the word CAT, while studying a list of TBR words? It might be supposed that the

item CAT is entered in the subject's memory store when he sees it on the list. But it is not necessary for the subject to 'learn' the word CAT. The item is already stored in the person's semantic memory.

What the subject is learning in the verbal learning situation is a description of an autobiographical event. At the most elemental level the subject might store such a description as 'The word CAT occurred in that context that I will call List-1.' He thus links together two known concepts in semantic memory, CAT and the 'List-1' context.

The recording of such facts by the establishment of new linkages between known concepts in semantic memory is what Tulving (1972) refers to as 'episodic memory'. Episodic memory receives and stores information about temporally dated episodes or events and the temporal-spatial relations among these events. The act of retrieval of information from the episodic memory store changes the contents of the store. The system is thus quite susceptible to transformation and loss of information. While the form in which the perceptual input is registered in episodic memory may be influenced by information in semantic memory, a phenomenon referred to as encoding, it is possible for the episodic system to operate relatively independently of the semantic system.

It should therefore be clear that while all laboratory verbal learning tasks involve episodic memory, they need not involve semantic memory. Semantic memory is only involved when the TBR material is subject to semantic encoding or organisation. A more detailed account of the role of semantic memory in verbal learning tasks is given in the next section, Section 1.4.c.

1.4.c Semantic memory in verbal learning tasks

Chapter 4 in the thesis reports the results of two verbal learning studies which were conducted on the assumption that the subjects could enhance their performance by the use of semantic memory. The two paradigms used were those of free recall learning and the paired-associate learning. In the following discussion of the role of semantic memory in verbal learning tasks, most of the research reported involves the free recall learning paradigm.

In free recall learning the subject is presented with a list of words to learn and is told to recall the items in the order in which he thinks of them. Interestingly, certain regularities appear in the order in which the items are recalled. For example, there is a tendency for items which are somehow related to one another to be recalled together even though these items were not contiguous during presentation (Shuell, 1969). This discrepancy between the order in which the items were presented and the order in which they were recalled is presumed to represent a tendency on the part of the subject to organise his recall on the basis of pre-experimental associations or conceptual relations. The associations and conceptual relations in terms of which the subject organises his recall are drawn from the semantic memory store. The tendency for related items to be recalled together is termed clustering (Bousfield, 1953). The tendency for subjects to recall groups of seemingly unrelated words in the same order consistently, despite the fact that the order of presentation changes on each trial, is termed 'subjective organisation' and was first observed by Tulving (1962).

The rationale for the use of the verbal learning experiments

in Chapter 4 of the thesis was therefore as follows: if the information about semantic memory structure provided by the repertory grids were valid and information drawn from the semantic memory store is used to organise recall in verbal learning tasks, then organisation of the TBR material according to the grid information should lead to better recall than random organisation of the TBR material. Specifically, it was thought that if the TBR material consisted of words which the grids had shown to be semantically highly related then recall should be better than for words which the grids had shown to be of low semantic similarity.

Further support for the role of semantic memory in verbal learning tasks will now be presented, and some theoretical interpretations discussed. Lines of support for the importance of semantic memory in verbal learning tasks derive from two main sources: studies of organisation of the TBR material and studies of release from proactive inhibition (Wickens, 1970). These two lines of support will now be discussed under separate headings.

Organisation

Evidence for the use of hierarchical semantic structure to aid recall is provided by Bower et al (1969). To one group of subjects they presented a hierarchically organised set of nested categories on a sheet of paper, so that the spatial display reflected the structure implicit in the set of words. To a control group they presented the same words in the same tree-like formations, but in this case the words were scrambled so that the spatial display did not conform to the class inclusion relations that existed among the words. Recall was two to three times better for the structured than the scrambled material.

Studies which have attempted to quantify the effect of organisation on amount recalled for uncategorised stimulus lists, have been confounded by the lack of a method of assessing degree of organisation independent of amount recalled. The measures of organisation which are in use, such as the SO (subjective organisation: Tulving, 1962) and ITR (inter-trial repetition: e.g. Fagan, 1968) measures for the recall of unrelated stimulus lists, are inadequate in that they are measures of output consistency. Postman (1971) maintains that organisation in the sense of the establishment of systematic inter-item dependencies will in many cases not be reflected in a high degree of output consistency. As a network of multiple inter-item relationships develops, the number of equivalent 'pathways' leading from one response to another increases. Since the SO and ITR measures do not reflect multiple associations, or the occurrence of units larger than two words, they are obviously inadequate. The clustering measure (Bousfield & Bousfield, 1966) cannot be substituted in such cases because it is only suitable for the recall of categorised material.

However, methods of assessing the amount of organisation independent of the amount recalled have been provided by Mandler (1967) and Seibel (1967). These methods also permit the measurement of larger size units than mere pairwise contingencies. The basic technique is to allow the subject to sort the TBR material into meaningful groups prior to learning and then to use these groupings as measures of the degree of organisation that the subject has imposed on the material. Output order can be examined to check how closely it reflects the initial groupings. Using this technique, Mandler (1967) showed that subjects who were only given sorting

instructions remembered as many words as subjects who were also told to memorise the words. In a study in which subjects were told to sort into a specific number of categories he found a positive correlation between amount recalled and the number of categories used by the subject in the initial sort. Unfortunately list length was held constant while the number of categories was varied, so number of categories was confounded by number of words per category, since the number of words per category would decrease as the number of categories increased. Seibel (1967) found that subjects whose free recall reflected their initial groupings of the TBR material were the ones who learned fastest. The results of Mandler's and Seibel's studies therefore support the view that organisation enhances recall.

A final source of support for the importance of organisation in recall comes from Tulving's study of transfer effects, (Tulving, 1966).

To assess the functional significance of the subjective organisation achieved during free recall learning, Tulving (1966) used a test of the transfer between successive free recall tasks. In a free recall experiment, requiring the learning of 18 words, half the subjects had previously learned 9 of the 18, while the other subjects had learned a different 9 word list. Although the subjects who had previously learned half the words showed initially higher performance on the 18-word list, their subsequent learning was much poorer than the other group. Tulving predicted this negative transfer effect, hypothesising that the operation of the subjective units formed during the learning of the original 9-word list would continue to be used, although they might not be suitable for encoding the extra 9 words. The control group, on the other hand, would be

forced to create a new set of subjective units, based on the entire set of 18 words.

Because this counter-intuitive theoretical prediction was supported by the experimental result, part-whole transfer effects are of considerable importance in supporting the hypothesis that organisation is important for recall. More specifically, they also suggest that there is an optimum organisation for a list of items which will serve to maximise the subject's performance. However, the origin of the interference in the persistent operation of inappropriate subjective units remains to be conclusively demonstrated.

Also, a recent study by Hicks and Young (1972) has shown that such proactive interference only occurs with lists of nouns, there being no negative part-whole list transfer for adjectives, but rather a positive transfer effect. Such findings do not support the organisational hypothesis. They fit better the traditional theories which emphasise response availability and contextual association. These would predict positive transfer during whole-list learning since half of the experimental items have already had their availability increased and been conditioned to the experimental context. The finding also suggests that different systems may operate in the storage and retrieval of nouns and adjectives.

Although the evidence for the importance of organisation for recall in verbal learning tasks is not totally unequivocal, the volume of supporting studies is sufficiently great for the conclusion that *organisation* is important to be drawn with a fair degree of confidence.

In order for this conclusion to lend credence to the use of verbal learning tasks in the thesis it must be shown that this organisation does depend in large part on semantic memory. The inter-

pretation of the role of organisation in free recall learning which follows should make its reliance on semantic memory more apparent.

As an interpretation of the phenomena of clustering and subjective organisation in free recall learning, Tulving (1962) introduced the ideas of recoding and chunking used by Miller (1956) in his analysis of immediate memory span. Miller pointed out that the capacity of immediate memory depended on the number of 'chunks' of information that it contained rather than on the amount of information per chunk. Tulving postulated that the organisation of the TBR material in free recall learning was a form of 'chunking', i.e. the recoding of the information into another form that was easier to remember, either because the message was physically shorter, (reduction coding) or because additional retrieval cues had been encoded, (elaboration coding).

Information in the semantic memory system plays a role in the organisation of verbal material into higher-order units, i.e. chunks, but the extent of its involvement depends on the basis of the organisation. For example, the three TBR items HAT, CAT and MAT may be organised into 'chunks' in terms of either their pronunciation or similarity of letter sequences. While in both cases the rules for organisation may originate in semantic memory, organisation of the three terms does not depend on the semantic relatedness of the words. On the other hand, if the three words were CAT, LION and TIGER, the encoding of each in terms of its semantic referent would provide the basis for their organisation into a higher order unit.

If the information about semantic memory structure provided by the repertory grids were valid, it should be possible to present subjects with, as TBR words, the words used as elements on the grid, selecting those which were shown to be highly semantically related.

This semantic similarity could be used by the subjects as the basis for forming higher-order units. If recall were better for such words than for words which were shown to be less semantically related then this would provide important support for the validity of the grid data as representing basic aspects of semantic memory organisation.

There is evidence that semantic encoding is the most efficient form of coding under most circumstances, and therefore subjects might be expected to select this form rather than any other which might be available, such as acoustic encoding. For example, Dale and McGlaughlin, (1970), conducted a free recall learning experiment in which lists of names were learnt. If each name was presented with a semantically appropriate occupation, e.g. Mr Law, the policeman, the number of names recalled was increased, although the labels used were not present at recall. That this form of elaboration coding depended on the semantic component was shown by the fact that no facilitation occurred when phonemically similar town names, e.g. Mr. Worth from Perth, were used.

Important evidence for the superiority of semantic coding in verbal learning tasks is also provided by the work of Wickens (1970, 1972) using 'the release from PI' technique. A brief account of his research will now be given.

The release from PI studies

As mentioned in Section 1.2.a Wickens is a proponent of the attribute theory of word storage. He argues (Wickens, 1970, p.1) that 'when a person hears or sees a word, the process of perceiving this word consists of encoding it within a number of different aspects, attributes, or conceptual psychological dimensions.' He uses the release from PI phenomenon as a technique whereby the attributes of

memory may be identified and evaluated. The technique evolved from studies of proactive interference in short-term memory tasks (e.g. Keppel and Underwood, 1962). Wickens et al. (1963) showed that proactive inhibition was specific to the class of material being presented. For example, if proactive interference was built up by presenting a series of three digit numbers then presentation of a consonant trigram would result in a release from the inhibition, i.e. the recall of the trigram would be better than that for the preceding three digit numbers. Wickens (1970) maintains that a change in the stimulus set that leads to a release indicates that this is a relevant dimension along which items are stored and retrieved from memory. The relative amounts of release obtained across different dimensions give an indication of their relative salience. When studying the major encoding dimensions for verbal items Wickens (1970) found a maximum amount of release to be associated with a variety of semantic characteristics of the stimulus items. Wickens (1972) reports that the subject is particularly reactive to the semantic attributes of words, mildly reactive to the mode of their physical presentation and almost impervious to their syntactical properties. He comments that the order of priorities with the superiority of semantic encoding seems to be a most adaptive one in a highly verbal culture such as our own.

Thus in evaluating the relative effectiveness of different attributes in enhancing retrieval of TBR information, Wickens has shown the semantic dimension to be the most salient. Although the experimental paradigm for the release from PI studies differs from the free recall learning and paired-associate learning paradigms used in Chapter 4, Wickens' results do suggest that the semantic dimensions

(e.g. semantic similarity) in the TBR material should be responded to by subjects in these tasks with similar priority.

1.4.d Individual differences in semantic memory function

In Section 1.2.d the examination of individual differences in semantic memory structure in the thesis was discussed. In fact the existence of these differences in structure was inferred from an investigation of individual differences in semantic memory function.

As was stated in Section 1.2.d, semantic distinctions of assured subjective salience were obtained for individual subjects by repertory grid technique. These were then compared in salience with the three distinctions for which a particular ordering in the semantic memory structure based on judgements of their relative salience had been proposed. If reliable differences in the salience of semantic distinctions within individual's semantic memory structures were found, then this would make the attempt to specify the ordering of such distinctions across subjects, in terms of salience, considerably more complex.

The method by which the self-generated semantic distinctions were compared in salience with the three distinctions for which the structure had been proposed was based on the assumption that the differences in salience would be evidenced by differences in the functioning of the self-generated and the experimenter-determined distinctions. This possibility was tested in Experiment 7.

The particular mode of functioning which was studied was the use of the distinctions in a semantic processing task. Specifically, an incidental learning situation was arranged in which subjects processed a set of words either in terms of three of their own distinctions or in

terms of the other three distinctions. The processing involved the subjects writing down for each word which of the poles of the three distinctions were true of that word.

Whether or not processing in terms of self-generated distinctions was more effective was determined by the amount of free recall of the words used as stimuli in the incidental learning session achieved by subjects in a single trial free recall session which immediately succeeded the incidental learning session.

The rationale for the use of amount recalled as an external criterion of salience derived from what is termed the 'levels of processing' approach to memory research (Craik and Lockhart, 1972). The basic tenet of this approach is that the persistence of the memory trace of a stimulus is a function of the 'depth' to which the stimulus is processed or analysed. The extra salience of the self-generated distinctions was presumed to involve a greater depth of semantic processing. Thus the words so processed would be better stored in memory and therefore better recalled.

The use of the term 'depth' in this sense derives from research on perceptual processing. Theorists such as Selfridge and Neisser (1960), Triesman (1964) and Sutherland (1968) view perception as involving the rapid analysis of stimuli at a number of levels or stages. Preliminary stages are concerned with the analysis of such physical and sensory characteristics as lines, angles, pitch brightness, while later stages are concerned with matching the input against stored abstractions from past learning e.g. pattern recognition and the extraction of meaning. This hierarchy of processing stages is often referred to as 'depth of processing' where greater depth implies a greater degree of semantic or cognitive analysis. In Craik

and Lockhart's view the memory trace is a result of this perceptual analysis.

After a word is recognised it may trigger associations, images or stories based on the subject's past experience of the word. It thus undergoes further processing by enrichment and elaboration leading to the formation of a more elaborate, longer lasting and stronger memory trace. The prediction on which Experiment 7 was based was that when subjects processed words in terms of self-generated semantic distinctions the words would be processed to a deeper level than when the three other semantic distinctions were used and would therefore be better remembered. The additional salience of the self-generated distinctions was thought to bring about the greater depth of processing.

Experiment 7 supported the predictions based on the 'levels of processing' approach. Processing in terms of self-generated distinctions resulted in reliably better recall scores. Thus by examining the functioning of self-generated semantic distinctions in terms of the amount of recall they promoted after an incidental learning task, the possibility of the existence of individual differences in semantic memory structure was investigated.

1.4.e Summary of approach to semantic memory function

The role of semantic memory is important in three types of learning experiment in the thesis: in the paired-associate learning experiment, Experiment 5, the free recall learning experiment, Experiment 6, and the incidental learning experiment, Experiment 7. The present section considered semantic memory function in these tasks.

The distinction between episodic and semantic memory (Tulving,

1972) was raised, and the point made that in tasks such as free recall and paired-associate learning it was possible for semantic memory to play a minimal role, serving to do no more than allow the subject to recognise as familiar words used as stimuli. It was possible for episodic memory to bring about recall with no greater reliance on semantic memory than this.

However, consideration of the role of organisation in memory i.e. that the subject may transform the TBR material into a form that is easier to remember by either reduction or elaboration coding, indicated that organisation in terms of semantic characteristics was an effective form of coding.

The work of Wickens (1970, 1972) on the encoding characteristics of verbal items was quoted as further evidence of the importance of their semantic characteristics in recall tasks. Thus the expectation that subjects would make use of semantic organisation present in the TBR material to aid recall was found to be justified by the available empirical data.

Finally the function of semantic memory in incidental learning was considered. Using this paradigm it is possible to determine to which characteristics of the stimulus material the subject attends by the particular experimental instructions. Thus the question of whether or not the subject would use semantic memory did not arise, the use of semantic memory being ensured by instructions to process the stimulus words in terms of particular semantic distinctions. What was in question in the consideration of the role of semantic memory in the incidental learning situation, was not whether the assumption that it would be involved was valid, but rather, whether the experimental task would reflect differences in the salience of the

semantic distinctions used to process the stimulus words. The framework on which differences were predicted in the amount of incidental learning (as indicated by amount recalled in a subsequent single trial free recall experiment) resulting from processing words in terms of semantic distinctions of differing degrees of salience was the 'levels of processing' framework of Craik and Lockhart (1972). According to this framework memory trace persistence is viewed as a function of depth of processing. Processing to a greater depth in terms of more salient semantic distinctions would therefore be expected to result in better subsequent recall. This prediction was tested in Experiment 7 reported in Chapter 5.

1.5 Introduction to the Experimental Chapters

The study of the organisation of non-hierarchical semantic information was developed in three steps in the following chapters. These were:

(1) The testing of a proposed ordering for the three semantic distinctions C/NC, S/L and N/M, based on differences in salience. Experiments 1, 2 and 3 reported in Chapter 2 were carried out to test this ordering. Experiment 1 tested whether the ordering predicted accuracy of assignment of the distinctions to partially known words. Experiment 2 tested whether it predicted retrieval speed when a set of common concrete nouns were sorted according to the three distinctions. Experiment 3 tested whether it predicted salience ratings for the distinctions for a set of common concrete nouns. The failure of these experiments to fully support the proposed ordering was discussed and the possibility raised that one reason for the lack of support might be the operation of confounding

individual differences in the salience of the three distinctions.

(2) The development of Repertory Grid Technique as a method of obtaining self-generated semantic distinctions from individual subjects which would be of assured salience. The main objective here was to provide evidence of the validity of the semantic distinctions generated by grid technique as basic conceptual meaning distinctions and therefore comparable with the C/NC, S/L and N/M distinctions. Consideration was also given to the introduction of grid technique as a general tool for studying semantic memory organisation. Experiment 4 reported in Chapter 3 was an exploratory study of the use of grid technique in the field of semantic memory research.

Experiments 5 and 6, reported in Chapter 4, were verbal learning studies, the first of which provided strong evidence of the validity of the grid as a means of tapping basic semantic memory structure.

(3) The testing of the relative salience of the self-generated and the C/NC, S/L and N/M distinctions. Experiment 7, reported in Chapter 5, compared the salience of self-generated with the experimenter-determined distinctions in terms of the amount of recall they promoted after an incidental learning task. The results indicated that self-generated distinctions were more salient, but the definite conclusion that this was due to individual differences in salience was not possible. A further study to clarify whether the result was due to individual differences in salience or due to the phrasing of the self-generated distinctions causing them to be generally more salient was proposed.

Chapter 2

Semantic Memory Organisation and the Age- of-Acquisition Index

Chapter 2

Semantic Memory Organisation and the Age- of-Acquisition Index

2.1 Introduction

The overall aim of the experiments reported in this chapter was to test the proposed ordering for the three semantic distinctions, 'Consumable/Nonconsumable' (C/NC), 'Solid/Liquid' (S/L) and 'Natural/Manmade' (N/M). The ordering was based on the age at which the distinctions were first acquired, as it was postulated that distinctions of maximum salience would be acquired earliest and that the semantic memory structure would be such as to give them maximum ease of access. The proposed ordering in terms of age of acquisition was C/NC, S/L to N/M, where the C/NC distinction was assumed to be acquired first and the N/M distinction last. This ordering was termed the age-of-acquisition index. Section 1.3 in the Introduction presents a full account of the theoretical background to the proposed ordering and of its implications for semantic memory structure.

Experiment 1 in the present chapter studied the accuracy of assignation of the three distinctions to low frequency words. The prediction was that if the most salient distinctions were acquired first, then for a partially known word the age-of-acquisition index should predict which distinctions could be supplied most accurately.

Experiment 2 tested whether the index would predict sorting times for a set of common concrete nouns according to the three distinctions. If ease of access was greater to distinctions acquired earlier in life then sorting times for these distinctions should be shorter.

Experiment 3 tested whether the index would predict salience

ratings for the distinctions for a set of common concrete nouns. Rated salience should decrease with the index, from a maximum for the C/NC distinction to a minimum for the N/M distinction.

2.2 Experiment 1

The assignation of three semantic distinctions to low frequency words

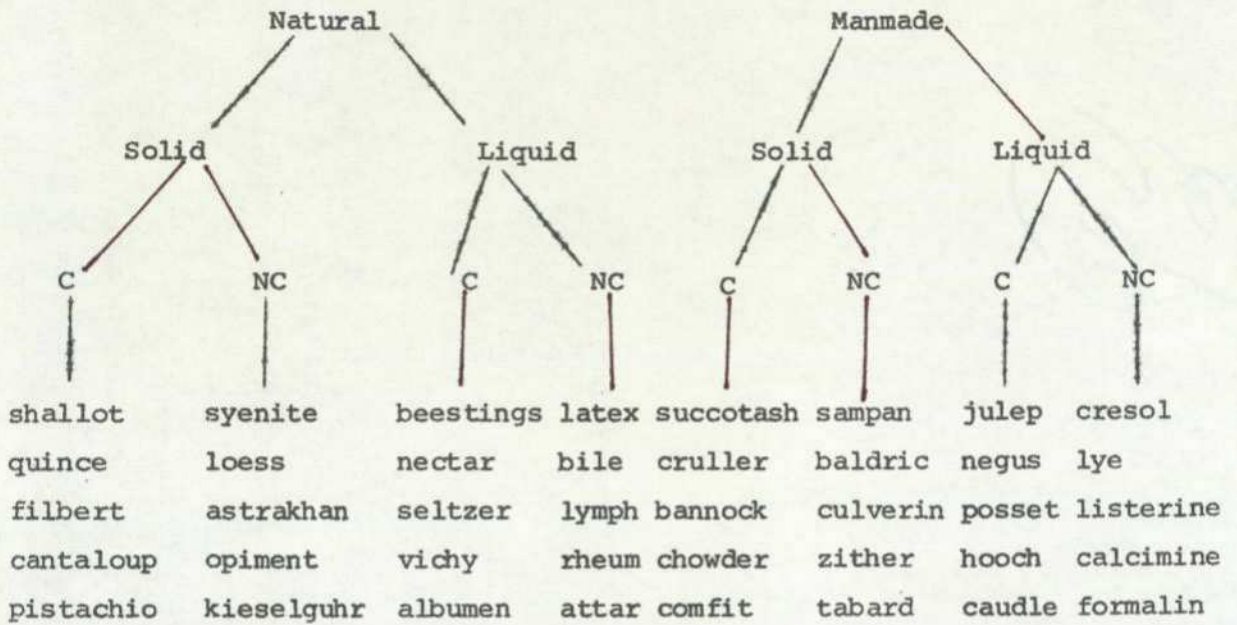
Johnson-Laird (1973) reported a study in which subjects categorised 48 rare words in terms of the three semantic distinctions, C/NC, S/L and N/M. It was found that errors increased in the direction predicted by the age-of-acquisition index. However, as Johnson-Laird (1973, p. 4) pointed out; the trend was: ' . . . to some extent misleading. Many of the words were completely familiar to our subjects, while others of them were sufficiently unfamiliar for subjects only to be able to guess at their meaning.'

The most interesting case to study would seem to be that in which the meaning of a word is partially known to the subject. If the distinctions which the subject could supply for the word were those acquired first according to the age-of-acquisition index, then important support for the validity of the index would be obtained.

The present experiment was therefore conducted with the aim of testing the prediction that for words whose meaning is only partially known, subjects should be able to supply semantic distinctions acquired earlier in life (according to the age of acquisition index) more accurately than those acquired later in life. Specifically, the prediction was that the C/NC distinction should be supplied most accurately, the S/L distinction moderately accurately and the N/M distinction least accurately.

TABLE 2.1: Experiment 1

The 40 low frequency words, broken down into categories according to the three dimensions



METHOD

The subjects were presented with 40 low frequency concrete nouns and were required to give two indications of how well they knew the words (1) by rating each word in terms of an A, B, C scale according to the scheme:

A: You know what the word means, could define it and could use it in the correct way.

B: The word is familiar but you are not sure of its meaning.

C: You have never heard of the word and have no idea what it means.

and (2) by writing a brief definition of each word. The subjects then differentiated between the words in terms of the three semantic distinctions C/NC, S/L and N/M, by writing beside each word the appropriate poles of the distinctions. The words could all be described by one of the poles of each of the dimensions. No time limit was set for the task.

Subjects

The subjects were 20 undergraduate student volunteers at the City University. They were tested in two group sessions.

Materials and Procedure

The 40 concrete nouns were selected from the Thorndike-Lorge (1944) word count. They were chosen from the list of words occurring at least once per 4,000,000 words but not so often as once per 1,000,000 words. They divided evenly into the eight categories formed by the interaction of the three semantic markers. To balance the categories it was necessary to include a few words whose frequency could not be checked in the Thorndike-Lorge. The words are shown in Table 2.1

Two sets of answer booklets were made up, in each of which the 40

METHOD

The subjects were presented with 40 low frequency concrete nouns and were required to give two indications of how well they knew the words (1) by rating each word in terms of an A, B, C scale according to the scheme:

A: You know what the word means, could define it and could use it in the correct way.

B: The word is familiar but you are not sure of its meaning.

C: You have never heard of the word and have no idea what it means.

and (2) by writing a brief definition of each word. The subjects then differentiated between the words in terms of the three semantic dimensions C/N, S/V and N/M, by writing beside each word the appropriate poles of the distinctions. The words could all be described by one of the poles of each of the dimensions. No time limit was set for the task.

Subjects

The subjects were 30 undergraduate student volunteers at the City University. They were tested in two group sessions.

Materials and Procedure

The 40 concrete nouns were selected from the Thorndike-Lorge (1944) word count. They were chosen from the list of words occurring at least once per 4,000,000 words but not so often as once per 1,000,000 words. They divided evenly into the eight categories formed by the interaction of the three semantic markers. To balance the categories it was necessary to include a few words whose frequency could not be checked in the Thorndike-Lorge. The words are shown in Table 2.1

Two sets of answer booklets were made up, in each of which the 40

words were presented, 8 words to a page. ^{Two} different random orderings of the words on each page were used and the ordering of the 5 pages in each booklet was also done on a random basis. Beside each word in the first set of booklets the letters A, B and C were printed. The front page of these booklets explained the use of the A,B,C rating scheme and instructed the subject to circle for each word which letter was appropriate. For each word a brief definition was also to be written in the space next to it. If the subject did not know the word, he was instructed to make a guess at its meaning and write that down.

The front page of the second set of booklets instructed the subject to write down next to each word whether the substance or object to which it referred was Consumable or Nonconsumable, Solid or Liquid and Natural or Manmade. Definitions of the three distinctions were given. The sequence of these three distinctions was presented in three different random orders in the instructions. Subjects were instructed to use guesses when they did not know.

The subjects were told that the experiment concerned the meaning of words, but was not a test of ability, the words having been selected intentionally to be unfamiliar to the subjects. The subjects were tested as a group, the second set of booklets not being distributed until the first had been completed, so that the distinctions written on the second set would not bias the writing of definitions in the first set.

The subjects' definitions were jointly scored by a panel of three judges as either correct, partially correct, or wrong.

TABLE 2.2: Experiment 1

The percentage of words for which the semantic distinctions
were assigned correctly expressed in terms of whether the words
received an A, B or C rating

<u>A (N = 237)</u>			<u>B (N = 172)</u>			<u>C (N = 391)</u>		
C/NC	S/L	N/M	C/NC	S/L	N/M	C/NC	S/L	N/M
91.9	91.6	90.6	68.9	65.1	66.9	52.2	56.0	50.4

TABLE 2.3: Experiment 1

The percentage of words for which the semantic distinctions
were assigned correctly expressed in terms of whether the
definitions given for the words were Correct, $\frac{1}{2}$ Correct or Wrong

<u>CORRECT (N = 213)</u>			<u>$\frac{1}{2}$ CORRECT (N = 102)</u>			<u>WRONG (N = 485)</u>		
C/NC	S/L	N/M	C/NC	S/L	N/M	C/NC	S/L	N/M
96.2	96.7	94.5	87.3	73.5	67.6	51.3	51.8	51.5

RESULTS

Two methods of assessing the subjects' familiarity with the words were used: the A,B,C familiarity rating and the accuracy of the subjects' definitions. The data for each of these methods is presented separately. Table 2.2 shows a breakdown of the data in terms of the A,B,C rating, giving for each distinction, the percentage of words for which the distinction was assigned correctly. Table 2.3 shows a breakdown of the data in terms of correct, partially correct and wrong definitions, giving for each distinction the percentage of words for which it was assigned correctly.

As an inspection of the tables shows, the percentage of words which received an A rating or which were defined correctly and for which the semantic distinctions were correctly assigned was very high, approaching a ceiling effect of 100% correct. This was to be expected; for well known words all the distinctions should be able to be assigned correctly. The percentage of words which received a C rating or which were incorrectly defined and for which the semantic distinctions were correctly assigned was low, very near chance level. This was also to be expected; for unfamiliar words the distinctions would not be able to be assigned with greater than chance correctness. Page's L trend tests within the four categories, A, C, Correct and Wrong, showed that in no case did a trend approach significance.

As argued in the introduction to the experiment, the significant group of words to examine for the presence of the predicted trend was the group whose meaning was only partially known. For these words, if the age-of-acquisition index were valid, a trend of decreasing accuracy in distinction assignment in the direction C/NC, S/L and to N/M was predicted. Table 2.2 shows that when the data for partially

known words were analysed in terms of the subjects' familiarity ratings (category B), the accuracy of distinction assignment did not follow the predicted trend. A Page's L trend test on this data was nonsignificant. However, when the data for partially known words were analysed in terms of the subjects' definitions (Table 2.3) the accuracy of distinction assignment did follow the predicted trend. A Page's L trend test on the data for the ' $\frac{1}{2}$ correct' words showed that the trend was reliable, $L(3,20) = 258\frac{1}{2}$, $p < 0.01$.

DISCUSSION

A reliable trend of decreasing accuracy in distinction assignment in the direction C/NC, S/L to N/M was found for words whose definition was $\frac{1}{2}$ correct. This trend was predicted by the age-of-acquisition index and therefore provided *good* empirical support for the index.

The predicted trend effect was not found for partially known words as indicated by the subjects' ratings of familiarity. The disparity between the two measures of word familiarity, i.e. the familiarity ratings and the subjects' definitions, might have been caused by the fact that the A,B, C rating scheme relied on 'feelings of knowing' (Hart, 1967), while the definition method required objective proof of knowledge. Confusion of a word with another that looked or sounded the same might have given rise to a 'feeling of knowing', when in fact the words was unknown. The subjects' definitions provide examples of such confusion e.g. mistaking cantaloup for cantilever. Such confusions would explain the greater number of words classed as A and B in comparison with the number of words for which a correct or partially correct definition was given (see Tables 2.2 and 2.3). In view of this possibility the definition data would seem to

provide the more accurate means of assessing word knowledge.

Although the differences in accuracy of distinction assignation did not support the predictions of the age-of-acquisition index for partially known words as specified by the familiarity ratings, the possibility of inaccuracies in the use of the rating method indicated that this negative result should not be given too much weight. As the subjects' definitions seemed to be the more reliable measure of word familiarity, the positive results of the analysis of accuracy of assignation in terms of this measure were upheld.

Experiment 1 had thus provided some empirical support for the validity of the age-of-acquisition index. It had been shown that the index could predict the ordering of the three semantic distinctions in terms of their accuracy of assignation to partially known words.

In order to obtain further support for the index the testing of its predictive power for the ordering of the three distinctions in terms of a behavioural measure other than accuracy of assignation was proposed. Accordingly, Experiment 2 tested whether the index could predict sorting times for a set of words according to the three distinctions. If the assumptions on which the index was based were valid, then the organisation of the distinctions in the internal lexicon should be such as to give greater ease of access to the more important distinctions. Thus retrieval speed in a sorting task was predicted to decrease in the direction C/NC, S/L to N/M, from a maximum for the most important distinction. Experiment 2 was carried out to test this prediction.

2.3 Experiment 2

Differences in retrieval time for three semantic distinctions in a sorting experiment

Experiment 1 provided some evidence of differences in accuracy of assignation of the three semantic distinctions C/NC, S/L and N/M which were predicted by the age-of-acquisition index. The next test of the index was to find out whether it predicted the speed with which the semantic distinctions were retrieved for common words.

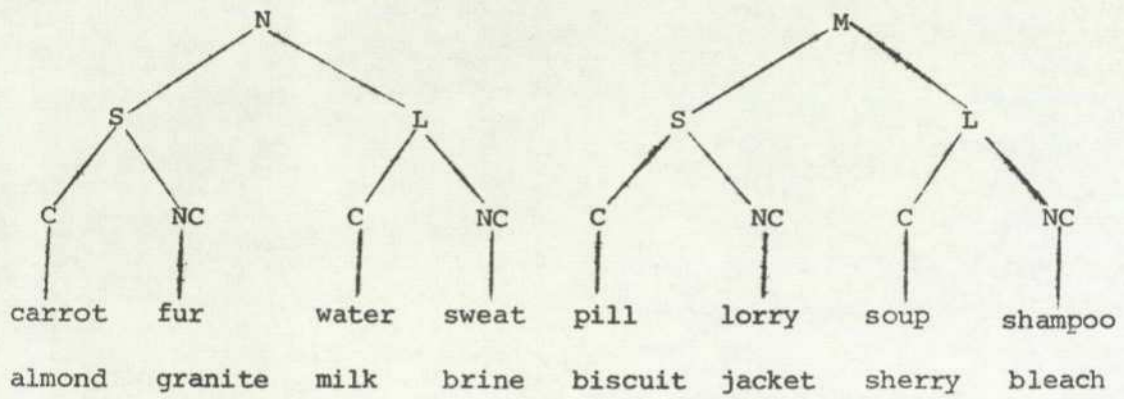
In a preliminary test of whether the index predicted the speed with which such semantic information was retrieved, Johnson-Laird and Shapiro (unpub.) examined sorting times for 24 concrete nouns. When subjects were asked to sort the words on each of the three distinctions, their sorting times showed a reliable trend in the required direction.

The present experiment was designed to replicate Johnson-Laird and Shapiro's results using a different set of words. Johnson-Laird and Shapiro's study had used nouns of both low and high frequency (in order to test an additional hypothesis). In the present experiment this source of variability was excluded, and only high frequency nouns were used.

The prediction was that sorting times would be fastest for the C/NC distinction, of an intermediate speed for the S/L distinction and slowest for the N/M distinction.

TABLE 2.4: Experiment 2

The 16 concrete nouns categorised according to the
3 semantic distinctions



METHOD

Design

The subjects sorted the same set of words according to the two poles of each of the three semantic distinctions. Since the same set of words was used for each sorting, any reliable difference in sorting time can be attributed to the relative difficulty of retrieving the various elements of meaning, rather than to the time taken either to recognise the words, or to make the appropriate physical movements.

Each of the 3 sorting tasks was performed twice, and the order in which the six sorts was executed was randomised both within and between subjects.

Materials

The set of words to be sorted consisted of 16 common concrete nouns, which broke down evenly into the eight categories formed by the interaction of the three logically independent distinctions. The words are shown in Table 2.4.

The words were typed on 5" by 3" library cards. They were sorted into 8" by 6" boxes. For each sort the fronts of the boxes were clearly labelled with the appropriate distinctions.

Subjects

The subjects were 18 undergraduate volunteers at the City University who were paid for their services.

Procedure

The subjects were tested individually. They were told that the experiment concerned the meaning of words, and that they would be required to sort a set of words a number of times into 2 distinct categories. This was to be done as quickly and accurately as possible.

The sorting task was not a test of their ability, but was intended to discover if different people categorised words in the same way. They were told that they would be timed and that they should sort the words as quickly as possible without making any mistakes. If they did place a card in the wrong box, they should not pause to correct this. The first sort was always the practice condition in which the subject sorted the words into those which denoted a substance which he 'liked' and those which he 'disliked'. During this sort the subject was instructed to indicate if any of the words were unfamiliar. In fact none of them were found to be unfamiliar.

After the practice condition, the subject performed the experimental sortings. For each of the experimental sortings the nature of the categories was explained to the subject and the two appropriately labelled boxes placed in front of him. The left-right positioning of the boxes for the two distinctions poles was randomised across subjects. The pack of cards to be sorted was placed face down in front of the subject. He was given time to orient himself to the new distinction. Then at a signal from the experimenter, he turned over the pack of cards and began sorting. The time in which he completed the sort was recorded using a stopwatch. The cards were thoroughly shuffled between trials.

RESULTS

As each subject sorted the words twice according to each distinction, the mean sorting time for the two sorts for each distinction was taken. The mean overall time in seconds taken to sort according to the three distinctions is shown in Table 2.5.

Table 2.5 shows that the sorting times for the three distinctions

TABLE 2.5: Experiment 2

The mean sorting time in seconds for the three
distinctions (N = 18)

C/NC		S/L		N/M	
<u>Mean</u>	<u>Standard</u> <u>Deviation</u>	<u>Mean</u>	<u>Standard</u> <u>Deviation</u>	<u>Mean</u>	<u>Standard</u> <u>Deviation</u>
16.30	3.60	15.77	3.65	19.58	3.94

did not increase in the direction predicted by the age-of-acquisition index, i.e. from a minimum for the C/NC, through the S/L to a maximum for the N/M distinction. In this case the S/L distinction would be placed first in the ordering, as it had the shortest sorting time.

An analysis of variance on the individual mean sorting times (Table A1) showed that the difference in sorting times between the distinction was significant $F(2,34) = 21.72, p < 0.001$. However unplanned comparisons using the Scheffé test showed that the difference in sorting time between the C/NC and the S/L distinctions was not significant, although the difference between each of these distinctions and the N/M distinction was significant, the F ratios being respectively $F(1,34) = 24.41, p < 0.001$ and $F(1,34) = 36.92, p < 0.001$. On the basis of this analysis therefore, the ordering of the distinctions would be C/NC and S/L first (having roughly equal sorting times) and N/M last (having a reliably longer sorting time).

It was noted that the analysis of variance (Table A1) also showed a significant subject effect $F(17,34) = 10.88, p < 0.001$. The possible *masking* of the results due to this effect is considered in the Discussion to the experiment.

The number of sorting errors, i.e. sorts in which a word was placed in the box of the wrong distinction pole, were also recorded for each distinction. As Experiment 1 had shown that accuracy of distinction assignment decreased according to the age-of-acquisition index, it was thought that sorting errors might increase with the index, i.e. from a minimum for the C/NC through the S/L to a maximum for the N/M distinction. Table 2.6 shows the ^{total} number of sorting errors

TABLE 2.6: Experiment 2

The number of sorting errors for each distinction

C/NC	S/L	N/M
32	14	81

for each distinction. The prediction of increasing sorting errors according to the age-of-acquisition index was not confirmed, as the least errors occurred for the S/L rather than the C/NC distinction.

The number of sorting errors was small, the total number of sorts for each distinction being 576 (18 subjects sorting 16 words twice for each distinction.) However, a Friedman two-way analysis of variance by ranks showed that the difference in number of errors between the distinctions was significant, $\chi^2 (2) = 10.75$, $p < 0.01$. A Wilcoxon matched-pairs signed-ranks test showed that for a two-tailed test the difference in errors between the C/NC and the S/L distinction was significant $T (12) = 12$, $p < 0.05$, that between the C/NC and the N/M distinction was significant $T (13) = 10.5$ $p < 0.02$, and that between the S/L and the N/M distinction was significant $T (16) = 5$, $p < 0.01$. Thus in terms of number of errors there was a reliable ordering in the direction of increasing errors from the S/L through the C/NC to a maximum in the N/M distinction and the prediction of increasing errors in the direction of the age-of-acquisition index, i.e. C/NC, S/L to N/M, was not supported.

A representative proportion of the subjects commented that it was easiest to sort according to the S/L distinction because there was only a unitary feature to consider i.e. the physical fact of whether a substance was solid or liquid. In the same way, the C/NC distinction was more difficult because certain complex features were involved. For example, certain substances may be deemed non-consumable by taboo rather than because they are physically harmful. Likewise, certain substances though not habitually consumed, might be difficult to categorise as totally nonconsumable. Many subjects remarked on the greater difficulty of sorting according to the N/M

distinction, pointing out that it not only involved a complex decision but also an ambiguous one. Almonds became manmade when they were sugar almonds, fur is now manufactured synthetically and milk is manmade in one sense when human milk is considered.

DISCUSSION

With regard to sorting times, the results indicated that there were consistencies in the speed with which subjects sorted words according to the three semantic distinctions. The C/NC and S/L distinctions elicited roughly equal sorting times while the N/M distinction elicited significantly longer sorting times. The reliably longer sorting time for the N/M distinction was predicted by the age-of-acquisition index, but as no difference between the S/L and C/NC distinctions emerged, the predictions based on the index were not fully supported.

With regard to sorting errors, the results indicated that the ordering of the distinctions should be S/L, C/NC, N/M in the direction of increasing errors from a minimum for the S/L distinction. The reliably greater number of errors for the N/M distinction was predicted by the age-of-acquisition index but the ordering of the S/L and C/NC distinctions in terms of number of errors was not. Thus as with sorting times, the sorting error data provided only partial support for the age-of-acquisition index.

In the assessment of the theoretical basis of the index, in Section 1.3.b, the point was made that the prediction that the distinctions C/NC, S/L and N/M would be organised in the semantic memory structure in such a way that ease of access to the C/NC and S/L distinctions would be greater than to the N/M distinction concurred

with Piaget's theory of cognitive development, but that predictions of a difference in ease of access between the C/NC and S/L distinctions were more difficult to justify theoretically. Both the C/NC and S/L distinctions could be learnt during the sensori-motor period (0-2 years), while the abstract N/M distinction would probably not be learnt until the period of concrete operations (6-7 years).

Reference to Piagetian theory therefore suggested that the age-of-acquisition index for the distinctions C/NC, S/L and N/M would be better represented as $C/NC = S/L < N/M$. Thus for example, considering sorting times as correlated with age-of-acquisition, no difference between the C/NC and S/L distinctions would be expected, but both of these distinctions would be different from the N/M distinction. Specifically, as ease of access was greater for the C/NC and S/L distinctions, these would be expected to elicit reliably shorter sorting times than the N/M distinction. A possible reason for the failure of the experimental result with regard to sorting times to fully support the age-of-acquisition index, might be the failure of the distinctions used to adequately represent a developmental sequence.

The finding that the number of sorting errors was least for the S/L distinction did not concur with either the predictions of the age-of-acquisition index or predictions based on a consideration of Piagetian theory. However, the subjects' introspective reports indicated that of the three distinctions, this was the only one that required a unitary simple decision. A decision about the physical fact of liquidity or solidity might be less prone to error than decisions for the other two distinctions. The subjects claimed that the C/NC distinction was not a unitary or simple one and the N/M distinction was felt to be similarly complex and ambiguous.

Perhaps for the child a clear set of consumable and nonconsumable substances exists, but as knowledge of the world increases, the distinction may become less precise. Similarly decisions about the N/M distinction might involve a consideration of a number of pieces of encyclopaedic knowledge. For example, when deciding whether to classify 'fur' as natural or manmade it might be necessary to weigh the abundance of synthetic fur fabrics against the fact of fur being the natural outer covering of many animals. Thus the greater error score for the C/NC and N/M distinctions might be due to the complex decision required in their often ambiguous assignation. That complex rather than unitary decisions might be involved was a further indication of the possible unsuitability of the distinctions chosen to adequately represent the age-of-acquisition index.

Johnson-Laird and Shapiro (unpub.) suggested another reason which might explain the present experiment's failure to fully support the age-of-acquisition index. They pointed out (p. 14) 'Whether consistencies in retrieval times are truly independent of the material used can only be settled by further research'. Reference has already been made to Paivio's work on imagery (Section 1.2.e). It was suggested that one type of effect that would be dependent upon the specific words used as sorting materials was differences in imagery value. Paivio (1971) stressed that words vary in their imagery value or 'concreteness'. If a word evoked a strong iconic image, then its perceptual features might dominate and so the S/L distinction would be retrieved faster than the other two distinctions. Thus for such words, sorting times might be shortest for the S/L distinction, even though the C/NC distinction was acquired first in childhood. The low error score in the present experiment for the S/L

distinction might be due to its clear imagery.

In order to assess the extent to which the experiment was confounded by effects specific to particular words, the examination of the response to the three distinctions for individual words was proposed. Experiment 3 was conducted to investigate this possibility.

The significant subject effect shown by the analysis of variance on the sorting times (Table A1) was noted. The possibility of individual differences in the salience of the three distinctions confounding the emergence of the overall trend was investigated in Experiment 7.

2.4 Experiment 3

Differences in the salience of three semantic distinctions for individual words

As pointed out in the Discussion section to Experiment 2, Johnson-Laird and Shapiro (unpub.) were aware that the consistencies in retrieval times which they observed for their three distinctions might have been an artefact of the particular words used as stimuli. This problem could only be clarified by a study in which response times to individual words were examined.

Consideration of specific word effects was also consonant with the point of view expressed in Clark's important paper (Clark, 1973) on the 'language-as-fixed-effect fallacy', in which he discusses the common empirical error of treating the stimulus words as a fixed rather than as a random effect. Researchers have often implicitly assumed that the sample of words used is representative of the complete population of words to which they wish to generalise. Rather than

overgeneralising, without providing statistical evidence which permits any generalisation beyond the specific materials, Clark in his section titled 'The Method of Single Cases', (Clark, 1973, p. 354) stresses the value of collecting and presenting data for each instance separately, pointing out that each word can be interesting in its own right.

Experiment 3 was therefore designed to investigate specific individual word effects. Basically, subjects were required to select, for individual words, which of the three distinctions, C/N, S/L, N/M was the most salient. The aim was to find out whether there were specific word effects which might confound the emergence of an overall age-of-acquisition trend. For example, reference to Paivio's work on imagery (Paivio, 1971) suggested that if a particular word evoked a strong iconic image, its perceptual (S/L) features might dominate, leading to a trend of decreasing salience from S/L through C/NC to N/M. The age-of-acquisition index would, of course, predict a trend of decreasing salience from C/NC through S/L to N/M for an individual word.

METHOD

Design

Subjects were presented with the same eight words three times. On each occasion the word was tagged with a different combination of two of the three relevant semantic distinctions. The subjects were required to circle the distinction which seemed most important for that word. Thus the word CARROT would appear with CONSUMABLE/SOLID, SOLID/NATURAL and NATURAL/CONSUMABLE. The order of presentation of the words was randomised. The order of presentation of the three combinations of distinctions was also randomised. The number of times a particular distinction was

stated first in the combination, and the distinction with which it was in combination when it was stated first, was balanced across the three distinctions.

Materials

Eight words were used; this was thought to be a sufficient number to show up any individual word effects. A greater number of stimulus words might lead ^{to} response saturation and fatigue. The words used were a subset of those used in the previous experiment. They are shown in Table 2.7. Only the appropriate poles rather than the complete bipolar distinctions were presented for each word to make the subject's choice of the most salient distinction easier. However, this revealed an interesting lack of balance in the importance of the two poles within each distinction.

24-page booklets were compiled in which the eight words were presented three times, one word per page. The words were typed in upper case and, beside each, one of the three relevant combinations of distinctions was typed in upper case. On the front of the booklets definitions of the three semantic distinctions were printed.

Subjects

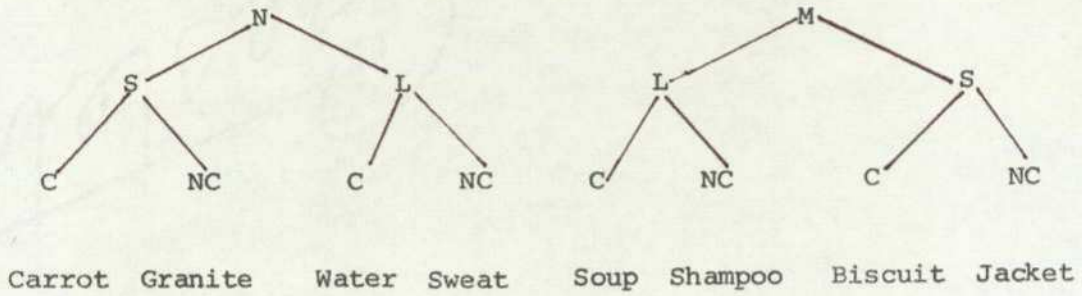
The subjects were 43 first year undergraduate students from the City University.

Procedure

The subjects were tested in two group sessions. The experiment took 5-10 minutes to complete. The booklets were distributed. Subjects were instructed to work quickly through the booklet, circling for each word which of the two distinctions written beside it seemed most important for that word. A notice stressing that subjects should circle which distinction 'STANDS OUT MOST' was written on the blackboard.

TABLE 2.7: Experiment 3

The eight stimulus words differentiated according to
the three semantic distinctions



RESULTS

It was thought possible that a strong response bias towards circling the first distinction in any combination might emerge, but a count of the number of distinctions circled first and second showed that this had not occurred. The distinctions were ranked in order of salience for each word. If the subject was consistent in his three choices of the most salient distinction in the combinations of two for each word, then he produced a system of circling in which one distinction was circled twice, one once and one not at all. For example, if for the word CARROT, a subject circled C out of the C/S combination, S out of the S/N combination and C out of the N/C combination then he would have indicated that the order of salience of the distinctions for that word was $C > S > N$. This ordering was transformed into ranks by giving C rank 3, because it had been circled twice (on both of the occasions on which it was presented) S rank 2 because it had been circled on one of the two occasions on which it had been presented and N rank 1 because it had been circled on neither of the occasions on which it was presented. If the subject was not consistent in his ordering of the three distinctions, then this resulted in each distinction receiving the rank of 2, because each would have been circled an equal number of times (i.e. once each). However, this did not occur often, a maximum of 5 inconsistent orderings being found for the word SWEAT. Table 2.8 shows the mean rankings of the three distinctions for the eight words. As can be seen from the column totals in this Table, when the rankings for each distinction were summed across words, no trace of the predicted trend was found. The total salience for each distinction was the same. Thus the hypothesis that there would be a trend of decreasing salience

from the C/NC distinction through the S/L distinction to a minimum in the N/M distinction was not confirmed. The predictions based on the age-of-acquisition index were not supported.

By contrast the prediction that there would be significant specific word effects was strongly confirmed. Kendall coefficients of concordance showed that the agreement among the subjects on the salience rankings was significant for each of the eight words. These results are summarised in Table 2.9. As can be seen from the Table, only the rank ordering for SOUP was in the direction predicted by the age-of-acquisition index. It would seem that the importance of the three semantic distinctions varied considerably from word to word.

It was possible, however, to draw out certain more general characteristics of the ranking by considering the rankings assigned to the two poles of each distinction separately. Table 2.10 shows the rank totals for the two poles of each distinction separately. A marked lack of equivalence in salience between the two poles of each distinction can be seen. The least salient of all the poles was the NC pole. Thus it could be predicted that for words for which this pole was appropriate, the poles of other two distinctions would receive higher salience rankings. This was the case for GRANITE and SWEAT, but not for SHAMPOO and JACKET. The S pole was also of low salience, and this could have replaced the NC pole as the lowest rank for JACKET. For SHAMPOO, the difference between the NC and the MM pole was small, both receiving considerably lower rankings than the L pole. In contrast to the NC pole, the C pole was of high salience. Thus for three out of the four words for which this pole was appropriate, it received the highest ranking.

TABLE 2.8: Experiment 3

The mean rankings of the three distinctions for the
eight words

	C/NC	S/L	N/M
Carrot	3	1	2
Granite	1	3	2
Soup	3	2	1
Shampoo	2	3	1
Water	1	3	2
Sweat	1	2	3
Jacket	2	1	3
Biscuit	3	1	2
Total	<u>16</u>	<u>16</u>	<u>16</u>

TABLE 2.9: Experiment 3

Kendall coefficient of concordance W, showing the degree
of agreement between the 43 subjects on the salience
rankings of the three distinctions for the eight words

K = 43, N = 3.

Word	Ranking of Distinctions			S	W	P
	<u>3</u>	<u>2</u>	<u>1</u>			
Carrot	C	N	S	1194	0.3154	0.01
Granite	S	N	NC	784	0.2119	0.01
Soup	C	L	M	1959	0.5162	0.01
Shampoo	L	NC	MM	446	0.1176	0.01
Water	L	N	C	739	0.1947	0.01
Sweat	N	L	NC	962	0.2536	0.01
Jacket	MM	NC	S	2236	0.5893	0.01
Biscuit	N	MM	S	1742	0.4593	0.01

TABLE 2.10: Experiment 3

The summed rankings of the salience of the three
distinctions for the eight words (N = 43)

Word	Distinction		Distinction		Distinction	
	C	NC	S	L	N	M
Carrot	103		58		97	
Granite		58	113		87	
Soup	109			100		49
Shampoo		80		103		75
Water	59			110	89	
Sweat		61		95	102	
Jacket		82	54			122
Biscuit	120		67			71
Total	391	281	292	408	375	317

DISCUSSION

It seemed that certain more general but complex effects due to the interaction of the distinctions were present in the data. General statements about the salience of particular poles of distinctions, when in combination with particular other poles, might be proposed. There was insufficient data to allow such proposals to be tested statistically, but a further experiment in which more data about particular combinations of distinction poles were collected would seem worthwhile.

Another point of interest in the data was that the high salience of the L pole was open to explanation in terms of Paivio's work on imagery, (Paivio, 1971). Paivio found that words which subjects rated as of high image-evoking capacity were better recalled in free recall learning tasks. In a similar way, the perceptual characteristics of the L pole may have been more potent than the perceptual characteristics of the S pole and thus contributed to its salience.

Finally, it should be said that although the presence of complex interactions between the poles of the distinctions might explain the seemingly individual and specific nature of the salience rankings for each word, there nevertheless appeared to be certain words for which certain of the poles were of particular salience. For example, from Table 2.10, GRANITE would seem to be particularly S, WATER particularly L, Jacket particularly M and Biscuit particularly C. Elucidation of any such specific word effects must await further research.

The findings of this Experiment offer a clear explanation for the failure of Experiment 2 to obtain support for the age-of-acquisition index. The emergence of differences in sorting times between the three distinctions may have been confounded by both differences in salience between the two poles of each distinction and by the effects on salience

of the complex interactions due to the combination of words in certain poles, quite apart from any differences in sorting time due to specific word effects.

2.5 General Discussion

The overall aim of the experiments reported in this Chapter was to test the age-of-acquisition index as a means of predicting accuracy of assignation (Experiment 1), response latency (Experiment 2) and salience (Experiment 3) for three semantic distinctions. If support for the index was obtained, then it might be used to specify the ordering of other semantic distinctions for which no natural a priori hierarchy was apparent. This would form a useful contribution to research attempting to map the structure of the internal lexicon. However, the overall pattern of results did not provide strong support for the index.

The index predicted that there would be a trend of increasing latency and decreasing accuracy and salience from the C/NC distinction, through the S/L distinction to the N/M distinction. For words whose meaning was partially known, Experiment 1 provided evidence of a reliable trend of decreasing accuracy in distinction assignation in the predicted direction. This positive result was obtained when the partially known words were selected with reference to the subjects' definitions of the words, but not when they were selected with reference to the subjects' familiarity ratings of the words. However the former measure of familiarity seemed more accurate, the latter being open to mistaken feelings of knowing consequent upon acoustic and visual confusion. The support for the index provided by the analysis of the data in terms of the subjects' definitions was therefore upheld.

In Experiment 2, as the index predicted, the N/M distinction produced a reliably longer latency, although no difference in sorting times was found for the C/NC and S/L distinctions. The number of sorting errors was at a maximum for the N/M distinction also, but here the trend was from a minimum for the S/L distinction, the C/NC distinction receiving reliably more deviant sorts than the S/L distinction.

In Experiment 3 analysis of the subjects' salience ratings for the three distinctions produced no trace of the age-of-acquisition trend, but evidence of significant specific word effects, showing reliable patterns of salience rankings for individual words, was obtained.

Taken together, these results provided some support for the age-of-acquisition index, but the overall pattern did not allow the validity of the index to be strongly confirmed.

A number of possible reasons for this lack of stronger support for the index were proposed. It could be that predictions based on age-of-acquisition variables did not maintain for adult subjects: the categorisation that was most useful for very young children might change as they grow older. This point made by Brown (1958) was discussed in Section 1.3.b in the Introduction. However, Carroll and White (1973) found that age-of-acquisition was an important determinant of naming latency in adults.

A second possibility was that the particular distinctions chosen did not adequately represent the developmental sequence on which the age-of-acquisition index was based. The choice of distinction was questioned on two accounts: (i) it was difficult to distinguish on theoretical grounds between the primacy of acquisition of the C/NC and

S/L distinction (ii) several subjects stated that the assignation of the C/NC and N/M distinctions required a complex decision rather than a unitary simple one. As the distinctions were supposed to represent basic elements of meaning, their assignation should by definition require only a unitary decision. If a complex decision were involved this suggested that subjects were considering complexes of features when using the C/NC and N/M distinctions. In order to test the index properly these distinctions should be replaced by ones for which a unitary decision was possible. Unfortunately the lack of a comprehensive list of semantic features makes replication with a different set of distinctions difficult.

A third possibility was that the emergence of the predicted trend was to some extent obscured by specific word effects. Experiment 3 confirmed this possibility and suggested the possibility of more complex interactions between the three distinctions. For example, if a substance were solid and consumable, then its consumable properties would be most important, but if the substance were liquid and consumable then the importance of these two properties would be much more equal. Further research into the operation of these complex interactions seemed worthwhile but was beyond the scope of the present thesis.

Consideration of the significant subject effect produced by the analysis of variance on the sorting data in Experiment 2 suggested yet a further source of variability. Individual differences in sorting times might reflect differences in the subjective salience of particular distinctions. Experiment 3 could be developed to test this possibility, by requiring subjects to indicate the relative salience of the three distinctions for a larger number of words. This

was not done initially, because it was felt that the subjects' sensitivity to differences in salience would rapidly become blunted when applied to too many words; the saturation point would soon be reached. Rather than attempting to modify Experiment 3 to enable the importance of subjective salience to be assessed, a new experiment was conceptualised in which subjects would supply their own distinctions for words such as the stimulus words used in Experiments 1, 2 and 3. These subjective distinctions could then be compared with the C/NC, S/L and N/M distinctions on some external criterion associated with salience and differences in salience thus assessed. Experiment 7, reported in Chapter 6, represents such an attempt to assess the importance of the subjective element in determining the salience of semantic distinctions.

The final explanation proposed for the lack of support for the age-of-acquisition index was that the phrasing of the distinctions by Johnson-Laird might not concord with the subject's personal expression of the distinction. For example, the term 'consumable' might not be the one used to express whether or not a substance could be eaten or drunk. If there were differences in expression, this might have led to the need to reinterpret the distinction into personalised categories, thus lengthening the response time. Differences in the phrasing of particular distinctions were examined in Experiment 7, but further research is needed to fully assess the importance of this factor.

Two major issues which arose from the consideration of the overall results of Experiments 1, 2 and 3 were (i) the question of the importance of individual differences in the salience of the distinctions and (ii) the question of whether the phrasing of the distinctions concurred with

the subjects' personal expressions of the distinctions; if it did not, this constituted a further source of variance.

Access to semantic distinctions generated by the subjects for a set of common concrete nouns such as that used in Experiments 2 and 3 was thought to be a useful preliminary step in answering these questions. In order to obtain subject-generated semantic distinctions the use of Repertory Grid Technique was proposed. If it could be shown to produce meaningful information about the semantic memory structure of individual subjects, through its elicitation of semantic distinctions from individual subjects, then it would constitute a most useful tool for semantic memory research. It would be of specific use in allowing differences in the salience and phrasing of individual subjects' semantic distinctions to be examined.

Chapter 3 reports an experiment (Experiment 4) in which the initial use of Repertory Grid Technique in this field of semantic memory research was investigated. Chapter 4 reports two experiments (Experiments 5 and 6) which tested the validity of the information about semantic memory organisation generated by the technique, and from which the tentative conclusion was drawn that it did provide information about basic and fundamental aspects of semantic memory structure. Chapter 5 reports an experiment (Experiment 7) in which individual subjects' semantic distinctions generated by grid technique were compared in salience with the C/NC S/L and N/M distinctions and in which the effects of differences in the phrasing of the distinctions were also considered. Clarification of these two issues was necessary for a more precise assessment of the validity of the age-of-acquisition index.

Chapter 3

Repertory Grid Technique and Individual Differences in Semantic Memory Structure

Chapter 3

Repertory Grid Technique and Individual Differences in Semantic Memory Structure

3.1 Introduction

The experiments reported in Chapter 2 failed to provide conclusive support for the proposal that three semantic distinctions might be organised in semantic memory according to the age at which they were first acquired. The proposal was based on the assumption that age of acquisition was an indication of a distinction's salience, distinctions of the greatest salience and utility being acquired first.

A suggested reason for the lack of support was that individual differences in the salience of particular semantic distinctions might have acted against the ordering predicted by the index. The experiment reported in the present chapter was an exploratory examination of individual semantic organisation, conducted in an attempt to find out how important the subjective element is in determining the salience of particular semantic distinctions. The first step in answering this question was to devise a means of tapping individual semantic memory organisation. To do this, Repertory Grid Technique, drawn from the field of clinical psychology, was used. A second aim of the present study was to provide a general test of the usefulness of Grid Technique as a means of studying semantic memory organisation. If it could be shown to generate information about basic semantic memory structure, then it could be used to test theories which specify the nature and organisation of particular semantic features. In the next two sections, Repertory Grid Technique and the theory on which it is based are described in more detail.

3.2 Personal Construct Theory

Kelly (1955) proposed his theory of Personal Constructs to account for the behaviour of patients whom he observed in the clinical setting. The Repertory Grid Technique was used as a means of tapping information about how the individual perceived and categorised his experience, which he might otherwise find difficult to articulate, or of which he might even be unaware. Kelly's basic postulate was that each individual evolves for his convenience in understanding the world an interlinked and more or less coherent construct system. A construct is a concept expressing the way in which some things are seen as alike and yet different from others. It is a personal bipolar abstraction, involving a particular basis for considering likenesses and differences, used to structure some aspect of a person's world; it is a tool to allow the discrimination and organisation of events.

According to Kelly, each person sees the world through his own personal construct system. People differ from each other in their construction of events because each of us lives in what is ultimately a unique world, uniquely interpreted and experienced. Just as constructs differ from person to person, so the way constructs are interrelated differs from one individual to another. As the grid is a means of studying individual construct systems, it cannot be used for comparing one individual with another without denying the unique nature of each person's construction. However, its use in the study of semantic memory structure should produce evidence of a fair degree of commonality between subjects, in that they share a common language.

Kelly is not alone in his emphasis on the study of the individual rather than the group. Dissatisfaction with the results obtained and the generalisations that can be made about individuals from nomothetic

methods has often been expressed. It has been demonstrated that only rarely do 'average curves' have the same form as any of the 'individual' curves from which they are derived. As Sidman (1952) and Bakan (1954) have pointed out, one can seldom draw meaningful inferences about individual processes from group statistics.

It is on grounds such as these that the differences between the grid and the semantic differential technique (Osgood et al., 1957) can be made clear. As these techniques are often perceived as similar, it may be useful to outline the differences, as an explanation of why one was used rather than the other. Both are concerned with the interweaving of concepts and scales, to use the semantic differential terminology, the approximately equivalent Kellian terms being 'elements' and 'constructs'. Both are measures of meaning. But the basis of the semantic differential was a nomothetic summing of the data from a large number of subjects, who rated a large number of concepts in terms of a large number of scales. Factor analysis of this data revealed three frequently occurring dimensions: the evaluative, activity and potency factors, and these were taken to be the basic dimensions of a semantic space in terms of which an individual's concepts about particular entities could be plotted. The problem is that these three nomothetic dimensions of meaning are not identifiable when analysed for one person instead of a group (Fransella, 1965).

Bannister and Fransella (1971) report a 'steady retreat' of semantic differential users from the original stand on the three 'universal' dimensions. For example, Warr et al. (1969) now express the belief that 'knowledge of only the average tells us nothing about the dimensions which underlie judgements by a particular individual'. Repertory grid technique, by contrast, is not based on nomothetic

generalisations.

A further difference between the two techniques, and the reason why grid technique was chosen in preference to the semantic differential in the present study, is that the grid allows access to additional information not obtainable from normal use of the latter method. The semantic differential is concerned with the placement of certain concepts in relation to a number of experimenter-determined scales. The grid method similarly allows the examination of elements in semantic space, but here the dimensions (constructs) are subject-generated. It is this generation of constructs which renders the technique so useful to the present study. In the case of the semantic differential the assumption is made that the subscales used to define the three dimensions will adequately express some aspect of the outlook of each new subject. It assumes generality of meaning in the words used to label these scales. The grid involves no assumptions about generality of meaning and subjects are encouraged to supply their own constructs.

As the intention of the present study was to tap subjective aspects of meaning, this approach seemed to offer an empirical means of obtaining such subjective information, which the semantic differential technique patently did not. Kelly would agree with Osgood et al. that we come to share certain meaning dimensions (constructs) because we live in a similar culture, but the implications of these constructs need not be identical because each individual's construct system is unique.

3.3 Repertory Grid Technique

The technique is an interviewing procedure which is defined in its general structure but can be varied within wide limits. The term 'repertory' is used because it taps constructs from the repertoire of

those which the subject uses to structure his world. The term 'grid' applies because numerical record of the comparison of certain elements in terms of certain constructs is made on a matrix of cells i.e. a grid. The elements in question are printed along the top of the grid and the constructs as they are elicited are written down the side. Scores for each element in terms of each construct are entered in the appropriate cells in the matrix. Any entities or concepts may be used as elements, from the names of close friends to the names of abstract qualities like fear or jealousy.

The central aim of the interview is to reveal the individual's particular construct patterning, rather than to relate this patterning to some established normative data. Any number of elements, specially chosen to reflect the particular concern of the interview, may be compared in terms of any number of constructs. The constructs may be supplied by the interviewer or elicited from the subject. The interviewer may supply some constructs to examine a particular topic, but in this case he cannot be sure that the construct is a meaningful one in the subject's repertoire.

Constructs are elicited by presenting the subject with a list of elements and asking him to select three and find some way in which two of them resemble each other and are different from the third. If, for example, the list of elements was composed of ten names of the subject's acquaintances and the subject remarked that he found two of the people interesting while the third was dull, then the construct 'interesting-dull' would be entered on the grid. The use of three of the elements to generate a construct is termed the minimum context form, as Kelly considered this to be the smallest number of elements from which a generalisation could be made regarding similarity

and relevant difference. The expression of the similarity between two of the elements which is stated first is called the emergent pole of the construct and the expression of the difference between these two elements and the third is called the implicit pole.

When one construct has been elicited the subject is asked to compare all the elements in terms of it and the numerical expression of this comparison is recorded in the appropriate cells on the matrix. The numerical terms in which the comparison is expressed are open to choice. A basic assumption underlying the method is that the psychological relationships between any two constructs for a given person are reflected in the statistical association between them, when they are used as judgemental categories. After having compared all the elements in terms of the first construct, the subject selects another set of three elements, including if he wishes some of those he has already used, and generates another construct. The procedure is repeated ideally until the subject's supply of relevant constructs is exhausted.

The elements may be compared in various ways. The simplest way is to sort them into two classes, those characterised by the emergent pole of the construct, and those not. This can be done by placing a tick in the appropriate cell of the grid for all elements characterised by the emergent pole, and leaving a blank for all elements characterised by the implicit pole. This method is rather inaccurate in that it does not distinguish between a cell being left blank because the implicit pole of the construct is true of that element, because the construct does not apply at all to that element, or because the element falls half way between the two poles e.g. an element might be sorted as 'of moderate illumination' on a construct such as 'light-dark'. This method also leaves open the possibility of deriving distorted

estimates of construct relationships if the ticks and voids in the matrix have a lopsided distribution. For example, if only one element was sorted as characterised by the implicit pole of one construct, and only one other element as characterised by the implicit pole of another construct, then there would be a significant correlation between these two constructs because of the unequal distribution of ticks and voids in the matrix.

This difficulty can be overcome by asking subjects to rank the elements, for example from 'most interesting' to 'most dull', or to rate them, a 5- or 7-point scale having been found to be appropriate (Bannister and Mair, 1968). In this way the inapplicability of a construct can be distinguished from the case where its implicit pole applies. Also the subject is given the opportunity of expressing finer distinctions between elements than is possible with an undifferentiated tick or blank.

The completed grid shows an experimentally selected part of a person's construct system, formed by the interaction of all the subject's elements with all his constructs. Even a relatively simple grid provides a large amount of material and reveals quite complicated interactions. Many types of scores have been derived from grids as extensions of basic matching or relationship measures. A full account of the variety of analyses which can be carried out on grid data is given in Bannister and Mair (1968). The methods of cluster analysis and factor analysis were used by the present writer. These procedures are described in the method sections of Experiments 4 and 7, respectively.

The technique is essentially descriptive and interpretation is made subjectively by the interviewer. The question of validity, that is, how meaningful or relevant is the information thus obtained, and the

question of reliability, that is, would the technique generate the same response on different occasions, are difficult ones. The recognised methods which have been developed for assessing reliability and validity are appropriate only for nomothetic tests and cannot be transferred to individually constructed grids. In the present study, empirical tests were made of whether the grids provided information about the basic structure of semantic memory. The rationale for the tests used is discussed in Section 1.4 of the Introduction. Specifically, tests were made of whether or not the grouping of words in terms of similarities between them as expressed on the grids resulted in better learning of these words than for a set of words not grouped in terms of similarity. This approach is based on the finding that words which are related in meaning are easier to learn than words that have no clear relationship to each other. An account of these empirical tests of the validity of grid technique as used in Experiment 4 is given in Chapter 4.

3.4 The Elicitation of Semantic Distinctions Using Repertory

Grid Technique

An exploratory examination of individual semantic memory organisation was conducted using repertory grid technique. The aim was to test the usefulness of grid technique as a means of tapping individual semantic memory organisation.

More specifically, the question was raised of whether the subjects would find the grid a meaningful task when completed under psychological laboratory conditions rather than in the clinical setting. The proposed use of the technique differed from its traditional clinical use in two ways: (i) common concrete nouns were to be used as elements rather than

elements of particular personal relevance to the subject (ii) the grid was to be administered in booklet form, the subject working through the booklet on his own, rather than the subject and experimenter working together in a lengthy interviewing situation, in which the interviewer elicited the constructs from the subject and supervised his grid completion. As grids from a large number of subjects were required the elimination of the time-consuming interviewing procedure would be an advantage.

It was also necessary to show that the semantic distinctions produced as constructs on the grids did represent basic dimensions of semantic memory structure, as the C/NC, S/L and N/M distinctions were assumed to do.

Only if the operational equivalence of constructs with self-generated semantic distinctions were proven could they be used to test the importance of the subjective element in determining the salience of particular semantic distinctions. The experiments reported in Chapter 4 were concerned with this validation of the grid data.

The present experiment focused on whether or not the subjects found grid completion in the novel experimental setting a meaningful task and on whether inspection of the completed grids and of cluster analysis of them appeared to have produced meaningful data.

A pilot experiment was first carried out to find out whether the subjects could generate constructs for a set of nouns which had not been chosen because of some particular personal relevance to the subject. The 30 nouns used as elements were chosen randomly from Roget's Thesaurus. Grids from ten subjects showed that subjects could fulfil the experimental requirements with such a set of elements, subjects supplying from 17 to 24 constructs for the set of elements in question.

This data was cluster analysed in three different ways. The method of cluster analysis used and the construction of dendrograms is described in the Results section of Experiment 4.

The three forms of cluster analysis carried out were:

- (1) Cluster analysis of the element x element similarity matrices for each subject. This showed the clustering of similar elements in terms of their ratings on the constructs.
- (2) Cluster analysis of the construct x construct similarity matrices for each subject. This showed the clustering of similar constructs in terms of how they were used to rate the elements.

and

- (3) Cluster analysis of the element x element similarity matrix summed across subjects. This showed the overall pattern of element clustering across subjects. Although Kelly's emphasis on the uniqueness of the individual subjects' constructions was recognised, this nomothetic analysis was based on the assumption that the individual subjects' understanding of the nouns used as elements would be sufficiently similar to render the cluster analysis meaningful.

The dendrograms presented in Tables 3.1 and 3.2 show the pattern of element and construct clustering produced by cluster analysis of one subject's data.¹ Only sections of the dendrograms are presented to show the emergent patterns of clustering more clearly. In the complete dendrograms the important clusters of highly similar items show less clearly because of additional links with less similar items.

Table 3.3 shows the complete dendrogram produced by cluster analysis of the element x element similarity matrix summed across subjects.

¹ The similarity coefficients shown in the fusion summaries of Tables 3.1 and 3.2 were not scaled to the range 0 to 99, as were those used in later analyses, because the initial computer program used did not necessitate this.

TABLE 3.1: Experiment 4

A section of the dendrogram for subject x showing the clustering for 17 of the 30 elements

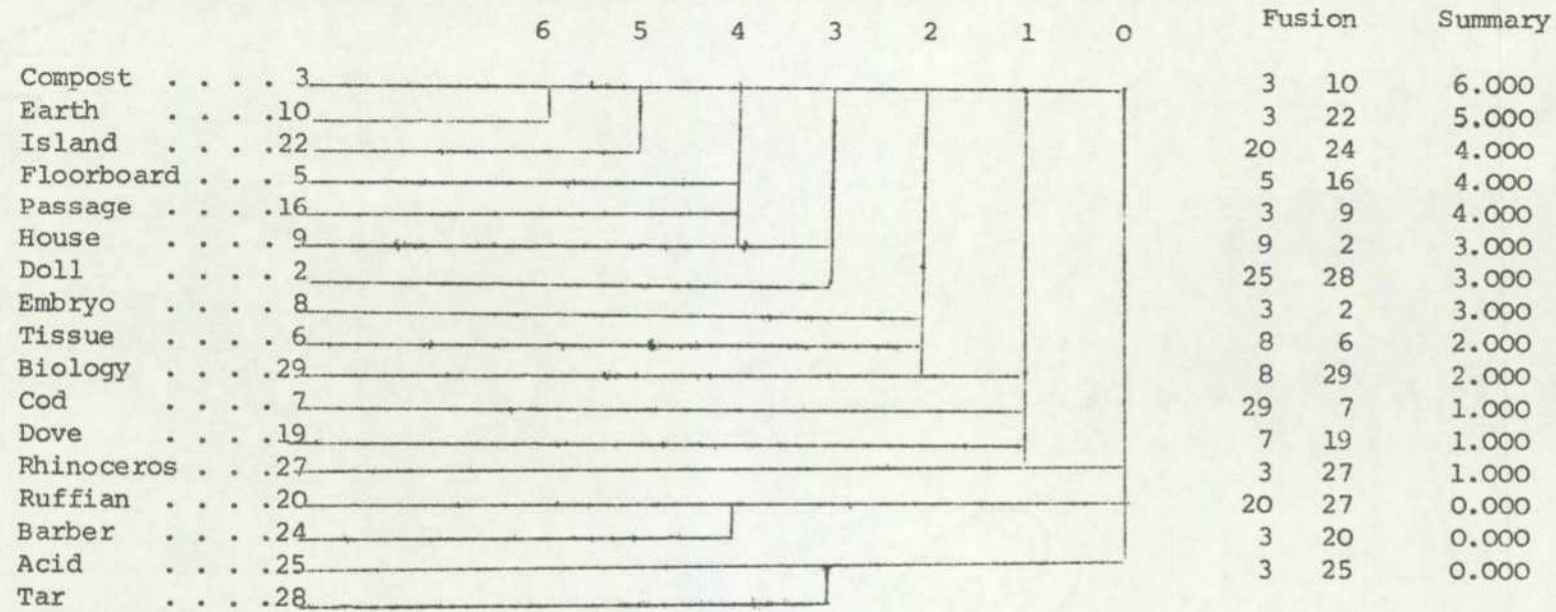


TABLE 3.2: Experiment 4

A section of the dendrogram for subject x showing the clustering for 10 of x's constructs

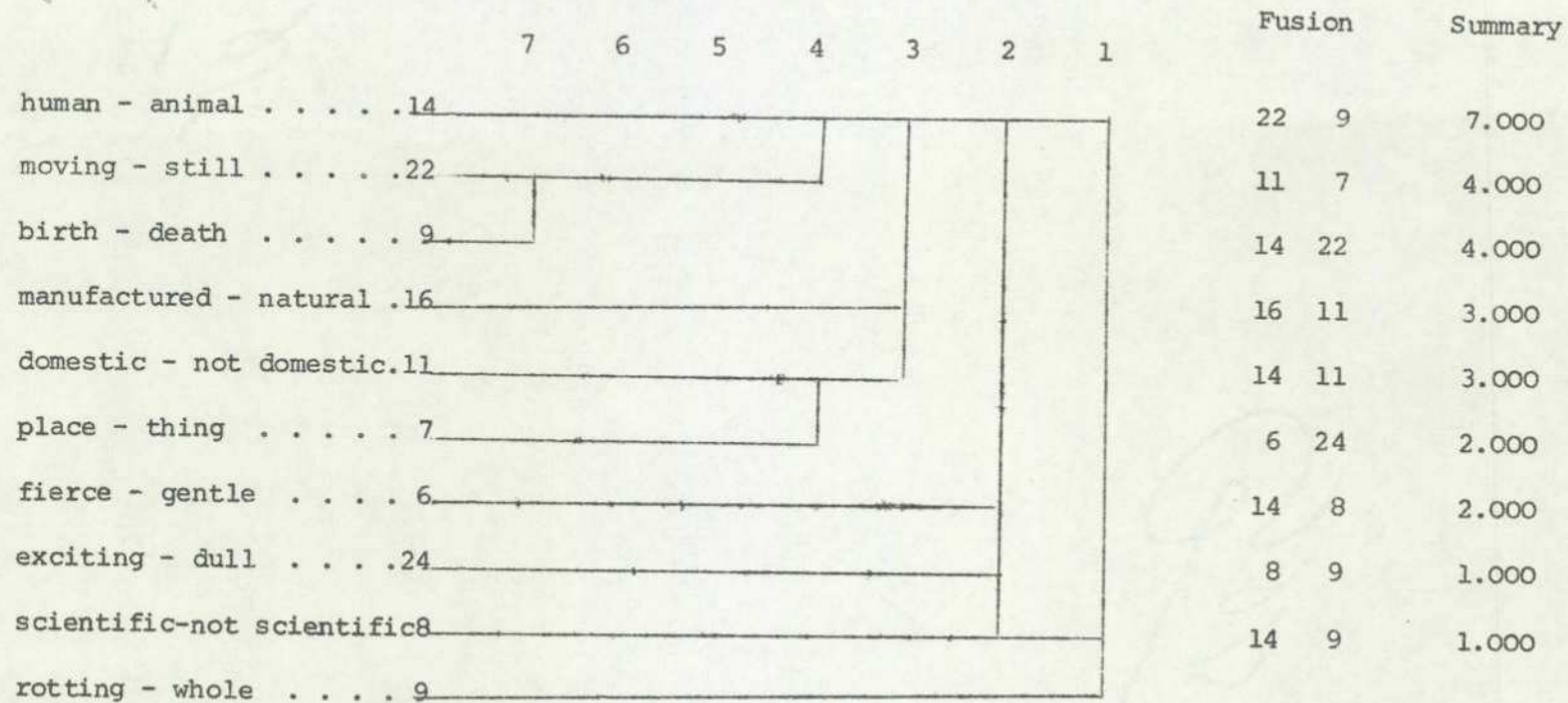
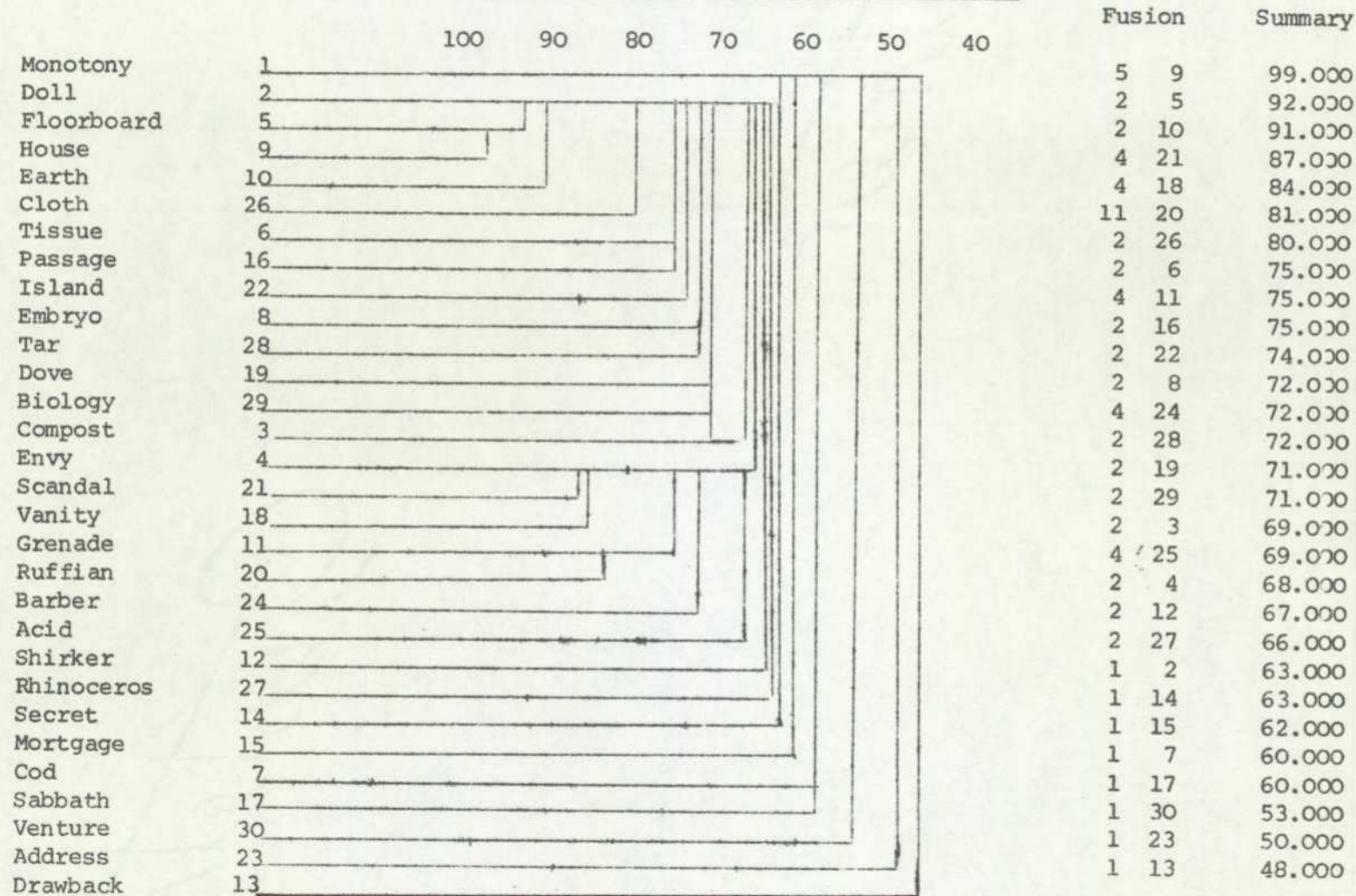


TABLE 3.3: Experiment 4

The complete dendrogram showing the clustering produced by analysis of the element x element similarity matrix summed across subjects



Inspection of the dendrogram presented in Table 3.1 shows certain patterns of element grouping that could be interpreted in terms of semantic distinctions such as 'animate - inanimate' and 'human - animal'. The dendrogram, presented in Table 3.2, which shows the construct clustering for the same subject shows that the 'human - animal' cluster was a central one because it had many links with the other constructs. The constructs 'moving-still' and 'birth-death' which might be subsumed under 'animate-inanimate' formed a tight cluster, indicating a high degree of similarity. The construct 'domestic-not domestic' would seem to account for the clustering of the elements FLOORBOARD PASSAGE HOUSE and DOLL shown in Table 3.1.

The pattern of element clustering for the summed similarity data shown in Table 3.3 was more complex, but the semantic distinction 'Concrete-Abstract' might be proposed to distinguish between two of the major clusters which were apparent.

Comparison of the element dendrogram for the individual subjects such as that shown in Table 3.1, with that for the summed data for all the subjects (Table 3.3) did reveal some commonality as was expected. However, the extent of the difference of the individual subjects' dendrograms from the mean dendrogram suggested that the investigation of individual differences in semantic memory structure was worthwhile.

Individual subjects dendrograms differed considerably in the patterns of clustering which they showed but the tighter clusters were generally comprehensible in terms of basic semantic concepts such as 'Human-Animal'. It seemed that the individual differences in semantic memory structure revealed by repertory grid technique might reside not in the nature of the basic semantic elements, but rather in the way they were organised.

The use of repertory grid technique in the field of semantic memory research therefore seemed to be practicable and to have heuristic potential. A full scale experiment was planned to study the dendrograms produced for a set of common concrete nouns used as elements. These nouns were chosen from the lexical domain represented by the semantic distinctions C/NC, S/L and N/M. It was hoped that their use as elements would encourage the generation of constructs that could then be compared in salience with the semantic distinctions C/NC, S/L and N/M. Experiment 4 was carried out to investigate this possibility.

3.5 Experiment 4

The Generation of Constructs for a Set of Common Concrete Nouns

METHOD

Design

The subjects were required to complete repertory grids in booklet form by supplying 20 constructs for a set of 40 nouns used as elements. Each subject received a unique set of 40 common concrete nouns selected randomly from the Toronto Word Pool (Murdock, 1968). The constructs produced in the pilot experiment had varied in their range of applicability to the elements in question. Some such as 'Moving-Still' had been of wide applicability, while others such as 'Fierce-Gentle' had been of more limited applicability. Because of this variability, all subjects in the full scale experiment were required to produce 20 constructs. This was thought to be a sufficient number to ensure that most of the constructs of widest applicability would be included in the list. Also as the TBR material in the verbal learning studies which were to be

carried out to test the validity of the grid data was to be the elements used on the grids, standardisation of the number of constructs generated by each subject was necessary to control for study time.

A unique set of words was used for each subject to prevent individual word effects such as those found in Experiment 3 from confounding the results of the proposed verbal learning studies. However, this did preclude the construction of a dendrogram for the element similarity data summed across subjects.

The subjects worked on their own to complete the grids, following the procedure from an explanatory leaflet.

Subjects

The subjects were 16 undergraduate and postgraduate volunteers from the City University.

Materials and Procedure

The booklets used for the grid completion had 20 pages. On each page a rectangular box was printed into which the subject slotted a detachable copy of his list of 40 elements. Beneath each element was a small box in which the rating for that element was to be written. Each subject was supplied with an explanatory leaflet which gave examples of how constructs were generated, how they were expressed in bipolar terms and how the numerical comparisons of the 40 elements in terms of the constructs were to be made and entered in the booklets. The method of numerical comparison used was a 5-point rating scale: +2 +1 0 -1 -2, where a rating of +2 was used to indicate that the emergent pole of the construct was strongly applicable to the element in question, where +1 indicated moderate applicability and where 0 indicated that the construct was inapplicable to that element. Similarly the minus ratings indicated that the implicit pole of the construct was

either moderately or strongly applicable.

Constructs were generated by the subject selecting any 3 words from the list of elements and finding an important way in which two of them were similar in meaning and different from the third. Successive sets of three words could include words used in previous sets. This did allow study time for the words to vary somewhat, but if a particular word evoked more than one construct, it seemed unnecessarily inhibiting to prevent the subject from making use of it again. As all 40 words were rated in terms of every construct the overall study times were standardised.

On thinking of a construct, the subject wrote it in the space provided on the lefthand side of a sheet in the booklet. He then placed the list of elements in position and rated all 40 elements in terms of it using the 5-point scale. The rating for each element was placed in the box immediately beneath it.

RESULTS

Cluster Analysis

Cluster analysis was used as the means of discovering the underlying semantic structure present in the large array of data provided by the grids. Specifically, the grouping or clustering of the elements in terms of the similarity between them shown by their ratings on the constructs was to be examined. Only element dendrograms were generated because the TBR material in the verbal-learning studies was to be organised according to the groupings of elements shown by these dendrograms. The number of elements was too great for the underlying pattern of grouping to be evident from simple inspection of the completed grids.

The completed grids were analysed individually for each subject. The first stage of the analysis required the construction of an element x element similarity matrix. This involved calculating a similarity coefficient for each element pair. This was done by adding together the score similarities for all the constructs. The score similarity was defined as: $\text{score similarity} = 4 - (\text{score 1} - \text{score 2})$. The similarity of an element with itself was designated 0. The absolute value of the subtraction of score 1 from score 2, regardless of sign, was used in the similarity equation. For example, if 2 elements received, for 3 constructs, ratings such as:

ELEMENT 1: +2 +2 -1

ELEMENT 2: 0 -2 -1

then the similarity coefficient would equal

$$\begin{aligned} & (4 - (+2 - 0)) + (4 - (+2 - (-2))) + (4 - (-1 - (-1))) \\ &= (4 - 2) + (4 - 4) + (4 - 0) \\ &= 2 + 0 + 4 \\ &= 6 \end{aligned}$$

The similarity coefficients for each element pair were entered in the subject's similarity matrix. These were then scaled to the range 0 to 99, because the card output from the computer program which was to be used was restricted to two digits, while with the data used, the similarity coefficients could be as large as 160. The scaling was done by (1) finding the largest similarity coefficient (M) and (2) multiplying all coefficients by $99/M$.

Cluster analysis, when applied to such an array of similarity measures, constructs a hierarchical system of clustering representations, ranging from one in which each of the elements is represented as a separate cluster to one in which all of the elements are grouped

together as a single cluster.

The technique used for estimating the relative similarity between two clusters and thus fusing them into one at a particular similarity coefficient was that of 'single linkage' or 'nearest neighbour'. According to this method, the similarity between two clusters is defined as the highest single similarity coefficient between two individual items, one from each cluster. A full account of this technique is given under the heading 'Minimum method' in Johnson (1967). A computer program was used to execute the cluster analysis. This was the Hierar hierarchical cluster analysis program, from a suite of Fortran IV programs named Clustan and developed by Wishart (Wishart, 1969).

The Hierar program constructed a hierarchical system of clustering representations, ranging from one in which each of the 40 elements was represented as a separate cluster to one in which all 40 elements were grouped together as a single cluster. It thus had to complete $(N-1)$ fusion steps (i.e. 39 steps where $N = 40$), linking those two clusters which were most similar until all 40 elements were grouped as a single cluster. It summarised the sequence in a 'dendrogram table' from which the associated tree diagram (dendrogram) could be drawn by hand. Part of one subject's dendrogram is shown in Table 3.4. The fusion summary on the right of the page indicates the value of the similarity coefficient at which two particular clusters fused. For each cluster the number of only one of the elements contained in it (the elements having been numbered from 1 to 40) is printed. Thus the line in the fusion summary which states '21 22 99.000' indicates that the cluster represented by element number 21 was fused with the cluster represented by element number 22 at a fusion point

TABLE 3.4: Experiment 4

A section of a typical dendrogram showing the clustering for 20
of the subject's 40 elements

	99	96	95	92	87	Fusion	Summary
Picture	20					1 20	85.000
Stocking	27					1 27	85.000
Receipt	33					1 33	85.000
Crystal	3					1 3	83.000
Bullet	4					1 4	83.000
Teacher	10						
Officer	16					10 16	96.000
Sheriff	32					16 32	99.000
Neighbour	21					10 21	96.000
Widow	22					21 22	99.000
Doctor	25					10 25	95.000
Robber	28					10 28	96.000
Beggar	36					28 36	95.000
Forehead	15					10 15	92.000
Genius	23					10 23	87.000
Collar	8					1 8	82.000
Angel	34					1 34	82.000
Moisture	38					1 38	81.000
Earth	26					1 26	80.000
Perfume	7					1 7	78.000

The 20 elements selected for display here were the 10 which clustered most tightly in terms of semantic similarity and the 10 which clustered least. The number written beside each element is simply the number by which it was represented in the computer program. For clarity only linkages for the 10 elements which clustered tightly are shown.

where the shared similarity was equivalent to a similarity coefficient of 99.000. In fact this is one of the tightest clusters shown, the only other cluster sharing an equally large amount of semantic similarity being that expressed by the fusion of elements 16 and 32.

For programming reasons the similarity coefficients had been scaled to the range 0 to 99, the maximum similarity being denoted by 99. Thus two clusters which met at coefficient '99.000' were the most closely linked, i.e. the most similar, while the final cluster in which all 40 elements were combined as similar would form at one of the lowest similarity coefficients.

DISCUSSION

Inspection of the individual subject's dendrograms, such as that shown in Table 3.4 suggested that semantically meaningful clusters of elements were being produced. Subjects varied considerably in the pattern of element clustering they produced, but in all cases particular groupings could be isolated which were open to interpretation in terms of such general semantic distinctions as Human-Nonhuman and Natural-Manmade.

Among the constructs themselves, a great range of distinctions were expressed. Certain distinctions did seem to recur across subjects, but even identical linguistic expression could not be assumed to indicate that the underlying meaning was the same for different individuals. However, as it had been the aim of the study to tap semantic information of such a personalised nature, considerations as to the generality of the findings were not at this stage necessary. Ultimately, of course, it was hoped to prove that the subjective element in semantic memory did affect the outcome of nomothetic studies of semantic processing such

as those reported in Chapter 2.

Before this could be done, however, it was necessary to find out how permanent and basic was the semantic information which grid technique generated. If the constructs which the subjects supplied were merely superficial transient semantic distinctions created specifically to fulfil the demands of the particular experimental task, then their use in subsequent tasks would not be justified. Alternatively, if it could be shown that the semantic organisation evidenced by the grids could form the basis on which predictions about the operation of semantic memory in other experimental tasks could be made, then assumptions about its relatively permanent and fundamental nature would be more justified. Individual semantic distinctions generated by grid technique could then be used in an experiment in which the importance of the subjective element in semantic processing was to be assessed.

This validation of the information provided by the grids was also necessary if the use of grid technique as an heuristic tool for testing theories of semantic organisation was to be evaluated.

3.6 General Discussion

The novel use of repertory grid technique in the laboratory rather than the clinical setting generated useful data. Although subjects had been required to supply constructs for common concrete nouns rather than for elements of high personal relevance, as is usual in the clinical studies, cluster analysis of the groupings of the nouns in terms of these constructs had produced seemingly meaningful dendrograms.

Subjects did vary in the pattern of element clustering they produced but in all cases particular groupings could be isolated which were open to

interpretation in terms of such general semantic distinctions as 'Human-Nonhuman' and 'Natural-Manmade'. These groupings were similar to those produced by Miller's sorting technique. Miller (1967) claimed that clusters which could be characterised by features such as 'Object-Nonobject', 'Living-Nonliving' and 'Human-Nonhuman' appeared to emerge from an intuitive analysis of his data. The use of Repertory Grid Technique to study semantic memory structure should lead to clusters corresponding to the same basic features being produced as with other techniques such as Miller's if (1) those features are basic and (2) the grid is an adequate tool for tapping semantic memory structure. If the latter point were proven then groupings of words provided by the technique could make possible the reaching of empirical judgements between rival theories of semantic memory structure, which stipulate how words should be related, by showing which theory best fits the empirical facts. The technique would thus permit a semantic analysis of a set of words, similar to that obtained by Miller's sorting technique, but which had the added advantage of allowing for the expression of finer shades of meaning and of subjects supplying their own labels (constructs) for the meaning dimensions they used.

In order to better assess the value of the semantic information generated by grid technique, tests of its predictive power for the outcome of other

tasks involving semantic processing were proposed. More specifically, it was reasoned that if semantic links within a set of words enhance their memorability compared to a set of less related words, then it should be possible to test the grids' validity by selecting sets of semantically related and unrelated words according to the dendrograms for each subject and examining learning performance with these two sets of stimulus material. If learning performance was better with sets which the grids had specified as semantically related then important validation for the grid data would have been obtained.

In Chapter 4 two verbal learning studies are reported in which the validity of the grid data was tested by this method.

Chapter 4

Semantic Memory in Verbal Learning Tasks

Chapter 4

Semantic Memory in Verbal Learning Tasks

4.1 Introduction

A major question regarding the elicitation of constructs using repertory grid technique was whether they were semantic distinctions drawn from the subject's basic semantic memory structure or were they merely task-specific transient semantic distinctions? If the former, then it should be possible to test their salience for the individual in various tasks involving the operation of semantic memory. If the latter, then their usefulness was limited to the task for which they were generated. The relationship between semantic memory organisation and memory performance was discussed in Section 1.4 of the Introduction. The experiments reported in the present Chapter were carried out to test whether or not the organisation of the words according to the constructs generated for them on the grids would predict memory performance when learning these words. Specifically, the question asked was whether or not words grouped as semantically similar according to the constructs would be easier to learn than words which the grids indicated to be less similar in meaning. The basic tenet here of course was that words which were related in meaning were easier to learn than unrelated words. If the semantic information provided by the grids could predict memory performance in this way, then support would be given to the possibility that the grids were tapping basic enduring semantic memory organisation.

Much of the research on organisation in memory has used the free recall learning paradigm as it allows the subjects more scope for imposing their own organisational structure than do most laboratory

memory tasks. However, even within the constraints of traditional verbal learning tasks such as paired-associate learning, the use of perceived semantic links between stimulus words is easily demonstrated. Subjects clearly employ retrieval rules which include such information as the type of material presented and any systematic features which may be used to aid performance. Subjects very rarely make totally inappropriate intrusions (e.g. digits are rarely supplied in experiments using letter sequences) which implies that the subject has perceived such characteristics of the input material and organised his memory accordingly. When material is designed to encourage such retrieval rules (e.g. by making each response rhyme with its respective stimulus, Bower (1970)) recall is enhanced.

Accordingly, in the present chapter, two episodic memory experiments are reported. Experiment 5 was a paired-associate learning situation which compared learning performance for each subject for word pairs shown to be highly linked in terms of similarity of meaning by repertory grid technique with word pairs which were randomly linked. Experiment 6 was a free recall learning situation which compared learning performance for each subject for a set of words which repertory grid technique had shown to be highly semantically related with a set of words which the grid had shown to be of low semantic similarity. If subjects learnt more easily word pairs or sets of words which were of high semantic similarity then this would indicate that the grids were tapping meaningful aspects of semantic memory organisation. The

constructs could then be used as self-generated semantic distinctions to test the importance of the subjective element in determining the salience of particular semantic distinctions.

4.2 Experiment 5

Paired-Associate Learning of Semantically Similar Versus Unrelated Word Pairs

The validity of the information about semantic memory organisation provided by repertory grid technique was tested using the paired-associate learning paradigm. This paradigm involved the presentation of stimulus items in pairs. The learning was tested by presenting the first member of each pair, to which the subject was to respond with the second. The stimulus lists were presented with the pairs in different random orders by means of a rotating memory drum.

The specific aim of the experiment was to test the hypothesis that a set of word pairs in which each of the two words in a pair were closely linked in meaning by the grid would be learnt more quickly than a set of word pairs in which the pairs were composed of minimally linked words according to the grid. The number of pairs correct over 10 trials for two such sets of word pairs was therefore determined.

METHOD

Design

The subjects completed an experimental session in which two lists of 10 word pairs were learnt. For each subject two unique lists of word pairs were drawn up by reference to the individual subjects' repertory grids. One list was composed of pairs shown to be similar in meaning by the grid. The other list was composed of random word pairs. Five different orders of each list were prepared.

The order of presentation of the related and random pair lists was determined by an ABBA design. Half the subjects began with the related list and half with the random list.

Materials

The word x word similarity matrix drawn up from the raw grid data prior to cluster analysis provided a simple means of selecting the ten most similar word pairs out of the total of 40 words used as elements on the grid. The similarity matrix was constructed by adding, for each word pair, the number of times that the two words received the same rating on a particular construct. The ten most similar pairs were those which had the highest score. These formed what were termed the High list. The Low list was formed by randomly combining the remaining 20 words into ten pairs. A High and a Low list was selected individually for each subject. A standard memory drum was used to present the lists of pairs.

Subjects

The subjects were the 16 undergraduate and postgraduate members of The City University who had completed the repertory grids in Experiment 4.

Procedure

The subjects were tested individually. The time lapse between grid completion and the paired-associate learning task was about four weeks. This length of time was thought to be sufficient for the actual grid completion to have been forgotten. Therefore if the grid data did predict the paired-associate learning performance this must be because it had reflected permanent semantic memory organisation.

The nature of the experiment was explained to the subject and he performed two practice trials in which the stimuli were lists of letter-

digit pairs. The experimental trials were then conducted. The pairs were presented at a 1 sec rate using the memory drum. The first item of the pair appeared first on its own for 1 sec and the complete pair was presented for a further second. The subject was required to respond with the second item during the period in which the first item was displayed alone. At the end of each list an arrow appeared. This signalled to the subject that he should commence the distractor task. This was a shadowing task in which the subject had to repeat digits which a recorded voice recited at a 2 per second rate. The distractor task was necessary to prevent rehearsal of the last presented word pair while the experimenter changed the list.

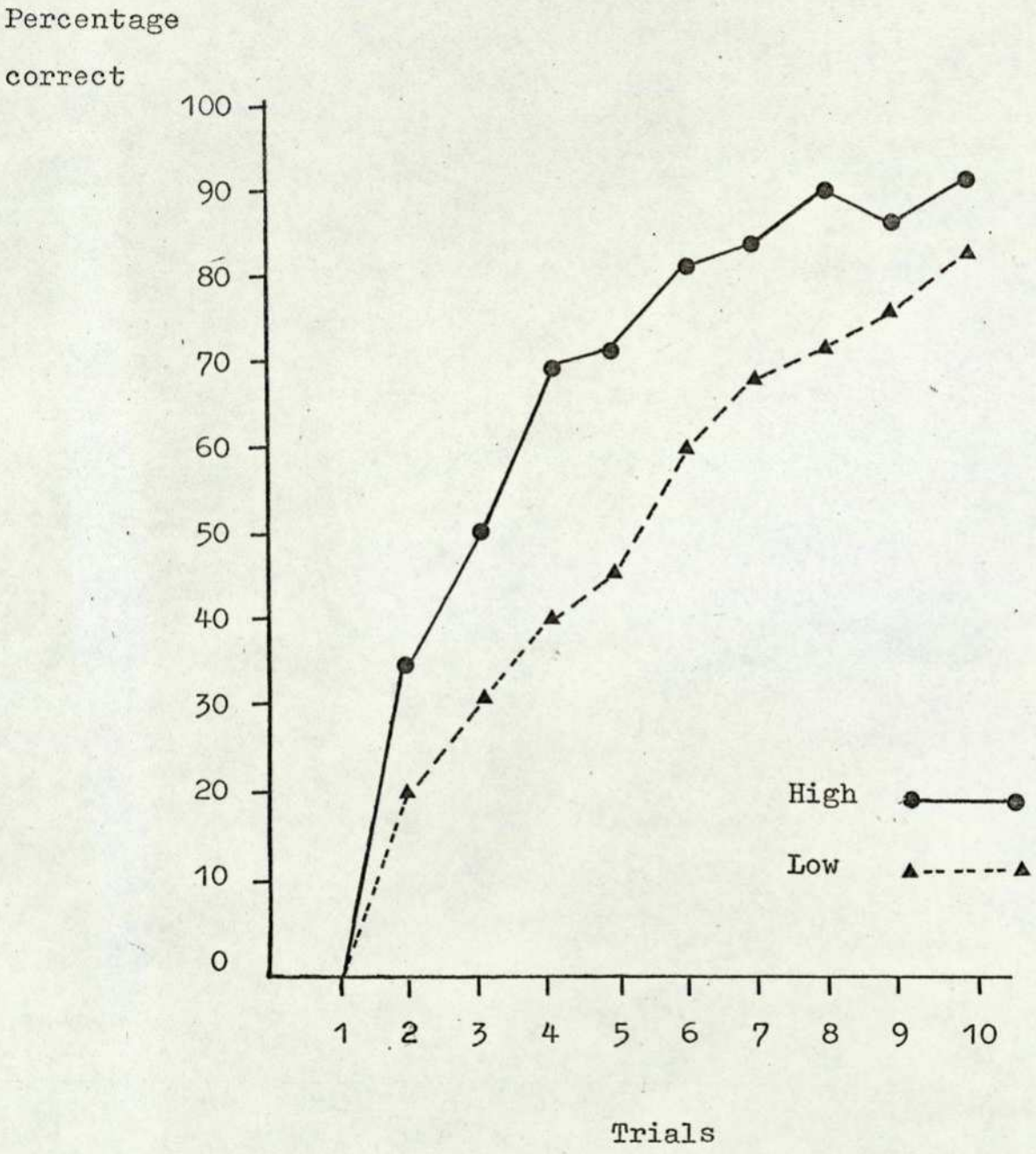
A record was kept of the number of pairs correct over ten trials for both the Low and the High lists.

RESULTS

The data of principal interest concerned the subjects' learning performance with the High and Low lists of word pairs, the prediction being that learning performance should be better with the High list. Figure 4.1 shows the learning curves for the two sets of material, the data from the 16 subjects having been pooled. Learning performance for the High list was clearly superior to that for the Low list. The recall scores are detailed in Table A.2 in the Appendix. An analysis of variance on the number of pairs correct per trial yielded a reliable difference between conditions, $F(1,15) = 12.57, p < .01$. As indicated by the learning curves, the trials effect was highly significant, $F(9,135) = 105.20, p < 0.001$. There was also a reliable trials x conditions interaction effect, $F(9,135) = 3.39, p < 0.01$, attributable to the fact that the subjects learnt the High list more rapidly.

FIGURE 4.1 : Experiment 5

Paired-associate learning curves for the High and Low stimulus lists.



The analysis is summarised in Table A.3 in the Appendix.

The experiment provided good evidence that the High list of semantically related word pairs were learnt better than the Low list of unrelated word pairs. Given this result, it seemed possible that within the High list, the word pairs shown to be most similar in meaning by the grids would be learnt more quickly than pairs which were less similar in meaning. Specifically, it was possible that there would be a positive correlation between recall scores for each word pair and the similarity coefficient for that pair in the similarity matrix.

Accordingly a subsidiary analysis was carried out. The ten pairs of words in the High list were ranked in terms of degree of similarity from most similar to least similar. This was done separately for each subject. A Kendall rank order correlation coefficient, ^{however,} showed no correlation between this and the rank order of the pairs in terms of the number of times the subject remembered them correctly.

DISCUSSION

The finding that the semantically similar word-pairs were reliably easier to learn than the random word pairs provided strong support for the hypothesis that the grids were tapping information about semantic memory organisation that was sufficiently permanent and basic to predict performance in a paired-associate learning task.

The subsidiary analysis in which the possibility was tested that word pairs with a higher similarity score would be learnt more easily than pairs with a lower similarity score did not produce a positive result. However it was evident from an inspection of the similarity

matrices that there were a large number of similarity links between words in the High set other than the specific pair linkage. This number of interlinkages other than the specific pair linkages would tend to obscure the emergence of a clear correlational effect.

As the number of interlinkages between words in the High list was greater than in the Low list, and as these links would interfere with the learning of the specific pairs, the finding that these pairs were better learnt than the random pairs appeared particularly significant.

The results of the paired-associate learning study provided strong support for the validity of the grid data as reflecting basic semantic memory organisation. Further support was next sought using a different verbal learning paradigm. Accordingly, Experiment 6 reports a test of the grid's validity using the free recall learning paradigm.

4.3 Experiment 6

The free recall learning of semantically similar and dissimilar sets of words

The aim of Experiment 6 was to test the hypothesis that words which were grouped as similar in meaning by cluster analysis of repertory grid data would be learnt more quickly than words which cluster analysis did not indicate to be similar in meaning. The learning paradigm used was that of multi-trial free recall. A further prediction made was that there would be differences in SO (subjective organisation) in the recall protocols between the semantically similar and dissimilar sets of words. The rationale for this was that if the subjects made use of semantic memory in organising the TBR material, then organisation would proceed to a greater extent for the

semantically similar set because more groupings in terms of semantic similarity were possible.

METHOD

Design

Each subject took part in two free recall learning sessions, the one succeeding the other after a few minutes pause. One of the sessions involved the learning of words grouped as highly similar in meaning (the High condition) while the other involved the learning of words grouped because of minimal similarities in meaning (the Low condition). The order in which the two sessions were performed was randomised across subjects, with the restriction that each condition should occur first and last an equal number of times.

Materials

The words for the two conditions were selected for each individual subject by perusal of his element dendrogram produced by cluster analysis of the repertory grid data. For the high condition, the cluster containing the ten most similar words was selected, while for the low condition ten words which linked as little as possible with each other were chosen. Thus each subject had a unique set of high and low words. Examples of sets of High condition words for three subjects were:

(1)	TICKET	(2)	SHERIFF	(3)	COLLAR
	CHIMNEY		BEGGAR		STOCKING
	COMPASS		TEACHER		SADDLE
	TUNNEL		GENIUS		NEEDLE
	CELLAR		WIDOW		VELVET
	STATION		NEIGHBOUR		LEATHER
	MACHINE		FOREHEAD		OVEN
	PENNY		OFFICER		BASIN
	RECEIPT		ROBBER		MIRROR
	PROGRAMME		DOCTOR		KETTLE

Examples of sets of Low condition words for three subjects were:

(1) ANGEL	(2) TRAFFIC	(3) SANDWICH
FOREHEAD	MACHINE	BANNER
INFANT	POISON	MUSCLE
PLANET	PATTERN	VOYAGE
NATIVE	FUNERAL	INSECT
FUNERAL	MONSTER	TUNNEL
MOUNTAIN	BUTTER	SUNSHINE
KEEPER	THUNDER	BASKET
BIBLE	INSECT	WRINKLE
FOREST	CANAL	CRYSTAL

For each condition a booklet was compiled in which the ten words were presented in ten different random orders. The words were presented in a vertical array and were printed in upper case letters. Each presentation list was followed in the booklet by a sheet on which subjects were to write the answers to an arithmetic distractor task and then by a sheet on which the subject was to free recall as many of the words from the presentation list as he could remember.

Subjects

The subjects were the 16 undergraduate and postgraduate members of The City University who had completed the repertory grids in Experiment 4, and the paired-associate learning task in Experiment 5.

Procedure

The subjects were tested in three group sessions. The time lapse between the paired-associate learning task and the free recall learning task was three months. Each ^{subject} received two practice trials with a list of monosyllabic common concrete nouns, to familiarise them with the procedure. The nature of the experiment was explained to the subjects and they were assured that it was not a test of how good their memory was relative to others in the group. The experimental format was as follows:

- (1) The subject turned to the first presentation list and memorised the words for 15 sec.

(2) The subject turned to the next page in the booklet and wrote down the answers to arithmetic problems such as $'97 + 8 = \quad'$ for 30 sec. The problems were presented at a constant rate by means of a Cassette tape recorder. Subjects were requested to work as quickly and accurately on this task as possible and not to rehearse the words. This distractor task was necessary to eliminate recency effects and to prevent ceiling effects in memorisation.

(3) The subject turned to the next page in the booklet and wrote down as many of the words as he could remember in any order. 30sec. was allowed for this.

When completed, the first booklet was collected in silence. The second booklet was then distributed and the procedure repeated.

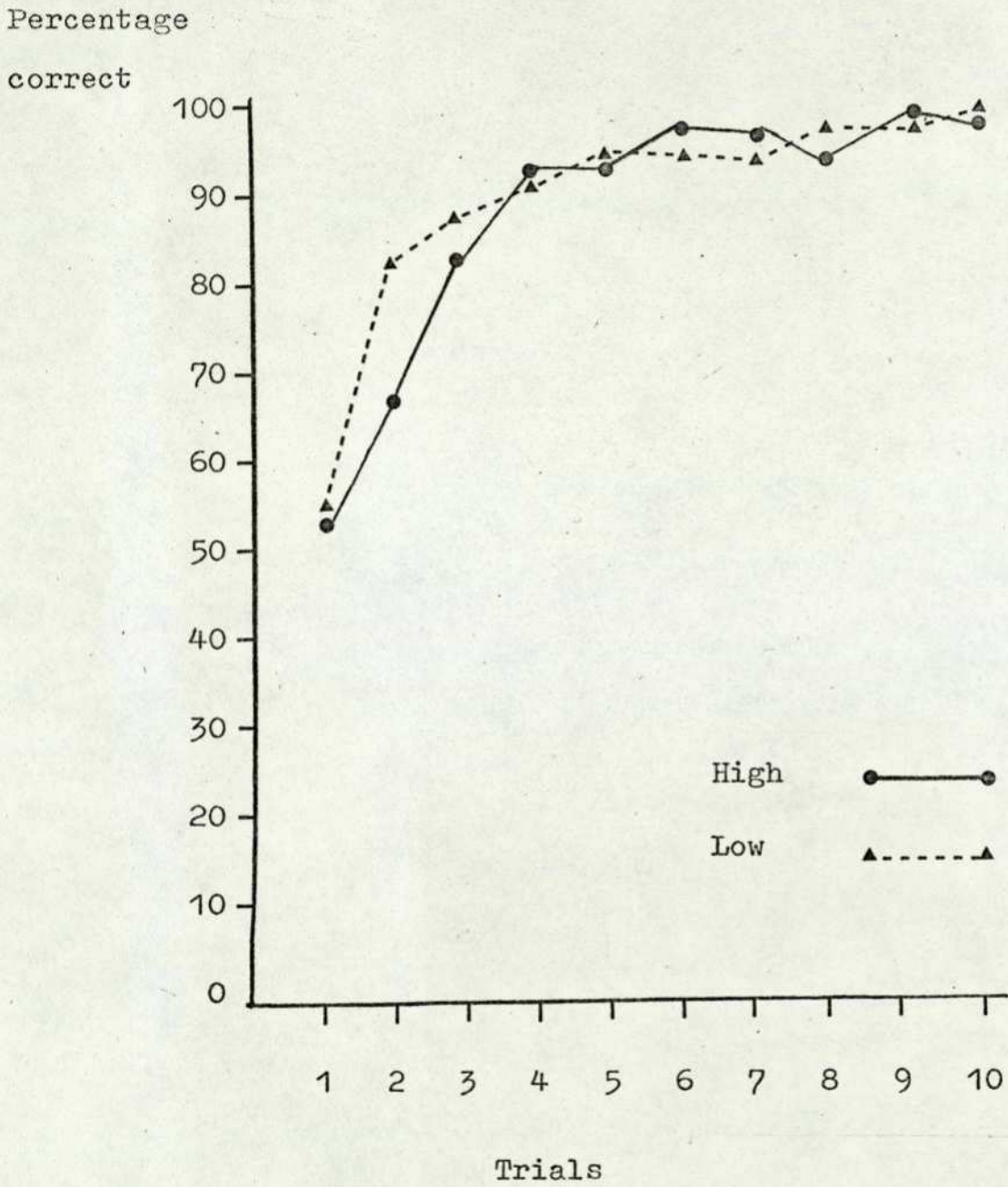
After the last recall in the second session, the subjects were required to indicate by bracketing any clusters of words which they had recalled together. Their first booklet was returned to them so that they could write out a similar specification for the words they had learnt first. This was to provide an indication of the organisational strategies in use; clustering on the basis of semantic, acoustic and phonological similarity might be identified.

RESULTS

The mean number of words recalled correctly per trial for the Low and High conditions is shown in Table A.4 in the Appendix. These data are plotted graphically in Figure 4.2. No difference in recall between the two conditions was apparent, although for both there was a marked increase in recall over trials. An analysis of variance on the recall scores for individual subjects showed that the difference

FIGURE 4.2 : Experiment 6

The free recall learning curves for the High and Low stimulus sets.



between the two conditions was nonsignificant, $F(1,15) = 0.64$. However, as indicated by the learning curves, there was a significant trials effect, $F(9, 135) = 77.75$, $p < 0.001$. The conditions $\times$ trials interaction was nonsignificant, $F(9, 135) = 1.12$. The analysis is summarised in Table A.5 in the Appendix. Thus the hypothesis that recall scores would be better for the High condition because of greater inter-word semantic similarities was not supported by the recall data.

In an attempt to find out if there was any difference between the conditions in the amount of subjective organisation used, ITR scores were calculated. The measure used was Fagan's ratio measure (Fagan, 1968), which he termed 'sequential consistency' (SC). Its basic advantage over the other ITR measures was that in being a ratio measure it allowed the maximum amount of clustering possible to be taken into account. The SC score is expressed as a percentage of the maximum possible score. The equation used to calculate Fagan's ratio measure is:

$$ITR_{(SC)} = \frac{O(ITR)}{c-1} - \frac{c}{hk}$$

where $O(ITR)$ = the observed number of intertrial repetitions

h = the number of items recalled on trial t

k = the number of items recalled on trial $t + 1$

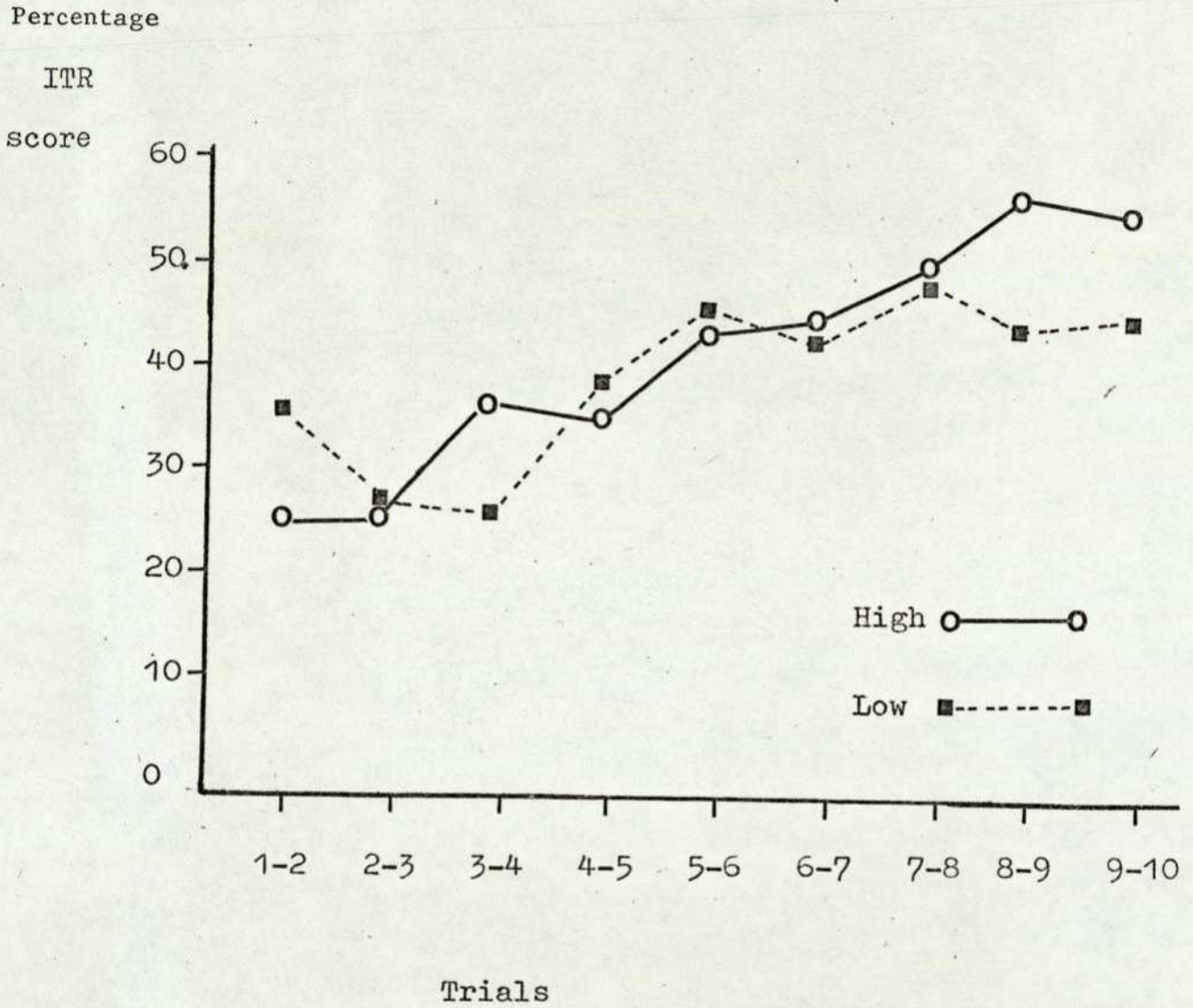
c = the number of items common to the two recalls

This was computed bidirectionally so that grouping effects in which on successive trials two words appeared together, but not in the same order, could also be scored.

The mean ITR scores for the two conditions are plotted graphically in Figure 4.3, and are presented in table form in Table A.6 in the

FIGURE 4.3 : Experiment 6

The increase in ITR scores over trials for the two conditions.



Appendix. As can be seen from the figure, the ITR scores did increase over trials, but there was no clear difference between the conditions. However, for trials 8-9 and 9-10 the scores for the High condition were considerably greater than those for the Low condition.

An analysis of variance carried out on individual subject's ITR scores showed that the difference between the two conditions was nonsignificant, $F(1,15) = 0.19$. As indicated by the increase in ITR scores over trials shown in Figure 4.3, there was a significant trials effect, $F(8, 120) = 3.97$ $p < 0.001$. The conditions $\times$ trials interaction was nonsignificant, $F(8, 120) = 1.05$. The analysis is summarised in Table A.7 in the Appendix.

Thus the ITR scores provided no reliable support for the assumption that subjective organisation would be greater in the High condition because the presence of greater inter-word semantic similarities would facilitate subjective organisation strategies based on semantic similarities in the TBR material. However, Figure 4.3 showed that the highest ITR scores were achieved in the High condition for trials 8-10. This suggests that subjective organisation proceeded to a greater degree for the High condition. Also the increase in ITR scores over trials confirmed that subjects were employing organisational strategies in their recall.

DISCUSSION

The results of Experiment 6 were somewhat surprising. They failed to show a difference between the High and Low word sets either in terms of amount recalled or in terms of degree of subjective organisation as instanced by Fagan's ratio measure.

One possible reason for this, was that the assumption that

subjects would use solely semantic organisation in organising the TBR material was invalid. It was by no means certain that subjects were ordering the TBR material by purely semantic strategies. Reference to the statements of which words were remembered together, provided by subjects at the end of the experiment, showed that acoustic and visual as well as semantic similarity was being used. Bower (1970) has shown that subjects can adopt phonemic coding when they consider it will aid performance, subjects learning rhyming word pairs more rapidly than phonemically unrelated pairs. Similarly Baddeley and Levy (1971) have shown that phonemic coding occurs when subjects are given unrelated words and insufficient time to produce a semantic link.

As the number of TBR items only just exceeded the span of immediate memory, such coding may well have been adequate to allow total recall. It was unfortunate that only 10 words per condition could be used, but this was necessary in order to clearly differentiate between the semantically linked and unlinked words drawn from the dendrograms.

A second possibility, which is apparent by inspection of Figure 4.2, was that differences between the High and Low conditions were masked by a ceiling effect. (Performance reached 90% plus at trial 4). The ceiling effect could have been overcome by the use of more words in each condition. However, as stated above, this was not possible. A very short amount of study time was allowed and a difficult interpolated task used, but this was not sufficient to prevent the ceiling being reached.

Although the ITR scores showed no reliable difference between conditions, the high scores for the last trials (trials 8-10) for the High condition did suggest that a higher degree of subjective organisation was possible for this condition.

It would seem that as a test of the predictive ability of the

semantic memory organisation illustrated by repertory grid technique, the present experiment was inadequate for two main reasons. First, it seemed possible that total recall could be achieved with little recourse to subjective organisation of the stimulus list; and second, even for subjects who did organise the input as a means of enhancing recall, it was by no means certain that they relied on semantic similarity to do this. In view of these inadequacies, the conclusion that the grids were not tapping basic semantic memory organisation did not seem necessary.

4.4 General Discussion

The basic question which Experiments 5 and 6 sought to answer, was whether the semantic distinctions generated by repertory grid technique were drawn from the subject's permanent semantic memory system or whether they were merely transient task-specific distinctions. It was postulated that if the former were the case then words grouped according to the distinctions should be more easily learnt than words chosen so that the semantic similarities between them were minimal.

The paired-associate learning experiment provided strong support for this hypothesis: the recall scores for word-pairs selected on the basis of high semantic similarity were markedly better than those for pairs formed by the random selection of words.

The results of the free recall learning experiment did not support the hypothesis. However, the experiment seemed to have been unsuitable for testing the hypothesis on two accounts: (i) it was not possible to use a large enough number of words per condition to prevent ceiling effects from being reached and (ii) the subjects organised their recall by organisational strategies other than those based on semantic similarity.

The subsidiary analysis carried out on the recall data for the High list of word-pairs in Experiment 5 did not show that degree of intra-pair semantic similarity correlated positively with amount recalled. This negative result mirrored the negative results found in a study which attempted to influence recall scores by varying the degree of ^{within-list} intra-pair free associative strength in the paired-associate learning situation. Wicklund, Palermo and Jenkins (1964) reported no differences in the recall scores for paired-associates in which the free associative strengths of the stimuli and responses were varied (by reference to word association norms) even when lists as long as 35 pairs were used. Stimulus-response pairs having associative strengths as weak as 0.2% according to the Russell and Jenkins norms (1954) were learned as rapidly as pairs with an associative strength of 69%, although all associated pairs were learned more rapidly than the ^{list of} control pairs. This suggested that a possible interpretation of the negative result of the subsidiary analysis was not that the differences in degree of similarity between the word pairs specified by the grids were artefactual, but that differences in degree of similarity as with degree of associative strength did not produce differences in ease of learning.

The possibility that the negative result of the free recall learning experiment was due to subjects organising their recall by strategies other than those based on semantic similarity suggested that this was another area in which the operation of individual differences might be important. Mandler (1967) and Tulving (1968) present some examples of subjective clusters which occurred consistently in free recall output in which the basis of organisation was not at all obvious, and Shuell (1969) has stressed that a whole variety of factors

must be assumed to determine which basis of organisation the subject will use in studies of this nature. Subjects may select a particular one on an idiosyncratic basis. As Shapiro (1970, p. 82) remarked ' . . while the term 'subjective organisation' had traditionally connoted subject-generated organisational units', perhaps the connotation of 'subjective' as being more or less idiosyncratic is also appropriate'.

The general conclusion drawn from the verbal learning studies reported in the present chapter was that the grids were tapping permanent and basic semantic memory organisation. This conclusion was justified by the strong support for the hypothesis provided by the paired-associate learning experiment and by the attribution of the negative result of the free recall learning experiment to its evident inadequacies as a means of testing the hypothesis.

This conclusion suggested that the use of the grid could be developed as a tool for gaining information about semantic memory organisation which could be used to test proposed theories of semantic memory structure.

The conclusion also justified the further use of repertory grid technique to investigate the importance of the subjective element in determining the salience of particular semantic distinctions. To recap, the use of grid technique was introduced in an attempt to find a means of tapping subjective semantic memory organisation, because it was felt that the failure of Experiments 1, 2 and 3 to obtain reliable response trends for Johnson-Laird's three semantic distinctions might in part be due to their lack of subjective salience. An experiment in

which semantic distinctions of assured subjective salience (generated by grid technique) were compared in effectiveness with Johnson-Laird's three distinctions was therefore planned. Chapter 5 presents an account of the execution of such an experiment, Experiment 7.

Chapter 5

The Importance of the Subjective Element in the
Saliency of Semantic Distinctions

Chapter 5

The Importance of the Subjective Element in the Saliency of Semantic Distinctions

5.1 Introduction

In considering possible reasons for the failure of the experiments reported in Chapter 2 to fully support predictions based on the age-of-acquisition index, it was suggested that individual differences in the saliency of particular distinctions might have prevented the emergence of the predicted effect. The structural organisation of the semantic distinctions in terms of saliency for certain subjects might differ from that predicted by the age-of-acquisition index. Particular past experiences with the semantic distinctions might have resulted in increased or decreased saliency in a direction not predicted by the index.

In order to test this possibility a comparison was made of the effectiveness of three subject-generated versus the three experimenter determined semantic distinctions in promoting incidental learning. Subjects' free recall performance would be compared following the incidental learning of a set of words for which the appropriate poles of either three self-generated or the three experimenter determined semantic distinctions, 'Consumable-Nonconsumable' (C/NC), 'Solid-Liquid' (S/L) and 'Natural-Manmade' (N/M), had been assigned as the orienting task.

The rationale for the use of this criterion of saliency, i.e. amount of incidental learning promoted, was discussed in Section 1.4.d in the Introduction. The basic argument derived from the 'levels of processing' approach to memory research (Craik and Lockhart, 1972). By this approach the strength or persistence of the memory trace is thought to be a function of the 'depth' to which the stimulus is

processed or analysed. Semantic distinctions of greater salience should result in a greater 'depth' of processing and hence better recall. Processing the words in terms of self-generated distinctions rather than experimenter-determined distinctions should allow a greater depth of processing to take place because in addition to the general semantic features of the words to which the subject would have cause to attend, he would also attend to extra features of the words due to the extra personalised aspects of meaning inherent in his self-generated distinctions.

A control group was included in the experiment to provide a baseline for amount recalled. According to the 'levels of processing' approach, attention to the phonological attributes of the stimulus results in a less deep level of processing than does attention to its semantic attributes. A control group which would process the words in terms of three phonological distinctions, such as whether or not the word contained a particular letter, was therefore included in the experimental design. It was predicted that amount recalled would be at a minimum for the control group at an intermediate level for the group processing the words in terms of the experimenter-determined semantic distinctions and at a maximum for the group using self-generated semantic distinctions.

The first step in the execution of Experiment 7, which was designed to test this prediction, was to find a method of obtaining salient semantic distinctions from individual subjects. The use of Repertory Grid Technique as a means of obtaining subject-generated distinctions was validated

in Chapter 4. It was proposed that the three most salient

distinctions for each subject could be found by factor analysis of the repertory grids. If a factor analysis program were set to draw out the three factors which accounted for the greatest amount of variance in the grid data, then the constructs which were most representative of each of the three factors could be isolated as the three most salient. Salience was therefore operationally defined in terms of the amount of variance accounted for in the data by a particular construct.

The learning paradigm chosen to test the amount of recall consequent upon different types of processing was the incidental learning paradigm. Traditionally incidental learning experiments have contrasted the characteristics of incidental learning with those of intentional learning. However, with the upsurge of interest in organisational aspects of memory, incidental learning has become an important technique in its own right, as a means of investigating coding processes. This is because it permits a greater control of the coding processes which the subject employs. When he is aware that he will have to recall, the subject may consciously employ a number of strategies. These confound any attempt to assess the differential effectiveness of particular types of experimenter-determined processing. The incidental learning situation, by contrast, permits the effect of particular modes of processing to be assessed by means of particular orienting tasks. The subject is directed to employ only a particular mode of processing in order to complete the incidental learning task. In view of the control over processing strategies which this technique allows it seemed appropriate to the aims of the present study.

Hyde and Jenkins (1969, 1973) have used the technique to show that orienting tasks which direct the subject's attention to semantic

attributes result in better subsequent recall than those which direct attention to structural features such as grammatical form class. In a similar way the present experiment hoped to show that attention to semantic characteristics would result in better subsequent recall than attention to phonological characteristics and more specifically, that processing in terms of self-generated semantic distinctions would result in better recall than processing in terms of semantic distinctions of no particular personal relevance.

To further enhance the differential effects of the different modes of processing, it was decided to cue the recall session with the three distinctions used during the incidental learning session. If processing in terms of the particular distinctions had caused particular features of the words to be encoded, then the presence of these distinctions during recall might aid retrieval.

Experiment 7

5.2 The Amount of Recall as a Function of Processing Words in Terms of Phonological, General Semantic or Self-Generated Semantic Distinctions

The aim of the experiment was to test the prediction that there would be increasing amounts recalled from the group processing in terms of phonological distinctions, through the group processing in terms of the semantic distinctions used to test the age-of-acquisition index (i.e. C/NC, S/L, and N/M), to a maximum in the group using self-generated

semantic distinctions.

METHOD

Design

The experiment involved three sessions. The first session was the completion of repertory grids by all subjects. All subjects performed this task, even though only one third of them were to use self-generated semantic distinctions, in order to equate study time for the words for all groups.

The second session was the incidental learning task in which the subjects were required to write down the stimulus words as they were read out by the Experimenter and to write beside each word which poles of a set of three bipolar distinctions applied to it. Subjects were randomly assigned to one of three groups: the group which were to use the set of three phonological distinctions, the group which were to use the set of three semantic distinctions, C/N, S/L and N/M, and the group which were to use a set of three self-generated semantic distinctions. These groups will be referred to as groups P, G and S respectively, for brevity. The stimulus material in this session was the set of words which had been used as elements on the grid. A two week period elapsed between grid completion and the incidental learning session in order to reduce possible ceiling effects in recall.

The third session was a single trial free recall session and it followed immediately after the second session. Subjects were asked to write down as many of the words used as stimuli in the incidental learning task as they could remember. They were cued with the three distinctions which they had just been using, Group P receiving an answer

sheet headed with the three phonological distinctions, Group G one headed with the three experimenter-determined semantic distinctions and Group S one headed with the individual subject's three most salient semantic distinctions.

Subjects

The subjects were undergraduate volunteers from The City University. 69 subjects completed the repertory grids, but only 44 of these were available two weeks later to complete the second and third sessions. The subjects were tested in three groups, with equal allocation of subjects to Groups P, S and G in each group.

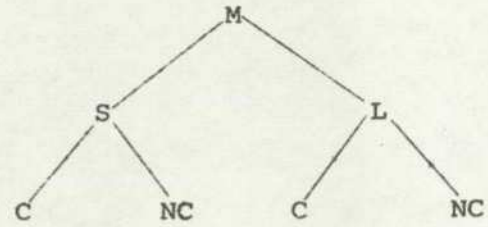
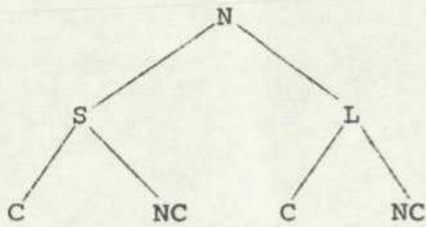
Procedure

In the first session, subjects were presented with blank repertory grids, along the top of which were printed the 24 words to be used as elements. These are shown in Table 5.1. The words chosen were common concrete nouns which broke down evenly into the eight categories formed by the interaction of the C/NC, S/L and N/M distinctions.

The subjects were required to supply twelve constructs for these elements by selecting sets of any three elements from the list and finding ways in which two of the elements were similar in meaning and different from the third. No restrictions were placed on the number of times an element could be used in a set of three. The word or phrase which expressed the similarity was written in the first blank row of the grid in the appropriately labelled column and served as the first pole of the bipolar construct. The word or phrase which expressed the difference was written next to it and served as the second pole of the construct. After producing a construct, subjects rated all 24 elements in terms of it, before proceeding to generate another. To do this they used a 4-point rating scale: +2 +1 -1 -2, where

TABLE 5.1: Experiment 7

The 24 words used as elements, displayed in terms of
the three semantic distinctions used by Group G



carrot	granite	water	oil	biscuit	jacket	sherry	bleach
almond	cotton	honey	sweat	cheese	lorry	soup	shampoo
trout	fur	milk	brine	pill	rifle	vinegar	paint

+2 indicated that the first-named pole of the construct was strongly applicable to the element under consideration, where +1 indicated that the first-named pole was moderately applicable, and where minus numbers indicated that the second-named pole was either moderately or strongly applicable. No zero point was used in an attempt to equate study time for each word by preventing subjects dismissing words from consideration on certain constructs by assigning to them zeros of nonapplicability. Subjects were requested not to leave any cells blank. No time limit was placed on the task.

The completed grids were subjected to a factor analysis program set to draw out for each subject the three major factors into which the 12 constructs could be grouped, in terms of similarity of applicability to the 24 elements. For each factor the construct with the highest loading was chosen to represent it. Thus for each subject the three constructs which accounted for the greatest amount of variance on his grid were isolated. These could be considered the three constructs with the highest degree of semantic salience for the subject.

The subjects were then randomly assigned to one of the three groups for the incidental learning session. For subjects who were assigned to Group S, individual task sheets were made out, at the top of which were written their three self-generated semantic distinctions. Similar sheets were printed for the other two groups.

The sheets for Group P were headed with the three phonological distinctions: (Has O - No O), (Has E - No E), (Has T - No T), The sheets for Group G were headed with the three semantic distinctions C/NC, S/L and N/M. These were written in full and their definitions were provided. Table 5.2 shows the distinctions used by the three groups.

TABLE 5.2: Experiment 7

The bipolar distinctions used by the three groups. For
Group S an example of one subject's set of
distinctions is given

<u>Group P</u>	<u>Group G</u>	<u>Group S</u>
Phonological Distinctions	General Semantic Distinctions	Self-Generated Semantic Distinctions
1) Has T - No T	1) Consumable - Nonconsumable	1) Viscous - Nonviscous
2) Has O - No O	2) Solid - Liquid	2) Manufactured - Natural
3) Has E - No E	3) Natural - Manmade	3) Tasty - Not Tasty

TABLE 5.3: Experiment 7

The mean recall scores for Groups P, G and S

(N = 44)

	<u>Group P</u>	<u>Group G</u>	<u>Group S</u>
	Phonological Distinctions	General Semantic Distinctions	Self-Generated Semantic Distinctions
	N = 14	N = 15	N = 15
Mean	13.57	16.40	18.00
Standard Deviation	3.21	2.93	2.56

TABLE 5.4: Experiment 7

A content analysis of the three constructs chosen for
each subject

<u>Distinction</u>	<u>Number of times present in top three constructs</u>
Solid/Liquid	56
Edible/Inedible	44
Manufactured/Natural	34
Like/Dislike	5
Coloured/Not Coloured	5
Animal/Vegetable or Mineral	5
Expensive/Cheap	4
Hot/Cold	4
Dangerous/Safe	4
Miscellaneous	46
Total	207

N = 69

The incidental learning task was conducted in three group sessions. In each session a different order of presentation of the stimulus words was used. The Experimenter read out the 24 words at a 1 per 30 sec. rate. The subjects wrote down each word on their test sheets as it was presented and then wrote beside it which poles of the three bipolar distinctions which they were using applied to it. All three groups were given four practice trials to familiarise them with the experimental format. Care was taken to ensure that all subjects understood the three distinctions they were to use. They were instructed not to use abbreviations or any shortened form such as ditto marks. This was to ensure that the time spent attending to each word in terms of the three distinctions was as great as possible. The session took just over 10 minutes to complete, subjects turning their task sheets face-down when they had finished.

The task sheets were then collected and answer sheets for the single trial cued free recall session were distributed. For each subject the three distinctions which he had just been using were printed at the top of his answer sheet. Before commencement the subjects were told of the provision of the relevant distinctions on their answer sheets. They were told that these were provided in order to help the subjects remember the words. Subjects then wrote down as many of the words used as stimuli in the second session as they could remember. Answer sheets were collected two mins. after the last word had been written. The session took about five minutes to complete.

RESULTS

The mean recall scores for the three groups are shown in Table 5.3. As can be seen, the predicted order of increasing amounts recalled, i.e. from a minimum for Group P to a maximum for Group S, was supported. An analysis of variance on the recall scores, summarised in Table A.8 in the Appendix, showed a reliable difference in amount recalled between groups, $F(2,41) = 8.00, p < 0.01$. A linear trend test showed a reliable trend of increasing amounts recalled in the predicted direction, $F(1,41) = 16.22, p < 0.001$. A deviation from the trend test was nonsignificant.

Thus the prediction that there would be increasing amounts recalled from Group P, through Group G to a maximum in Group S, was strongly supported. Processing words in terms of self-generated semantic distinctions resulted in better recall than processing in terms of experimenter-determined semantic distinctions or phonological distinctions.

In order to obtain an estimate of the variety of semantic distinctions generated by the subjects, a content analysis was conducted of the three most salient distinctions (constructs) selected for each of the 69 subjects by factor analysis of their repertory grids. The distinctions were assigned to groups by the joint decision of a panel of three judges. The content analysis is shown in Table 5.4. As can be seen, by far the most frequently occurring distinctions corresponded closely to those originally proposed by Johnson-Laird and Shapiro, i.e. C/NC, S/L and N/M. The distinction chosen to label each group of distinctions was chosen as most representative of that group by the panel. Thus the distinction 'Edible-Inedible' was used by many subjects and was therefore used to label the category

FIGURE 5.1: Experiment 7

The percentage of subjects in Group G whose self-generated distinctions were similar to the general distinctions C/NC, S/L and N/M, as judged by a committee of three judges

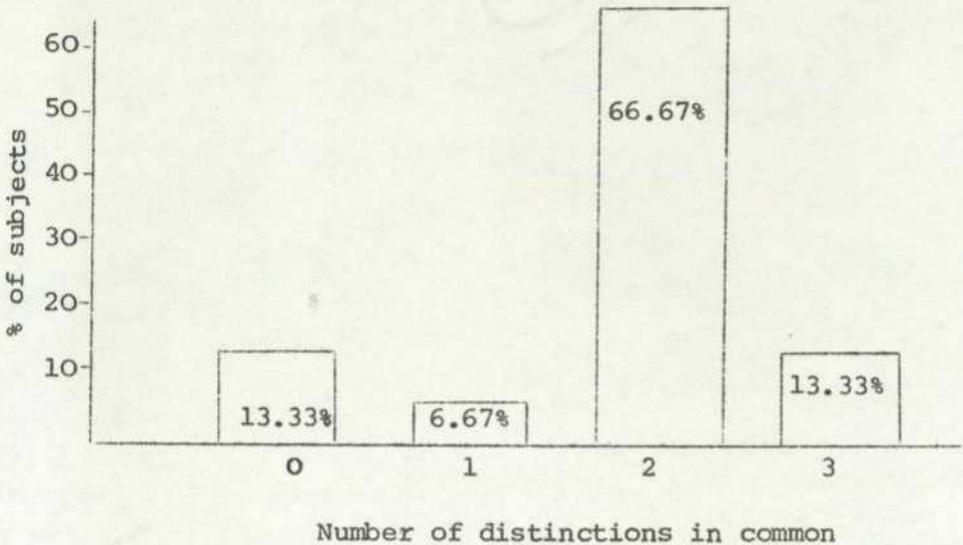
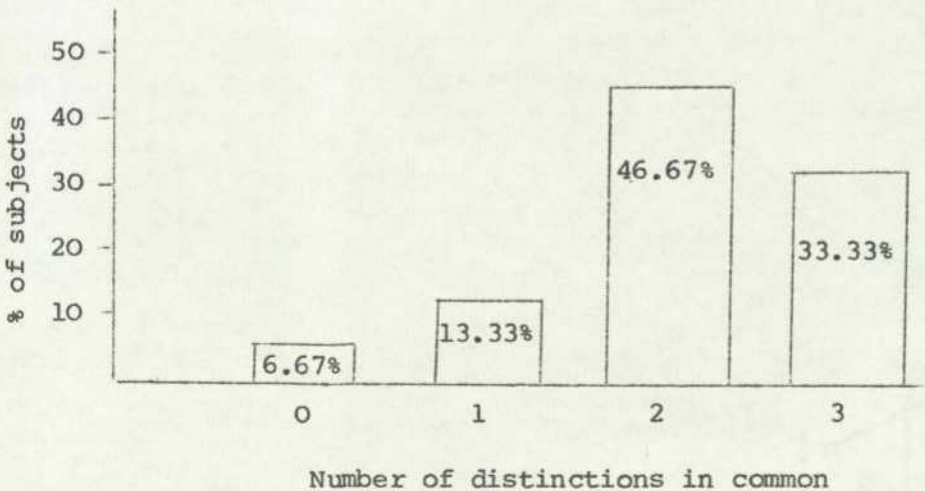


FIGURE 5.2: Experiment 7

The percentage of subjects in Group S whose self-generated distinctions were similar to the general distinctions C/NC, S/L and N/M, as judged by a committee of three judges



which most closely corresponds to Johnson-Laird and Shapiro's 'Consumable-Nonconsumable' distinction. It must be stressed that such generalisations in which individual subject's constructs are grouped together in general categories denies the individualised interpretation which underlay each subject's verbal expression of a construct. The content analysis therefore can only present a rough estimate of the most frequently occurring distinctions.

In order to clarify the question of whether subjects in Group G had generated distinctions which were the same as the experimenter-determined distinctions they had used, a count was made of the number of distinctions in common. This was important because if the number in common was high, then it would seem that Group G had, to a large extent, been using their own most salient distinctions, and their experimental experience would approximate more closely to that of Group S. The major difference between G and S would then be that Group S used self-phrased and self-generated distinctions, while Group G used experimenter-phrased but self-generated distinctions. The percentage of self- and experimenter-generated distinctions which Group G had in common is shown in Figure 5.1. Figure 5.2 presents a similar count for Group S, i.e., the percentage of the self-generated distinctions which they were using which were the same as the general distinctions. From Figures 5.1 and 5.2 it can be seen that for 80% of the subjects in both Group G and Group S, their three most salient self-generated distinctions were judged to be expressing the same concept as two or three of the general distinctions used by Group G. However, only 27% of all subjects (including those in Group P) had all three distinctions in common with the three experimenter-determined distinctions. As with the content analysis, the committee of three

judges decided jointly on whether or not individual subject's distinctions corresponded to the general distinctions, as the phrasing was rarely identical. As stressed above, such assumptions about generality of meaning should only be taken to provide a rough estimate of the similarity between the individual subject's distinctions and the general distinctions.

DISCUSSION

The hypothesis that there would be a trend of increasing amounts recalled from Group P, through Group G to a maximum in Group S was confirmed. Processing the TBR material in terms of semantic distinctions led to greater recall than processing in terms of phonological distinctions. For the two groups using semantic distinctions, the group who used self-generated distinctions recalled more. These findings fit well with the 'levels of processing' approach to memory storage, a greater depth of processing being seen as resulting in greater trace persistence. It seemed that self-generated semantic distinctions led to the greatest 'depth' of processing.

Although the self-generated semantic distinctions promoted maximum incidental learning, it was not clear whether this was due to (i) their additional subjective salience, or (ii) the fact that they were phrased differently from the three experimenter-determined distinctions, C/NC, S/L and N/M, and that this made them generally more salient; the particular phrasings of the distinctions used by Group S subjects might have induced a form of processing which led to better recall, or (iii) their representing different and more salient semantic concepts.

This last possibility could be dismissed if the underlying

concepts for the C/NC, S/L and N/M distinctions could be shown to be equivalent to the underlying concepts for the 3 most salient self-generated distinctions for Group S subjects.

It was not possible to ascertain precisely how closely the concepts underlying the subjects' verbal expressions of semantic distinctions matched those underlying the C/NC, S/L and N/M distinctions. The emphasis on the uniqueness of the individual's construction in the use of grid technique could not be overlooked. However to the extent that all subjects shared the same language, a substantial amount of generality in the underlying semantic concepts could be assumed. The individual differences would probably be more important in the salience of particular concepts, rather than in their denotative meaning. As the content analysis, Table 5.4 shows, distinctions resembling the C/NC S/L and N/M distinctions were generated with a very high frequency, relative to that of any of the other distinctions generated. This supported the contention that the C/NC, S/L and N/M distinctions had general salience.

Figure 5.2 shows that 80% of the subjects in Group S used distinctions of which two or three were judged to express the same concepts as the C/NC, S/L and N/M distinctions. The superior performance of Group S subjects was therefore unlikely to be due to their using different and more salient distinctions, but rather to some additional salience accruing from the fact that the distinctions were self-generated.

Similarly Figure 5.1 shows that 80% of the subjects in Group G used distinctions which were judged to express the same concepts as two or three of the distinctions which they had generated. Therefore the superior recall performance of Group S would appear not to have been

due to the fact that Group G was using distinctions that expressed a different and less salient concept than those which they had generated and therefore led to poorer recall scores.

The point should be made that it was not the mere fact of self-generation which might have facilitated recall for Group S, because all subjects had had the opportunity to generate relevant distinctions for the stimulus words during the completion of the repertory grids. Rather, it was argued that it was the use of such distinctions during the incidental learning session and their representation as cues during the recall session that enhanced recall.

Although the full significance of the effect of the subjective element on semantic processing tasks could not be conclusively assessed from the findings of the present experiment, the experiment provided further evidence that conflicted with the trend of decreasing salience for the three distinctions, C/NC, S/L and N/M predicted by the age-of-acquisition index. Inspection of the content analysis (Table 5.4) shows that once again the most important or outstanding distinction, this time in terms of the number of times it was generated, was the S/L distinction and not the C/NC distinction as the index would predict. Thus the inadequacy of the index as a means of specifying the ordering of the three semantic distinctions was again indicated.

5.3 General Discussion

The experiments reported in the present thesis have been directed towards an assessment of the age-of-acquisition index as a means of specifying the organisation of three semantic distinctions. If its specifications could be verified with respect to these distinctions then the organisation of other parts of the internal

lexicon might be elucidated by the same method.

The index had failed to gain clear empirical support, one of the factors which might have prevented the emergence of the predicted effect being individual differences in the salience which the three distinctions held for subjects. The present Chapter focused on this issue of subjective salience. Experiment 7 attempted to compare subject-generated distinctions with the three distinctions used to test the age-of-acquisition index on an external criterion of salience. This was the amount of recall consequent upon processing words according to the two types of distinction. Subjects who used self-generated distinctions showed a clear superiority in recall, but the exact nature of the facilitation due to the inclusion of this subjective element was not clear. It seemed that either the use of a self-expressed distinction per se allowed the words to be processed in a more meaningful way and hence facilitated greater recall, or else the particular phrasings of the self-generated distinctions increased the meaningfulness of the distinctions and hence facilitated greater recall. The resolution of this question must await further research.

The clarification of the importance of the particular phrasing of the distinctions would be of value because it was possible that the cumbersome phraseology of the three distinctions was impeding the emergence of the predicted trend. Ten subjects out of the 15 in Group S in Experiment 7 used a semantic distinction which the committee judged to be similar in conception to the consumable-nonconsumable distinction, but none of them used this phrase, 'edible-inedible' being much more frequent. This suggests that in other experiments in which the three distinctions have been used, such as Experiments 1, 2 and 3 in the present thesis, even though the Experimenter may have

ensured that the subject understood the meaning of a distinction, such as the C/NC one, if the subject had to reinterpret this into 'edible-inedible' every time he used it, it would obviously detract from his performance.

If the phrasing of the distinctions were found to be important, then this would mean that repertory grid technique provided a means of obtaining semantic distinctions phrased in terms which led to more effective semantic processing, in the sense that recall was greater using such distinctions.

The present findings, although indicating that subjects can supply distinctions that are in some way more meaningful than the C/NC, S/L and N/M distinctions, nevertheless confirms the general salience of these distinctions, in that they were by far the most frequent distinctions^o generated, in terms of underlying conception rather than manifest phraseology. However, the order of frequency with which they occurred did not concur with the age-of-acquisition index, S/L rather than C/NC being the most frequent.

Also, although many subjects included at least two of these distinctions in their self-generated set, only 27% of subjects had all three of their most salient distinctions in common with the general ones. For a few subjects their three most salient distinctions were judged not to include any of the general ones.

This suggests that there may be slight differences in salience for most subjects from that predicted by the age-of-acquisition index, and that for a few subjects concepts of C/NC, S/L and N/M may not obey the index at all because of their low salience in the individual's semantic system.

Chapter 6

Concluding Discussion

Chapter 6

Concluding Discussion

6.1 Introduction

A major objective of the thesis was to investigate the possibility of specifying the organisation in memory of semantic distinctions for which no a priori hierarchy was apparent. The semantic distinctions in question were Consumable/Nonconsumable (C/NC), Solid/Liquid (S/L) and Natural/Manmade (N/M). Four sets of studies, relevant to this objective, have been described in the preceding experimental chapters. In Section 6.2 the main relevant findings and interpretations are reviewed. This is followed by a general discussion of the organisation of non-hierarchical semantic distinctions in Section 6.3. Lines of further development are considered in Section 6.4, and finally, in Section 6.5, the principal findings and their implications are summarised.

6.2 The Major Relevant Findings and Interpretations

The main findings relevant to the organisation of non-hierarchical semantic distinctions arise from the results of four sets of studies: the studies testing the age-of-acquisition index, the study testing the use of Repertory Grid Technique in the field of semantic memory research, the studies showing that the semantic information generated by the grids represented basic and permanent aspects of the internal lexicon, and the study using such self-generated information to examine the importance of the subjective element in determining the salience of semantic distinctions in episodic memory performance. These studies are reviewed under the headings Age-of-Acquisition, Repertory Grid Technique, Tests of the Validity of the Grid Data, and Salience and Self-

Generation, respectively.

The age-of-acquisition studies

In these studies the hypothesis tested was that parts of the internal lexicon might be organised according to the age at which the semantic information was first acquired. This proposal hinged upon the notion that survival needs would ensure that the child would acquire first information that was of primary relevance to survival. It would seem logical that information of such primary importance would be readily accessible in the semantic system, while information of secondary importance would be less readily accessible. Johnson-Laird and Shapiro (unpubl.) derived an ordering of three semantic distinctions in terms of age-of-acquisition. The ordering of these distinctions, in order of decreasing importance and increasing age-of-acquisition, was C/NC, S/L, N/M. This was termed the age-of-acquisition index in the thesis, and was put to experimental test. In the first experiment, in which the assignation of semantic distinctions to low frequency words was studied, it was found that for words whose meaning was partially known as indicated by subjects writing word definitions which a panel of judges judged to be partially correct, subjects assigned the three distinctions with an accuracy which decreased according to the index from a maximum for the C/NC distinction through to a minimum for the N/M distinction.

The second study tested times taken to sort common words into the two poles of each of the three distinctions. This study supported the age-of-acquisition index in the results for the N/M distinction. The N/M distinction required a reliably longer sorting time but no difference in sorting time was found for the C/NC and S/L distinctions. As the index predicted, the error score was at a maximum for the N/M distinction but was lowest for the S/L rather than the C/NC distinction.

The third study tested the possibility that specific word effects

were masking the emergence of response trends in the direction predicted by the index. This was a study in which subjects rated the importance of the three distinctions for a set of ^{COMMON} words. The index predicted a trend of decreases in rated salience from a maximum for the C/NC distinction through to a minimum for the N/M distinction.

The predicted overall trend did not emerge, but each of the individual word effects was significant, i.e. there was significant agreement among the subjects in their salience ratings for the distinctions for individual words. It was postulated that more general effects were operating in addition to these specific effects. For example, if two distinction poles of fairly low salience, such as NC and S combined for a word, then the third distinction pole would be rated more salient than it would in combination with other more salient distinction poles. This might explain why the pole M was so salient for the word JACKET. However these effects seemed to be of a complex nature, depending on the interactions of particular combinations of distinctions, and there was insufficient data for further analysis.

In addition to confounding factors, discussed above, a further possibility for the failure of these studies to fully support the index was the presence of an interaction between subjects and salience. The remainder of the thesis was devoted to the development of a means of testing the importance of this subject effect in determining the salience of the semantic distinctions. If it did affect their salience, then another explanation of the failure of the index would have been provided.

In summary, these studies found some support for the age-of-acquisition index. Consistent differences across experiments were found between the C/M and the two other distinctions. These differences were in line with the age-of-acquisition index. One reason suggested for the

lack of complete support was the operation of specific word effects. A second reason was thought to be that individual differences in the salience of the distinctions were sufficient to confound the results. The remainder of the thesis focused upon testing this possibility.

The repertory grid technique study

Repertory grid technique was introduced as a means of obtaining subject-generated semantic distinctions, known as 'constructs' in grid terminology. The salience of such distinctions to the individual subjects would be assured. Their salience as indicated by the amount of recall promoted by semantic processing in terms of them, could be compared with that of the three distinctions which had been used to test the age-of-acquisition index. If they were of greater salience, then the importance of the subjective element in determining salience would have been supported.

The first use of the technique aimed to find out if it could provide useful data when applied in the field of semantic memory research rather than in a clinical setting. It was found that subjects could supply constructs and complete the grids by rating elements in terms of them, when the elements in question were common concrete nouns of no particular subjective relevance and when the subjects worked on their own to generate the constructs, rather than supplying them through an interviewing procedure.

If the grids could be shown to be tapping basic semantic memory organisation then the technique could be used to provide information about the semantic memory structure of individual subjects and to study idiosyncratic differences in semantic organisation, such as those that might arise from differences in the subjective salience of

particular semantic distinctions.

It could also be used to study semantic memory structure in general. For example, cluster analysis of the combined element similarity data, summed across subjects, showed a fair degree of commonality between subjects, as was to be expected if the grids were tapping basic semantic features.

Thus the technique would allow for both the investigation of general aspects of semantic memory structure and for the more detailed examination of individual differences in semantic memory structure.

Objective empirical validation of the grid as a means of tapping basic semantic memory organisation was felt necessary and two validating studies were next conducted. In these the validity of the grids was tested by their predictive power for the outcome of episodic memory tasks which involved the operation of semantic memory.

Tests of the validity of the grid data

It was argued that if the semantic information provided by the grids did represent aspects of the individual's permanent and basic semantic memory structure, then groupings of words according to this information would be meaningful to subjects in episodic memory tasks involving a semantic component. Two verbal learning tasks were therefore conducted in which superior performance was predicted when the TBR material was selected so as to reflect the semantic memory organisation evidenced by the grids.

The first of these studies was a paired-associate learning study. This provided strong support for the validity of the grid information, the recall scores for word pairs selected in terms of maximum semantic similarity according to the grids being reliably superior to those for

word pairs which were randomly created.

The second test was a free recall learning study. This failed to validate the grid semantic information, the recall scores for words selected on the basis of maximum similarity not being superior to those for words selected on the basis of minimum similarity. A number of reasons for this failure other than the inadequacy of the grid information were apparent, such as the occurrence of ceiling effects and the subjects' use of non-semantic organisational strategies.

In summary, the validity of the grid semantic information was tested in two verbal learning studies which aimed to find out if it could predict differences in episodic memory performance. The results of the paired-associate learning experiment provided strong support for the validity of the information. But no support was provided by the results of the free recall learning study. It was concluded that the semantic information provided by the grids did represent aspects of the individual's permanent and basic semantic memory system.

Salience and self-generation

Empirical confirmation of the age-of-acquisition index was thought to have been impeded in part by individual differences in the salience of the distinctions used. To test this possibility, distinctions which were self-generated (by grid completion) and therefore of assured personal salience were compared in salience with the three distinctions used to test the age-of-acquisition index. If the former distinctions proved to be more salient in predicting episodic memory performance, then the importance of the subjective element in determining salience would receive more confirmation.

An external criterion of salience, that of the amount of recall

consequent upon processing words in terms of the distinctions, was chosen. Greater salience was thought to result in a greater depth of processing and hence greater memory trace persistence.

An incidental learning study showed that subjects who used self-generated distinctions to process words recalled reliably more than subjects who had used phonological distinctions or the three age-of-acquisition distinctions. However, it was not clear whether this superior facilitative effect was due to the fact that the distinctions were self-generated per se, or to the fact that the phrasing of the distinctions promoted greater recall.

The high salience of the three age-of-acquisition distinctions for the set of words in question did receive confirmation, as most subjects generated distinctions at least two of which were judged to be expressing the same concepts as these. Their proposed ordering was not confirmed, the S/L rather than the C/NC distinction having the greatest frequency of occurrence.

In summary, the self-generated distinctions were found to be of greater salience than the age-of-acquisition distinctions in predicting episodic memory performance. Of the three most salient self-generated distinctions, defined in terms of the amount of variance in the grid data for which they accounted, the most frequently occurring were judged to be the same as the age-of-acquisition distinctions in underlying semantic conception. This indicated that the self-generated distinctions did not promote better episodic memory performance because they expressed different and more salient semantic distinctions. Their greater salience was thought to derive either from the fact of their being self-generated, or because they were expressed in a more salient way. Support was provided for the salience of the

semantic concepts underlying the C/NC, S/L and N/M distinctions, but the proposed ordering based on a decrease in salience from a maximum for the C/NC through to a minimum for the N/M distinction was not fully supported. The N/M distinction occurred least frequently in the three most salient self-generated distinctions for each subject, as the proposed ordering predicted. However the S/L rather than the C/NC distinction occurred most frequently and therefore was the most salient in these terms.

6.3 General Discussion of the Organisation of Non-Hierarchical Semantic Distinctions

The organisation of the internal lexicon is poorly understood. Recently a number of psychological accounts of its structure have been proposed, but empirical tests of these have tended to centre on the semantic elements for which a natural hierarchy is apparent.

A basic question which has arisen concerns the nature of the fundamental semantic units in the semantic memory system; are these best conceptualised as words, or as more elemental components of meaning? Evidence, although far from conclusive, indicates that the conception of units more elemental than words can account for more of the empirical data on semantic memory structure. Thus the meaning of words can be thought of as breaking down into a number of underlying unitary semantic concepts, the specific meaning of individual words being determined by the addition of idiosyncratic distinguishers.

Three basic components of meaning which have been proposed to represent a small part of the semantic memory structure are the distinctions C/NC, S/L and N/M. The organisation of these distinctions relative to each other cannot be couched in any a priori hierarchy, and the

present thesis attempted to specify their organisation in terms of an age-of-acquisition index. It was also suggested that the structure of other parts of the internal lexicon might be elucidated by the same means.

The ordering proposed (Johnson-Laird and Shapiro, unpub.) was such that ease of access to the distinctions, within the semantic *memory* structure, would be at a maximum for the C/NC distinction and would decrease from the S/L to the N/M distinction. The justification for this claim rested on judgements about the importance of the distinctions in understanding the world. It was postulated that children acquire such information according to its usefulness and that access would be fastest to information of high survival value.

Support for the later acquisition of the N/M distinction includes the results of Piaget's (1955) studies, which show that abstract concepts such as N/M are grasped later than more concrete concepts. However there was no clear support for the ordering of the C/NC and the S/L distinctions in the literature.

The findings reported in the thesis have tended to support the ordering with respect to the N/M distinction, but findings with respect to the other two distinctions were ambiguous. A number of reasons have been proposed for this lack of conclusive support:

- 1) It might be that the C/N and S/L distinctions do not represent a developmental sequence, but rather are acquired at about the same time in childhood. Certainly both could be expressed through 'enactive representation' (Bruner et al., 1966), being learnt during Piaget's 'sensory-motor period'.
- 2) The high imagery value of the S/L distinction (Paivio, 1971) might increase its rated salience and retrieval speed above that

of the C/N distinction.

- 3) According to the subjects' reports neither the C/N nor the N/M distinctions were unitary concepts, but required the consideration of a number of factors. For example, modern technology manufactures synthetically many products which once only occurred naturally, e.g. fur. The N/M distinction has become accordingly more complex. As the distinctions were assumed to represent unitary components of meaning, this would obviously confound the results. Unfortunately it is difficult to select distinctions which are ostensibly independent but for which a set of words assignable to one of the two poles for each distinction is available.
- 4) The phrasing of the distinctions may not have corresponded to the subjects' phrasings. This was felt to be most likely for the C/NC distinction, i.e. no subject produced this phrase in grid generation, but phrases such as edible/inedible were common.
- 5) The sorting experiment (Experiment 2) in particular would also have been confounded by the imbalance between the poles of the distinctions, revealed by Experiment 3. Differences in salience between the poles may have been large enough to outweigh any differences in salience between the distinctions.
- 6) Experiment 3 provided confirmation of the notion that specific word effects were operating against the emergence of the overall trend. However, some regularities in the salience rankings were apparent and it was postulated that further research might reveal that the salience of individual distinction poles were determined by the combination of other distinction poles with which they were grouped for particular words. This suggested that the assumption

of the independence of the distinctions needed reconsideration.

Although they were independent in denotative meaning, they interacted to determine the salience which accrued to the particular meaning components.

- 7) Experiment 7 showed that subject-generated distinctions were more salient in determining episodic memory performance than the three C/N, S/L and N/M distinctions. Further research is necessary to clarify whether their superior salience was due to the phrasing of the subject-generated distinctions or to the use of self-generated distinctions per se. If the latter were the case then attempts to specify semantic memory organisation for adult subjects in terms of differences in the importance of semantic distinctions would become contentious as large differences in subjective salience became apparent.

The conclusion that must be drawn from these considerations is that specification of the semantic organisation of distinctions based on age-of-acquisition and initial salience is too simple a concept to adequately represent the complexities of adult semantic structure. It might be possible to show that a simple age-of-acquisition trend does operate for children up to a certain age. Brown (1958) stressed that the semantic categorisation which is most useful for very young children may change as they grow older. However, for young children the specification of the structure of parts of their semantic memory system in terms of general principles of utility might be a possibility.

The proposed ordering for the three semantic distinctions for which no natural hierarchy was apparent was not wholly supported by the experimental results, but several interesting lines of further research were generated.

6.4 Lines of Further Development

In this section, consideration is given to three lines of development arising from the results of the present experiments. The first concerns the use of grid technique as a tool in the field of semantic memory research. The second concerns the possibility of complex differences in the salience of semantic distinctions for particular words. These differences are thought to depend upon the particular combinations of distinctions which define a word's meaning. This has implications for the specification of the meaning of individual words. The third concerns the finding that subject-generated distinctions were of greater salience in predicting episodic memory performance than were the three experimenter-supplied distinctions. The origin of this superior salience requires clarification. The three lines of development are presented under separate sub-headings.

Repertory grid technique in semantic memory research

A secondary objective of the thesis was to develop repertory grid technique as a general means of studying semantic memory organisation. One problem was whether subjects would find the grid a meaningful task when conducted under psychological laboratory conditions rather than as an interviewing procedure and when the elements for which constructs were to be supplied were not selected because of any particular relevance to the individual subject. Inspection of the completed grids and of the cluster analysis of this data suggested that meaningful information about semantic memory organisation was being produced. Cluster analysis of the summed element similarity data across subjects indicated a fair degree of commonality between subjects. This was to be expected if the grids were tapping basic semantic memory structure.

The possibility is therefore open for the use of this technique as a means of testing theories of semantic memory organisation. Hypotheses about the structure of particular lexical domains could be tested by supplying subjects with words drawn from the domain as elements on a grid. The subjects could then generate constructs for these elements. This would serve as a useful means of specifying the distinctions operating in that domain. Cluster analysis of the completed grids could reveal the structural groupings of the elements in terms of the constructs and of the constructs in terms of the elements. Thus predictions about word organisation and also about the structural hierarchies of the underlying semantic distinctions could be tested.

As Kelly's approach emphasises the unique quality of each individual's construct system, nomothetic analyses of grids from a group of subjects would seem to ignore valuable intra-subject data. Accordingly, grid technique was used in experiments reported in the thesis as a means of studying subjective semantic memory organisation, in an attempt to assess the importance of the subjective element in semantic processing tasks.

A full assessment of the technique as a means of isolating basic conceptual elements of meaning is beyond the scope of the present thesis, although evidence that it does reveal fundamental aspects of semantic memory organisation has been presented. Such an assessment would require the initial clarification of the degree of equivalence that exists between semantic features and constructs. Certainly the words used to specify the poles of the construct must not be confused with the underlying conceptual distinction. The question that should be answered conclusively is whether the constructs express unitary or

more complex meaning dimensions. The tests of the basic nature of the constructs used in the thesis were not sufficiently stringent for their unitary nature to be assumed in general.

The unique characteristics of the grid as a technique for idiographic rather than nomothetic research mean that generalisations from the data of individual subjects would be problematical. However, the technique would be of particular use in clarifying to the subject his use of a particular semantic structure and a variety of language problems might lend themselves to this sort of therapy. In addition, if the equivalence of constructs and semantic features were confirmed, the actual generation of semantic distinctions by the technique should provide the psycholinguist with explicit examples of basic meaning dimensions on which to base his theories.

Complex effects due to particular combinations of distinctions

Experiment 3 showed significant agreement among subjects in ranking the salience of the three distinctions for each of the eight words studied, although between the words there was no trace of any general trend. Unfortunately the numbers involved were insufficient for statistical analysis, but inspection of the data did suggest that certain more general characteristics of the salience rankings might be isolated.

For example, it seemed that the combination of distinct poles in a word such as JACKET, i.e. NC, S and M, might be expected to result in a very high salience ranking for M, because of the generally low salience of the other two poles. This would obviously have an important effect on the meaning of that particular word, its manmade qualities being markedly potent. This suggested that complex interactions

between the denotatively independent distinctions occurred in the determination of their individual levels of salience.

In order to examine such complex effects empirically, it would be necessary to obtain salience rankings for distinctions for a large number of words. The various combinations of distinctions which led to the exaggeration or diminution of the salience of particular poles could then be calculated.

If meaningful effects of particular combinations of distinction poles could be specified, this would open up a very challenging new area of semantic memory research. The specification of the meanings of individual words would become not just a matter of specifying their set of semantic distinctions and idiosyncratic distinguishers, but also of specifying the interaction of their distinctions and the resultant dominance and weakness of the salience of these particular distinctions in the word's meaning.

It would also mean that attempts to specify the organisation of parts of the internal lexicon by means such as the age-of-acquisition index does not allow the complexities of adult semantic memory structure to be adequately taken/into account. Ease of retrieval of particular distinctions for a set of words would not be dependent solely upon their individual levels of salience, but on the complex interactions of the particular combinations of distinctions in each individual word.

Additional salience through self-generation

Experiment 7 showed that subject-generated distinctions were of greater salience in predicting episodic memory performance than the three experimenter-supplied distinctions. This had important implications for the use of experimenter-supplied distinctions in semantic memory

research.

The question arose of whether the greater salience was due to the self-generation per se, or to changes in phrasing consequent upon self-generation. The resolution of this problem is necessary before either line of research can be developed further.

A simple experiment to answer this question would involve two groups of subjects processing words in terms of semantic distinctions in an incidental learning task, and then completing a single trial free recall session. One group would use the three experimenter-supplied distinctions (used in Experiment 7) while the second group would use the distinctions generated and used by Group S in Experiment 7. This would remove the element of self-generation but retain any additional salience accruing from improvements in the phrasing of the distinctions.

If recall was better for the group using the differently phrased distinctions, this would suggest that characteristics of the phrasing of the experimenter-supplied distinctions were detracting from their salience. One distinction in particular, the C/NC distinction, was never phrased as such by Group S, phrases such as 'Edible/Inedible' and 'Tasty/Not Tasty' being used instead. This would imply that subsequent studies using experimenter-generated distinctions should conduct a preliminary survey of the modes of expression of the particular distinctions common among their subject population, thereby ensuring maximum consensus between experimenter and subject on the meaning of the semantic distinctions under investigation.

If recall was not better for the group using the differently phrased distinctions, this would suggest that it was the use of self-generated distinctions that was responsible for the increased

salience. All groups had generated their own distinctions during repertory grid completion, so it would be the actual use of the self-generated distinctions during the incidental learning session that was significant.

Such a result would have far reaching implications for research such as that attempting to specify semantic organisation in terms of differences in the importance of semantic distinctions. Generalisations about levels of importance would become more contentious as large differences in subjective salience became apparent.

A possibility here might be that studies with children rather than adults would be less prone to being confounded by this factor. As Roger Brown (1958) made clear in his work on the acquisition of word meaning according to utility, terms which are most useful for very young children may change as they grow older. It might be possible to specify the structure of parts of the child's internal lexicon in terms of general principles of utility. To do this for adult subjects would prove a more complex and difficult task.

6.5 Summary of Principle Findings and their Implications

The principal findings reported in the thesis come from four sets of studies: the age-of-acquisition studies, the Repertory Grid Technique study, the studies testing the validity of the Repertory Grid data and the study investigating the salience of the self-generated distinctions.

The assumption which underlay the age-of-acquisition studies was that the earlier in life a distinction was acquired, the greater would be its salience, i.e. distinctions of importance for survival or of utility in coping with the environment would be acquired first.

A trend of decreasing salience was predicted for the three distinctions in the order C/NC, S/L to N/M. Accuracy of assignation of the distinctions followed the trend but there was no reliable difference in sorting time for the C/NC and S/L distinctions, although for both this was reliably less than that for the N/M distinction. No trace of the overall trend was obtained from the subjects' rankings of the salience of the distinctions for a set of words. However the ranking study suggested interesting possibilities with regard to the occurrence of complex interactions between the supposedly independent distinctions.

Several reasons for the failure of the studies to adequately support the index were proposed, notably the presence of specific word effects, the occurrence of complex interactions between the distinctions resulting in marked alterations in salience, the inadequacy of the distinctions chosen to represent the supposed developmental sequence and the uncertainty of the predictive power of the developmental sequence for adult subjects.

A further confounding variable was thought to be individual differences in salience. To test for this, Repertory Grid Technique was introduced as a means of obtaining subject-generated and therefore supposedly salient semantic distinctions. Two verbal learning studies were conducted to test its validity in predicting the outcome of episodic memory experiments involving a semantic component and the conclusion was drawn that the semantic information provided by the grids represented aspects of permanent and fundamental semantic memory organisation.

This finding supported the use of Repertory Grid Technique as a tool for semantic memory research. It provided detailed information about individual semantic memory organisation, but could also be used for

studying commonalities between subjects in semantic memory structure.

The salience of the subject-generated distinctions produced by grid technique was then compared with the salience of the three distinctions used in the earlier studies. The subject-generated distinctions were shown to have superior salience in determining episodic memory performance. It was not clear whether this derived from the use of self-generated distinctions per se, or from their different and more meaningful phrasing. The implications of both these possibilities were discussed.

The attempt to specify the organisation of three semantic distinctions for which no a priori hierarchy was apparent by means of the age-of-acquisition index received some support from the results of the present studies. But it was concluded that the age-of-acquisition index, in its present form, was too simple a concept to adequately represent the complexities of adult semantic memory structure. However, it might be possible to show that a simple age-of-acquisition trend operated for children up to a certain age. In addition several interesting lines of further development were generated, especially the investigation of changes in the salience of the distinctions due to their complex interactions. The use of the age-of-acquisition index to study the organisation of the three distinctions led to the proposal of some important new variables in semantic memory research, such as the superior salience of self-generated over experimenter-determined semantic distinctions and the effects on distinction salience of complex interactions between denotatively independent distinctions.

References

- ANGLIN, J.M. (197.). The Growth of Word Meaning. Cambridge, Massachusetts: MIT Press.
- BADDELEY, A.D. & LEVY, B.A. (1971). Semantic coding and short-term memory. *J. exp. Psychol.*, 89, 132-136.
- BAKAN, D. (1954). A generalisation of Sidman's results on group and individual functions, and a criterion. *Psychol. Bull.*, 51, 63-4.
- BANNISTER, D. & FRANSELLA, F. (1971). *Inquiring Man*. London: Penguin Books.
- BANNISTER, D. & MAIR, J.M.M. (1968). *The Evaluation of Personal Constructs*. London: Academic Press.
- BIERWISCH, M. (1967). Some semantic universals of German adjectives. *Foundations of Language*, 3, 1-36.
- BIERWISCH, M. (1969). Certain problems of semantic representations. *Foundations of Language*, 5, 153-84.
- BOUSFIELD, W.A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.*, 49, 229-240.
- BOUSFIELD, D.A.K. & BOUSFIELD, W.A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychol. Rep.*, 19, 935-942.
- BOWER, G.H. (1967). A multicomponent theory of memory trace. In K.W. Spence and J.T. Spence (Eds.) *The Psychology of Learning and Motivation*, Volume I. New York: Academic Press.
- BOWER, G.H. (1970). Organisational factors in memory. *Cog. Psychol.*, 1, 18-46.
- BOWER, G.H. (1971). Mental imagery and associative learning. In L.W. Gregg (Ed.), *Cognition in Learning and Memory*. New York: Wiley.
- BOWER, G.H., CLARK, M.C., LESGOLD, A.M. & WINZENZ, D. (1969). Hierarchical retrieval schemes in the recall of categorised word lists. *J. verb. learn. verb. Behav.*, 18, 323-343.

- BROWN, R. (1958). How shall a thing be called? *Psychol. Rev.*, 65, 14-21.
- BRUNER, J.S. & OLVER, R.R. (1963). The development of equivalence transformations in children. Monograph of the Society for Research in Child Development, 28 (whole no. 86) 125-141.
- BRUNER, J.S., OLVER, R.R. & GREENFIELD, P.M. (1966). *Studies in Cognitive Growth*. New York: Wiley.
- CARROLL, J.B. & WHITE, M.N. (1973) Word frequency and age of acquisition as determiners of picture-naming latency. *Quart. J. exp. Psychol.*, 25, 85-95.
- CHOMSKY, N. (1965a). *Current Issues in Linguistic Theory*. The Hague: Mouton.
- CHOMSKY, N. (1965b). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
- CLARK, H.H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *J. verb. Learn. verb. Behav.*, 12, 335-359.
- COLLINS, A.M. & QUILLIAN, M.R. (1969). Retrieval time from semantic memory. *J. verb. Learn. verb. Behav.*, 8, 240-247.
- COLLINS, A.M. & QUILLIAN, M.R. (1970a). Facilitating retrieval from semantic memory: The effects of repeating part of an inference. *Acta Psychol.*, 33, 304-314.
- COLLINS, A.M. & QUILLIAN, M.R. (1970b). Does category size affect categorisation time? *J. verb. Learn. verb. Behav.*, 9, 432-438.
- COLLINS, A.M. & QUILLIAN, M.R. (1972). How to make a language user. In E. Tulving and W. Donaldson (Eds.) *Organisation of Memory*. New York: Academic Press.
- CRAIK, F.I.M. (1972). A 'levels of analysis' view of memory. Paper presented at the 2nd Erindale Symposium on Communication and Affect.

- CRAIK, F.I.M. & LOCKHART, R.S. (1972). Levels of processing: a framework for memory research. *J. verb. Learn. verb. Behav.*, 11, 671-684.
- DALE, H.C.A. & McGLAUGHLIN, A. (1970). Enhancing memory with redundant labels. *Nature*, 227, 411-412.
- FAGAN, J.F. (1968). Measuring verbal recall II: The ITR score expressed as a ratio. *Psychon. Sci.*, 11, 205.
- FILLENBAUM, S. (1971). Psycholinguistics. *Ann. Rev. Psychol.*, 22, 251-308.
- FILLENBAUM, S. & RAPOPORT, A. (1971). *Structures in the Subjective Lexicon*. New York: Academic Press.
- FRANSELLA, F. (1965). The effects of imposed rhythm and certain aspects of personality on the speech of stutterers. Unpublished Ph.D. thesis. University of London.
- FRINCKE, G. (1968). Word characteristics, associative-relatedness and the free recall of nouns. *J. verb. Learn. verb. Behav.*, 7, 366-372.
- HART, J.T. (1967). Memory and the memory-monitoring process. *J. verb. Learn. verb. Behav.*, 6, 685-691.
- HICKS, R.E. & YOUNG, R.K. (1972). Part-whole list transfer in free recall: A reappraisal. *J. exp. Psychol.*, 96, 328-333.
- HYDE, T.S. & JENKINS, J.J. (1969). Differential effects of incidental tasks on the organisation of recall of a list of highly associated words. *J. exp. Psychol.*, 82, 472-481.
- HYDE, T.S. & JENKINS, J.J. (1973). Recall for words as a function of semantic orienting tasks. *J. verb. Learn. verb. Behav.* (in press).
- JOHNSON, S.C. (1967). Hierarchical clustering schemes. *Psychomatrika*, 32, 241-254.

- JOHNSON-LAIRD, P.N. (1973). Meaning and the mental lexicon. Conference on Current Research in Aspects of Long-Term Memory. University of Dundee, July, 1973.
- JOHNSON-LAIRD, P.N. (1974). Experimental Psycholinguistics. *Ann. Rev. Psychol.* (in press).
- JOHNSON-LAIRD, P.N. & SHAPIRO, D. The meaning of words. Unpublished research report. University College, London.
- KATZ, J. (1966). *The Philosophy of Language*. New York: Harper and Row.
- KATZ, J. & FODOR, J. (1963). The structure of a semantic theory. *Language*, 39, 170-210.
- KELLY, G.A. (1955). *The Psychology of Personal Constructs, Volume I and II*. New York: Norton.
- KEPPEL, G. & UNDERWOOD, B.J. (1962). Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.*, 1, 153-161.
- KINTSCH, W. (1972). Notes on the structure of semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organisation of Memory*. London: Academic Press.
- MANDLER, G. (1967). Organisation and memory. In K.W. Spence and J.T. Spence (Eds.), *The Psychology of Learning and Motivation, Volume I*. New York: Academic Press.
- MARSHALL, J.C. (1970). A note on semantic theory. In G.B. Flores D'Arcais and W.J.M. Levelt (Eds.) *Advances in Psycholinguistics*. London: North Holland.
- MILLER, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 63, 81-97.

- MILLER, G.A. (1967). Psycholinguistic approaches to the study of communication. In D.L. Arm (Ed.) *Journeys in Science*. Albuquerque: University of New Mexico Press.
- MILLER, G.A. (1969). A psychological method to investigate verbal concepts. *J. math. Psychol.*, 6, 169-191.
- MILLER, G. A. (1972). English verbs of motion: A case study. In A.W. Melton and E. Martin (Eds.), *Coding Processes in Human Memory*. New York: Wiley.
- MORTON, J. (1970). A functional model for memory. In D.A. Norman (Ed.), *Models of Human Memory*. New York: Academic Press.
- MURDOCK, B.B. Jr.(1968) Modality effects in short-term memory: Storage or retrieval. *J. exp. Psychol.*, 77, 79-86.
- NORMAN, D.A. (1969). *Memory and Attention*. New York: Wiley.
- OSGOOD, C.E., SUCI, G.J. & TANNENBAUM, P.M. (1957). *The Measurement of Meaning*. University of Illinois Press.
- PAIVIO, A. (1968). Effects of imagery instructions and concreteness of memory pegs in a mnemonic system. *Proc. 76th Ann. Convention of Amer. Psychol. Assoc.*, 77-78.
- PAIVIO, A. (1971). *Imagery and Verbal Processes*. London: Holt, Rinehart and Winston.
- PIAGET, J. (1955). *The Construction of Reality in the Child*. London: Routledge and Kegan Paul.
- POSTMAN, L. (1971). Organisation and interference. *Psychol. Rev.*, 78, 290-302.
- RUSSELL, W.A. & JENKINS, J.J. (1954). The complete Minnesota norma for responses to 100 words from the Kent-Rosanoff word association test. Tech. Rep. No. 11, ONR Contract N8, onr-66216. University of Minnesota.

- SCHAEFFER, B. & WALLACE, R. (1969). Semantic similarity and the comparison of word meanings. *J. exp. Psychol.*, 82, 343-346.
- SCHAEFFER, B. & WALLACE, R. (1970). The comparison of word meanings. *J. exp. Psychol.*, 86, 144-152.
- SEIBEL, R. (1967). Organisation in learning. Project report, U.S. Dept. of Health, Education and Welfare, Pennsylvania State University.
- SELFRIDGE, O.G. & NEISSER, U. (1960). Pattern recognition by machine. *Sci. Amer.*, 203, 60-68.
- SHAPIRO, S.I. (1970). The commonality of subjective organisation in free recall. *Psychon. Sci.*, 21, 20.
- SHUELL, T.J. (1969). Clustering and organisation in free recall. *Psychol. Bull.*, 72, 353-374.
- SIDMAN, M. (1952). A note on functional relations obtained from group data. *Psychol. Bull.*, 49, 263-269.
- SUTHERLAND, N.S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proc. Royal Soc. Series B*, 171, 297-317.
- THORNDIKE, E.L. & LORGE, I. (1944). The teacher's word book of 30,000 words. New York: Columbia University Press.
- TRIESMAN, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *J. verb. Learn. verb. Behav.*, 3, 449-459.
- TULVING, E. (1962). Subjective organisation in the free recall of 'unrelated' words. *Psychol. Rev.*, 4, 344-354.
- TULVING, E. (1966). Subjective organisation and effects of repetition in multitrial free-recall learning. *J. verb. Learn. verb. Behav.*, 5, 193-197.
- TULVING, E. (1968). Theoretical issues in free recall. In T.R. Dixon and D.L. Horton (Eds.), *Verbal Behavior and General Behavior Theory*. New Jersey: Prentice-Hall.
- TULVING, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organisation of Memory*. London: Academic Press.

- TULVING, E. & OSLER, S. (1967). Transfer effects in whole/part free recall learning. *Canad. J. Psychol.*, 21, 253-262.
- UNDERWOOD, B.J. (1969). Attributes of memory. *Psychol. Rev.*, 76, 559-573.
- VOSS, J.F. (1972). On the relationship of associative and organisational processes. In E. Tulving and W. Donaldson (Eds.) *Organisation of Memory*. London: Academic Press.
- WARR, P.B., SCHRODER, H.M. & BLACKMAN, S. (1969). The structure of political judgement. *Brit. J. soc. clin. Psychol.*, 8, 32-43.
- WEINREICH, U. (1966). Explorations in semantic theory. In T.A. Sebeok (Ed.), *Current Trends in Linguistics, Volume III*, The Hague: Mouton.
- WICKENS, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychol. Rev.*, 77, 1-15.
- WICKENS, D.D. (1972). Characteristics of word encoding. In A.W. Melton and E. Martin (Eds.), *Coding Processes in Human Memory*. London: Wiley.
- WICKENS, D.D., BORN, D.G. & ALLEN, C.K. (1963). Proactive inhibition and item similarity in short-term memory. *J. verb. Learn. verb. Behav.*, 2, 440-445.
- WICKLUND, D.A., PALERMO, D.S. & JENKINS, J.J. (1964). Effects of associative strength and response hierarchy on paired-associate learning. *J. verb. Learn. verb. Behav.*, 3, 413-420.
- WISHART, D. (1969). *Clustan 1A User Manual*. Computing Laboratory, University of St. Andrews.

Appendix

TABLE A.1: Experiment 2

The analysis of variance on the individual mean sorting
times for the three semantic distinctions

Source	DF	SS	MS	F	P
Subjects (S)	17	656.83	38.64	10.88	0.001
Main Effect (A)	2	154.21	77.10	21.72	0.001
S x A	34	120.73	3.55		
Total	53	931.76			

TABLE A.2: Experiment 5

The mean recall scores for the High and Low lists of
word pairs

	Trials									
Condition	1	2	3	4	5	6	7	8	9	10
Low	0	2.00	3.25	4.00	4.63	6.00	6.88	7.13	7.57	8.31
High	0	3.69	5.13	7.00	7.31	8.38	8.44	9.19	8.56	9.19

TABLE A.3: Experiment 5

Analysis of variance on the recall scores for the high and
low lists of word pairs

Source	DF	SS	MS	F	P
Subjects (S)	15	761.24	50.74		
Condition (C)	1	226.13	226.13	12.57	0.01
C x S	15	269.82	17.99		
Trials (T)	9	2247.01	249.67	105.20	0.001
T x S	135	320.73	2.37		
C x T	9	58.78	6.53	3.39	0.01
C x T x S	135	259.72	1.92		
Total	319	4143.44			

TABLE A.4: Experiment 6

The mean number of words recalled correctly per trial for
the Low and High conditions

Trials										
Condition	1	2	3	4	5	6	7	8	9	10
Low	5.75	8.31	8.88	9.38	9.69	9.63	9.56	9.75	9.81	9.94
High	5.44	6.88	8.31	9.44	9.56	9.75	9.69	9.56	9.94	9.81

TABLE A.5: Experiment 6

The analysis of variance on the recall scores for the Low
and High conditions

Source	DF	SS	MS	F	P
Subjects (S)	15	81.72	5.45		
Condition (C)	1	1.13	1.13	0.64	NS
C x S	15	26.42	1.76		
Trials (T)	9	531.75	59.08	77.75	0.001
T x S	135	102.00	0.76		
C x T	9	6.65	0.74	1.121	NS
C x T x S	135	89.30	0.66		
Total	319	838.97			

TABLE A.6: Experiment 6

The mean ITR scores for the Low and High Conditions,
expressed as a percentage of the maximum possible score

Trials									
Condition	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
Low	37.49	26.34	25.43	38.67	44.97	41.05	47.71	41.08	41.64
High	25.56	25.71	37.23	36.36	42.11	43.98	48.09	55.03	50.97

TABLE A.7: Experiment 6

The analysis of variance on the ITR scores for the Low
and High conditions

Source	DF	SS	MS	F	P
Subjects (S)	15	90498.16	6033.21		
Condition (C)	1	379.02	379.02	0.19	NS
C x S	15	29864.19	1990.95		
Trials (T)	8	17003.98	2125.50	3.97	0.001
T x S	120	64231.24	535.26		
C x T	8	4308.09	538.51	1.04	NS
C x T x S	120	61624.45	513.53		
Total	287	267909.13			

TABLE A.8: Experiment 7

The analysis of variance on the recall scores of Groups
P, G and S

Source	DF	SS	MS	F	P
Between Groups	2	148.80	74.40	8.00	0.01
Within Groups	41	381.20	9.30		
Total	43	530.00			