128

<u>An investigation of the role of context in retrieval of information from semantic memory.</u>

by

PAUL JAMES SHANAHAN

Thesis submitted to the University of Stirling for the degree of Doctor of Philosophy, February, 1976.

(awarded June 1976)

ACKNOWLEDGEMENTS:

I would like to express my gratitude to Dr. Philip Smith for his patient supervision, to all past and present members of the Psychology Department for the help they so willingly provided, to Mrs. J. W. Drane for coping with my hand-writing and to my parents whose support enabled me to complete this thesis.

CONTENTS

FIGURES.

Please note: There is no Table 2.

INTRODUCTION

The research reported here examines how a person's knowledge of the world is used in language recognition and production. Essentially it is concerned with the importance of a word's meaning as a factor in its recognition by a listener or reader and is its production by a speaker or writer. This area of research overlays with a great many areas in psychology, drawing upon research in attention, pattern recognition, memory, psycholinguistics and thought.

It is necessary to give some working definitions of the terms used. The definition of semantic memory used here is that supplied by Tulving (1972, p 386): "Semantic memory is the memory necessary for the use of a language. It is a mental thesaurus, the organized knowledge a person possesses about words and other verbal symbols, their meanings and referents, about relations among them and about rules, formulas and algorithms for the manipulation of these symbols, concepts and relations."

The contents of semantic memory are typically what a person would say that he "knows" rather than what he "remembers". e.g. a person might say "I know canaries are yellow" whereas

"I remember canaries are yellow" would not "sound right" to most native English speakers. This also illustrates an important property of semantic memory. The knowledge it contains is to a large extent common to members of a given culture. There will of course be individual differences but a sufficient body of knowledge will be shared in order to allow communication between persons.

Retrieval from semantic memory is used here to refer to any process that involves making use of such stored knowledge. This may range from simply deciding that a particular sound pattern has occurred in speech before to verifying complex propositions.

Context is restricted here to linguistic context. The question asked is how information provided by previous linguistic input affects processing of later input or output of language. The view of language comprehension taken here is similar to Goodman's (1967) approach to reading. This approach is described as follows: "...Reading is a psycholinguistic guessing game. It involves an interaction between thought and language. Efficient reading does not result from precise perceptions and identification of all elements but from skill in selecting the fewest, most productive cues necessary to produce guesses which are right first time. The ability to anticipate that which has not been seen, of course, is vital in reading, just as the ability to anticipate what has not yet been heard is vital in listening." (p 260)

It is assumed here that a person's ability to anticipate is dependent upon the knowledge stored in semantic memory. The way this knowledge is used will in turn depend upon how it is organized. Since the Ancient Greeks the importance of organization in memory has been recognized but it is only relatively recently that psychologists have attempted to determine the principles underlying this organization. Since Quillian (1966) a number of models of how semantic memory is organized have been proposed. These will be discussed in the following sections.

Many of the experiments reported here are concerned with what might be called "micro-context", that is how individual words, phrases and sentences affect recognition of incoming stimuli. Of course, the use of context goes far beyond the immediately preceding input but as yet there are no satisfactory theories, linguistic or psychological, that can deal with these wider aspects of language use. In fact there is still considerable disagreement over the processes involved in the recognition of single words, (see for example Rubenstein, Lewis and Rubenstein, 1971; Baron, 1973).

The approach taken to word recognition here is similar to Norman (1968) and Morton (1969). The notion which is central to both these authors and Goodman (see above) is the realization that no process can be analysed in isolation. The language system cannot decode the incoming sensory information without reference to stored knowledge. As Norman (1969, p 3) describes the role of memory, "it provides the information about the past necessary for proper understanding of the present". Thus context indicates to the memory system what knowledge is relevant to the analysis of the current input.

To summarize this approach the information provided by context (immediate past) is referred to semantic memory (past) which in turn helps to produce the best guess as to the nature of the current sensory input (present) or even the nature of input which has not yet arrived (future).

The problem examined in this research is how the organizational structure of knowledge in semantic memory influences this guessing process. Whether such guessing is an active process as suggested by some investigators (e.g. Liberman, Stevens and Halle) or a passive process suggested by others (e.g. Morton, Treisman) will be discussed in a later section.

Review of the Literature

In the first part I shall discuss the linguistic approaches to semantics that form the background to the models of semantic memory that are discussed in the second part. In the third part I shall try to relate the models of semantic memory to some models of speech production and recognition. No attempt is made here to review the more technical aspects of the linguistics. The theories are only considered from a psychological point of view.

The Linguistic Background

Some indication of the problem facing linguists dealing with semantics is given by the fact that Ogden and Richard's (1923) in their classic book "The meaning of meaning" were able to give 22 definitions of meaning. Meaning here will be restricted largely to what Leech (1974) calls "conceptual meaning or sense" which refers to a word's "logical, cognitive or denotative content" in contrast to other aspects of meaning such as a word's connotative or stylistic meaning.

In spite of the problems Ogden and Richards took an optimistic view of likely progress in semantics. In contrast Bloomfield, ten years later wrote "The statement of meanings is therefore the weak part in language study, and will remain so until human knowledge advances very far beyond its present state." (1933, p 140). This attitude dominated linguists thinking on semantics for over twenty years.

Since the 1950s psychologists have taken a growing interest in the work of linguists. Much of this interest can be attributed to Chomsky's (1959) demonstration of the inadequacy of traditional S-R theories of psychology to account for language behaviour. For present purposes the most notable aspect of Chomsky's (1957) early work is his belief that a syntactic theory could be constructed in-dependently of semantics. In this belief Chomsky still reflected the influence of Bloomfield. However, in his later work Chomsky (1965) makes some concession to the role of semantics in grammar. "In fact, it should not be taken for granted, necessarily that syntactic and semantic considerations can be sharply distinguished". (p 77).

One of the reasons for this shift in position was the publication of "The structure of a semantic theory" by Katz and Fodor (1963), described by Bouveresse (1974) as "the official reintroduction of semantics". The paper by Katz and Fodor was an attempt to produce a theory of semantics within the general framework of transformational grammar. It has been developed by Katz and Postal (1964) and Katz (1972). The account here is based on the original 1963 paper.

Although attempting to produce a semantic theory Katz and Fodor's famous statement "linguistic description minus grammar equals semantics" seems to reflect Chomsky's (1957) thoughts on semantics. "Meaning tends to be used as a catch-all term to include every aspect of language that we know very little about." (pp 103-104). In spite of their rather negative definition of semantics Katz and Fodor were prepared to try to produce a semantic theory consistent with Chomsky's transformational grammar. They describe their aims as follows:

"A semantic theory describes and explains the interpretative ability of speakers: by accounting for their performance in determining the number and content of the readings of a sentence; by detecting semantic anomalies; by deciding upon paraphrase relations between sentences; and by marking every other semantic property or relation that plays a role in this ability." (p 486).

Katz and Fodor contend that the basic fact a semantic theory must explain is that a fluent speaker can determine the meaning of a sentence in terms of the meaning of its constituent lexical items. There are two components in the semantic theory to achieve this end. The first component is a dictionary of lexical items of the language and the second component is a system of rules which operates on the full grammatical descriptions of sentences and on the dictionary entries to produce semantic interpretations for every sentence in the language.

Our main concern here is with the first component, the dictionary but some comment will be made about the rules later. (For a critical review of the operations of the rule system see Weinreich, 1966; Savin, 1973).

A dictionary entry in Katz and Fodor's theory consists of two parts, a grammatical portion which provides part-of-speech classification and a semantic portion which represents each of the distinct senses the lexical item has in its occurrence as a given part of speech. A word is represented as a string of semantic markers. This method derives from the technique of componential analysis used by anthropologists to describe kinship terms, (see for example, Wallace and Atkins, 1960; Romney and Andrade, 1964). Leech (1974) describes componential analysis as "a technique for describing interrelations of meaning by breaking each concept down into minimal components, or features, which are distinctive in terms of a semantic opposition or dimension of contrast. So 'woman' can be defined by the features + HUMAN, + ADULT, - MALE in such a way as to discriminate it from the related concepts 'girl', 'man', 'child', 'cow', etc." (p 124).

Such a technique is an attempt to reproduce in semantics the success of Jakobson and Halle (1956) in phonology in describing phonemes in terms of a limited number of distinguishing features. It seems unlikely that it will be possible to represent all lexical items in terms of a finite set of binary features. In phonology the range of possible phonemes is sharply restricted by the capabilities of the human speech organs, yet in semantics there is no comparable restriction on the range of possible lexical items. We need as many features as are necessary to produce a unique representation of each lexical item. Such a set must be open-ended. Furthermore many of the dimensions needs to represent lexical items are not binary (e.g. colour, shape) which means that simple presence or absence of a feature in a lexical item's feature list will not be sufficient to characterize that item on that dimension.

Yet it must be admitted that the idea that an item's meaning can be broken down and represented in terms of its consitituent features is attractive and some such approach is used in all the models of semantic memory discussed in the next section.

To return to the relationship between the two components of Katz and Fodor's theory (the dictionary and the projection rules) semantic interpretation involves combining semantic features of individual words to produce a description of the entire sentence. This formulation was followed by Chomsky (1965) and has been called "interpretive semantics" since the meaning of a sentence is obtained by applying semantic rules that interpret a syntactic base. Thus in classical transformational grammar syntax has 'priority' over semantics, in that the generation of a deep structure is presumed to be independent of meaning. This position may be tenable in a competence model but is certainly not acceptable in a performance model, where as a general rule the object of communication is meaning.

Since 1965 there has been a considerable movement towards granting semantics a more central role in linguistic theory. Anderson and Bower (1974) have divided this movement into the Neo-Chomskians who accept the general framework of transformational grammar and the generative semanticists who claim that Chomsky's view of semantics is inadequate even for a competence model.

The most important of the Neo-Chomskian developments is the case grammar presented by Fillmore (1968). Case grammar is designed to deal with the fact that the subjects of sentences such as John runs, John is afraid, The window broke, The medecine cures, Chicago is hot, are all treated alike in transformational grammar whereas they all have different semantic roles (examples from Anderson and Bower, (1974). In case grammar these different roles are made explicit in the deep structure by assigning cases to the items (e.g. agent, passive object, instrument etc.). Fillmore argues that by emphasizing these "semantically relevant syntactic relations" it is easier to produce a semantic interpretation of the deep structure. Case grammar has been used in a number of recent models of memory (e.g. Rumelhart, Lindsay and Norman, 1972; Anderson and Bower, 1974).

The generative semanticists (McCawley, Lakoff, Ross) differ more radically from transformational grammar. While accepting the necessity for the base, transformational, semantic and phonological components of language they question Chomsky's assumption of the deep structure as a separate level. Crystal (1971) summarizes their question as follows: "If the whole point about talking about deep structures at all in the first place was to take account of meaning... then why should not these meaning-problems be incorporated as part of the study of other meaning problems which the semantics component has to face anyway?" (p 235).

The most significant feature of the generative semanticists from the point of view of the present research is that they take semantics as their starting point for studying language rather than the classical starting point of syntax. At present there are considerable technical difficulties in using a generative semantics approach in a model of human language use. As yet no model has made extensive use of generative semantics although Rumelhart et al. make use of certain rudimentary concepts similar to those of generative semantics.

Models of Semantic Memory

While the growth in interest in semantics was occurring in linguistics there was a parallel development in psychology in the interest in meaning as a factor in memory. The "association" as the basic unit in memory has been fundamental to psychology since the British Empiricists who themselves derived the idea from Aristotle. For a long time the prevailing picture of memory was of a hotch-potch of associated ideas that arose from the accidental contiguities of experience. More recently there has been a growing recognition within experimental psychology of the fact that structure is imposed on the contents of memory. The existence of such organization had been apparent to certain analytical psychologists for a considerable length of time. In particular Jung had observed "...the tendency of ideas to become associated around certain nuclei." (from

Fordham, 1953, p 22). Within experimental psychology this aspect of memory was largely ignored. The major factor in association was held to be contiguity in time. Deese (1965) comments that "attention to this property of temporal order has led to the neglect of structure." (p 1). However, during the last 30 years, beginning with the pioneering work of Bousfield and continued by Mandler, Tulving, Bower, Deese and others, the structure of associations and the organization of memory has become one of the major topics in memory research. (For a review of this development see Deese, 1965).

The models of semantic memory considered here have all developed out of this interest in organization of knowledge. Their common approach has been described by Anderson and Bower (1974) as "neo-associationist". Neo-associationism is described as a "profane union" of methodological empiricism and methodological rationalism. "The result is a theory that irreverently intermixes connectionism with nativism, reductionism with wholism, sensationalism with intuitionism, and mechanism with vitalism." (Anderson and Bower, 1974 p 4).

One of the earliest and most influential models of semantic memory is Quillian's Teachable Language Comprehender (1966, 1967, 1969). The early versions of the model are computer based but psychological implications of the model have been examined by Collins and Quillian (1969, 1970). The model is designed only to hold denotative factual information. Information is represented in the form of an "association network". A concept is represented in the network as a node connected to other modes by different kinds of associative links. (See figure 1). The meaning of the concept is defined in terms of the other concepts to which it is linked. To obtain the meaning of a concept a search is started at the node representing that particular concept. The search spreads out along all the links leading from the original mode to the connected nodes. The search will then proceed along all the links from these nodes. In this way the meaning of the concept is defined as "...all the nodes that can be reached by an exhaustive tracing process". (Quillian, 1967, p 413). Such an exhaustive search involves rejection of Katz and Fodor's assumption that a word can be defined by a limited number of features. In Katz and Fodor's model only a subset of knowledge is called upon in defining meaning. In Quillian's model the whole of a person's world knowledge is used.

The most important determinant of memory organization in Quillian's model is the need to avoid redundancy. Quillian regarded the space available for storage as limited and assumed that information would be stored in ways that would minimize the demands on storage space. Quillian proposed that the most efficient means of storing information is in a hierarchy of superordinates. Concepts in the model are grouped into categories. (see figure 1). The properties that are shared by

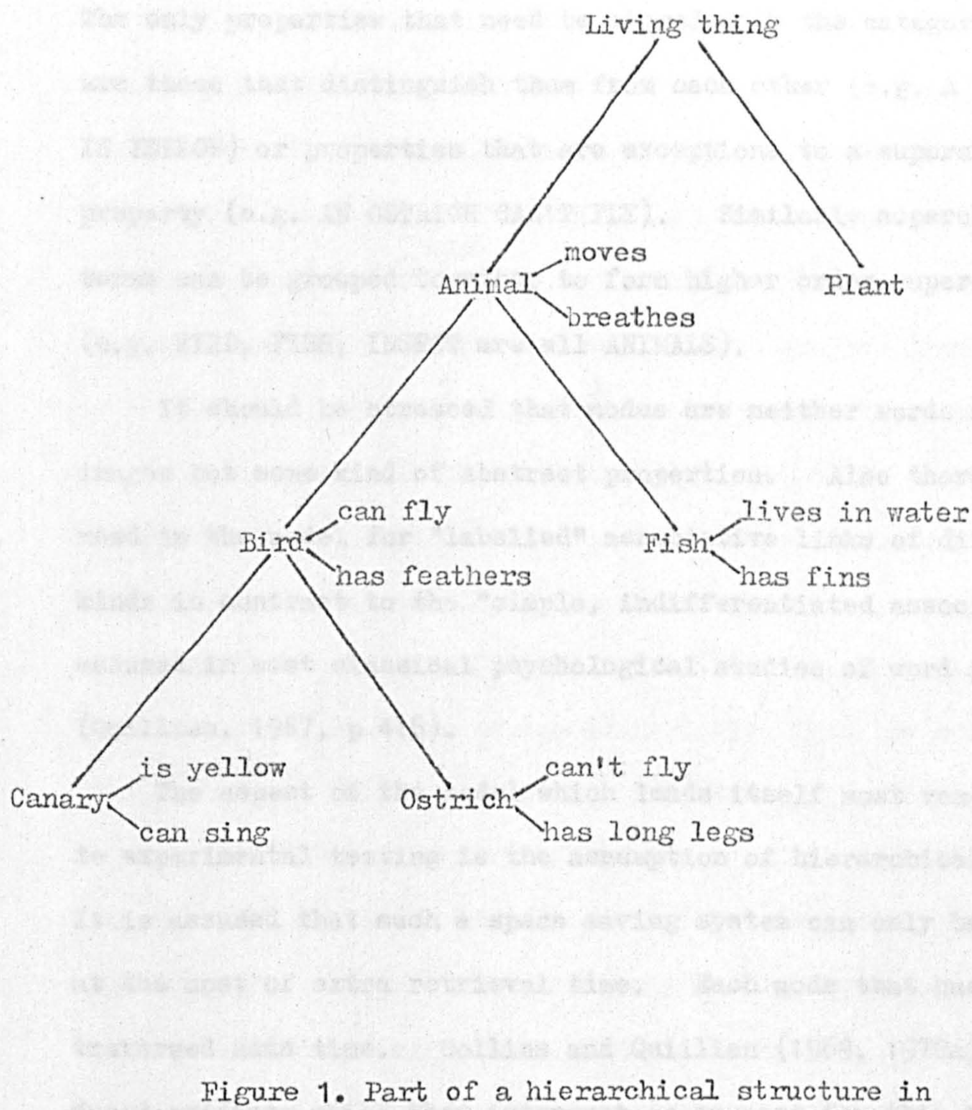


Figure 1. Part of a hierarchical structure in long term memory. Based on Collins and Quillian (1969).

all members of a category (e.g. ALL BIRDS CAN FLY) are stored only once with the superordinate. Retrieval of such a property given a category member (e.g. CAN A CANARY FLY?) would involve first retrieving that the instance belonged to that category. The only properties that need be stored with the category members are those that distinguish them from each other (e.g. A CANARY IS YELLOW) or properties that are exceptions to a superordinate property (e.g. AN OSTRICH CAN'T FLY). Similarly superordinate terms can be grouped together to form higher order superordinates (e.g. BIRD, FISH, INSECT are all ANIMALS).

It should be stressed that nodes are neither words nor images but some kind of abstract properties. Also there is a need in the model for "labelled" associative links of different kinds in contrast to the "simple, indifferentiated associations assumed in most classical psychological studies of word associations". (Quillian, 1967, p 416).

The aspect of the model which lends itself most readily to experimental testing is the assumption of hierarchical storage. It is assumed that such a space saving system can only be achieved at the cost of extra retrieval time. Each node that has to be traversed adds time. Collins and Quillian (1969, 1970a) produced evidence which they interpret as support for this assumption. People were quicker to verify sentences such as A CANARY IS A BIRD than A CANARY IS AN ANIMAL which in turn was verified faster than a CANARY IS A LIVING THING. There have since been a number of studies that question this finding in particular and Quillian's cognitive economy principle in general.

Schaeffer and Wallace (1970) found that it takes subjects longer to decide a word does not belong to a category the closer the word and the category appear to be. E.g. It took longer to reject A DAISY IS A BIRD than GOLD IS A BIRD. Such a finding is difficult for Quillian's model to explain. (see however Collins and Quillian, 1972). Wilkins (1971), Conrad (1972) and Kosslyn and Nelson (1972) have all found measures (e.g. conjoint frequency, association norms) that predict results such as Collins and Quillian's better than the hierarchical model. Other investigators (e.g. Landauer and Meyer, 1972, Wilkins, 1971) have argued that Collins and Quillian confounded hierarchical structure with category size. Landauer and Meyer (1972) further criticize Collins and Quillian for relying on a small number of semantic categories selected from the same domain. Rips, Shoben and Smith (1973) have presented evidence showing that category membership is not an all-or-none relationship but that instances vary in how "close" they are to the superordinate. This finding can only be accounted for in Quillian's hierarchical model in an ad hoc manner. There is thus considerable evidence against Collins' and Quillian's formal hierarchical structure.

However, in a recent model of memory (Anderson and Bower, 1974) Quillian's initial assumption that storage space is limited and that information will be organized to make the best use of this space at the expense of retrieval time is completely reversed. Anderson and Bower assume that in human memory storage space is unlimited but time spent on retrieval is precious.

Although the hierarchical principles of Quillian's model do not appear to be reflected in memory his use of association networks has been imitated in more recent models of memory (e.g. Rumelhart et al, 1972; Kintsch, 1972; Anderson and Bower, 1974). These three models share a common approach in that they assume that knowledge is represented in the form of propositions. A proposition is built up from concepts connected together by labelled associations. Also all three models use a version of Fillmore's case grammar in producing propositions. These models differ from Quillian's model by drawing a sharper distinction between general world knowledge and dictionary information about what a word means. More recently Collins and Quillian (1972) also make this distinction although the implications for their model have not been formally stated.

Out of these three models the only one which has been subjected to extensive experimental testing is Anderson and Bower's model HAM. The other two models cases are largely unproven, although Kintsch (1974) has recently produced evidence consistent with his general assumptions. Anderson and Bower claim that HAM is the only model "...in which the psychological meaning of the networks have been carefully developed. It is not enough merely to construct an intuitively satisfying graph and assert that it represents certain information in memory. Such graphs acquire psychological meaning only after one has addressed himself to the necessary task of defining the functional properties of the network." (1974, p 510).

Recently Smith, Shoben and Rips (1974) have drawn a distinction in models of semantic memory between network models (e.g. Collins and Quillian; Rumelhart et al) and set-theoretic models (e.g. Meyer, 1970; Schaeffer and Wallace, 1970; Clark, 1970). In the set-theoretic models a concept is represented as a unique list of features or attributes. This approach is derived from Katz and Fodor's theory discussed in the previous section.

The difference between the classes of models can be illustrated by the way they treat category membership. In a network model (e.g. Rumelhart et al., 1972) the statement A CANARY IS A BIRD is verified by finding a link between the concept 'canary' and the concept 'bird' which has an ISA label. In the set-theoretic models the attributes belonging to 'canary' would be compared against the attributes of 'bird' to see if a sufficient number matched. Smith et al., (1974) make a further distinction between defining and characteristic features. They present evidence that is difficult for simple network or simple feature models to account for (e.g. Lakoff's (1972) analysis of hedges). However, as noted in the discussion of Katz and Fodor's theory there appears to be problems with the feature approach. It seems implausible that a word's meaning should be restricted to a finite set of features. On the other hand network models do seem to capture this open-ended aspect of meaning. What is perhaps needed is an approach that defines features in a more flexible manner than the Katz and Fodor +MALE , +ADULT kind of feature. There is really no fundamental conflict between feature and network models. As Kintsch (1972) points out, defining a word by specifying the semantic relations that it enters into (as in the network models) ultimately amounts to specifying a word's features. It should be quite possible to produce a model which contains the characteristics of the set-theoretic/feature models in a network system. Indeed as Smith et al. concede Quillian's model is in some respects just such a model. Although undeniably a network model Quillian (1967) says "...what begins as the English definition of a word seems better viewed after encoding as a completely structured bundle of attribute values." (p 421).

More will be said about this problem later.

Models of word recognition

Our main concern here is not directly with how a person's world knowledge is organized but with how such stored knowledge is used in the processing of language. It is assumed that the structure of semantic memory will be an important factor in how the contents of memory interact with other parts of the language processing system. We shall first outline two classes of models of speech recognition and discuss the role of semantic memory within each model.

The models can be divided into active and passive models. The active model is presented by Halle and Stevens (1962) and Liberman, Cooper, Shankweiler and Studdert-Kennedy (1967). The passive model is presented by Norman (1968) and Morton (1969). The research reported here is not directly concerned with this controversy. These experiments concentrate mainly on aspects of Morton's Logogen model. However it is valuable to contrast this class of model with the alternative active models.

The active model is described as a "recognition model in which mapping from signal to message space is accomplished largely through an active or feedback process. Patterns are generated internally in the analyzer according to a flexible or adaptable sequence of instructions until a best match with the input signal is obtained." (Halle and Stevens, 1962, p 155). The main argument in favour of the active model is that there is not a one-to-one relationship between the psychological and the physical events, e.g. although we hear speech as a series of discrete phonemes it is not produced as such. (For a review of the evidence showing the disparity between speech-as-spoken and speech-as-heard see Corcoran, 1971).

What is the role of semantic memory in this model? Halle and Stevens argue against the notion of a dictionary containing word definitions. The role of past knowledge is in guiding the internal generation of the comparison signal. Halle and Stevens describe its operation as follows: "This information is utilized by the control component to formulate strategies that would achieve convergence to the required result with as small a number of trials as possible." (1962, p 157). Exactly how this generation process is guided is not made clear.

There are a number of problems with active model, e.g. how do children understand language before they can talk. Miller (1962) argues that such a model would require an unrealistically high speed of decision making. The active model also has problems in explaining the "cocktail party phenomenon". For a fuller account of these and other problems see Norman (1969).

The passive models have arisen largely out of work on attentional problems such as the cocktail party phenomenon. Central to this approach is the notion of stimulus analysing mechanisms. Stimulus analysing mechanisms are neural units which are sensitive to certain features of the incoming information. Their most important property is that they can combine evidence. Physiological evidence for such mechanisms has been provided by Lettvin, Maturana, McCulloch and Pitts (1959) and Hubel and Wiesel (1959, 1962). Early theories making use of such mechanisms were Selfridge's (1958) Pandemonium model and Sutherland's (1959) pattern recognition model.

In the passive models perception of speech is built up by combining the outputs of a hierarchy of stimulus analysing mechanisms. Norman (1969) summarizes the process as follows: "...information about in-coming signals is abstracted by a number of different analysing mechanisms. As this information is processed by the nervous system the outputs of the analysers may be successively combined, forming a hierarchical process whereby the outputs of one level of analyser are analysed by yet another. Presumably the types of analysers are limited but the ways in which they can be combined are not." (p 38).

Treisman (1964) demonstrated that context could have an effect by biassing tests on the incoming stimuli towards the expected stimuli, (equating tests with stimulus analysing mechanisms). The exact workings of context in such a system are described in greater detail by Norman (1968) and Morton (1969). Norman's and Morton's models are similar in many respects so we shall only discuss in detail Morton's Logogen model here.

Morton has presented his model in a number of papers (e.g. Morton, 1964, 1969, 1970). The outline of the model given here is based mainly on the 1969 paper. Morton (1969) summarises the model as follows:

"The basic unit of the model is the logogen. A logogen is a device which accepts information from sensory analysis mechanisms concerning the properties of linguistic stimuli and from the context producing mechanisms. When the logogen has accumulated more than a certain amount of information, a response (in the present case the response of a single word) is made available. Each logogen is, in effect, defined by the information it will accept and by the response it makes available. Relevant information can be described as the members of the set of attributes (Si), (Vi), (Ai), these being semantic, visual and acoustic sets respectively." (p 165). A diagram of the model is shown in Figure 2.

A logogen simply counts the number of its attributes that are (a) in the stimulus, (b) provided by the context system. The system is solely concerned with number of attributes and makes no distinction between the sources. If the attribute count of a logogen exceeds its threshold the logogen makes its response available to both the semantic system and the output buffer. It is assumed that high frequency words will have lower thresholds than low frequency words. The input to the logogen system from the sensory systems is assumed to be relatively abstract acoustic or visual features which are the products of analysis of the stimulus by "lower order" analyzing mechanisms. The input from the context system is assumed to be in the form of semantic attributes that have been extracted from the previous input to the context system. In terms of the working of the logogen system attributes from context are treated identically to attributes from the sensory analysis. The effect of context-provided attributes on logogens which contain these attributes as part of their defining sets is to

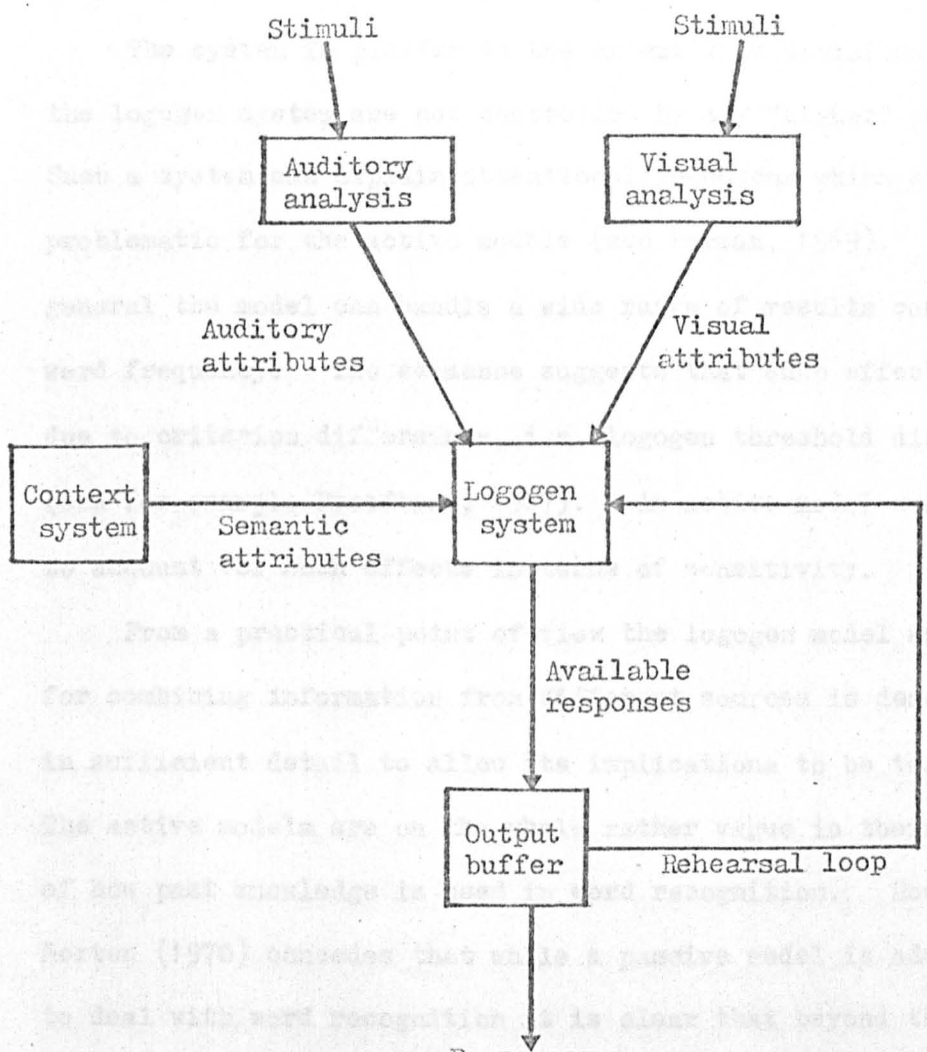


Figure 2. Diagram of Morton's Logogen Model.

reduce the number of attributes from the sensory analysis necessary to reach threshold.

The system is passive to the extent that decisions within the logogen system are not controlled by any "higher" process. Such a system can explain attentional phenomena which are problematic for the active models (see Norman, 1969). In general the model can handle a wide range of results concerning word frequency. The evidence suggests that such effects are due to critérion differences, i.e. logogen threshold differences (see for example Broadbent, 1967). An active model would have to account for such effects in terms of sensitivity.

From a practical point of view the logogen model as a means for combining information from different sources is described in sufficient detail to allow its implications to be tested. The active models are on the whole rather vague in their accounts of how past knowledge is used in word recognition. However, Morton (1970) concedes that while a passive model is adequate to deal with word recognition it is clear that beyond this level speech recognition involves active "constructive procedures". A passive approach is plausible if the to-be-recognized set is finite. Words, although a large set, can be regarded as finite in number. A passive model is inadequate to deal with potentially infinite sets such as sentences. Sentence recognition/ comprehension must involve the use of rules, i.e. an active process.

At the level of word recognition the difference between the active and passive models can be distinguished in the role of attention. In the active models attention is necessary throughout the speech recognition process, including word recognition. In the passive model attention is not required at the level of word recognition. The passive model is consistent with the view of skilled adult reader whereas the active model may be more consistent with the unskilled child reader. It is a characteristic of skilled performance that as much of the task as possible is reduced to an automatic level of processing, i.e. making minimal demands on attention. Goodman's (1967) approach to reading as a skill has already been mentioned (see Introduction). More recently Laberge and Samuels (1974) have discussed the role of automatic information processing in reading. Laberge and Samuels put forward as the goal of fluent reading a state where "...the reader can maintain attention continuously on the meaning units of semantic memory, while the decoding from the visual to the semantic systems proceeds automatically". (p 313).

It has been known for some time that semantic context influences word recognition (e.g. Taylor, 1956; Miller and Isard, 1963; Rubenstein and Pollack, 1963; Tulving and Gold, 1963). It is implied in the papers of Goodman (1967) and Laberge and Samuels (1974) that the more the reader can automatically make use of contextual information the more fluent will be his performance. Out of the models of word recognition Morton's logogen model gives the clearest and most formal account of context operating in just such an automatic manner. Morton's model is also amenable to experimental testing. The present research was an attempt to develop and test the logogen model's account of the role of context in word recognition.

A Brief Outline of the Research in this Thesis.

The intention of the research reported here is to examine the hypothesis suggested by Morton's Logogen Model that the function of contextual information is to facilitate the decoding from the visual to the semantic representations of words.
The research has drawn upon the work of a number of investigators (in particular Meyer and his colleagues) who have not been included in the review. Where a previous finding is closely involved with work reported here it is discussed in the experiments to which it relates.

The first four experiments are concerned with the demonstration, discussion and analysis of a number of context-produced priming effects. Experiments 5 and 6 examine such priming effects in relation to one of the central functions of context, namely the resolution of ambiguity. Experiments 7 and 8 investigate contextual priming effects in processes other than recognition in an attempt to demonstrate the generality of the model presented.

EXPERIMENT ONE

Cued sentence verification:

Introduction:

Quillian (1966) presented a computer model of semantic memory. His model of semantic memory is a highly structured network of concepts, words and images capable of making references and comprehending language. In the model each word is represented by a particular "node" in the network. Each word has stored with it a configuration of pointers to other words. This configuration represents the word's meaning.

The idea of supersets plays a fundamental role in Quillian's model, since this reflects the overall hierarchical organization of the memory system. Quillian proposes that the grouping of concepts into categories saves storage space. Properties that are shared by all members of category need only be stored with the superordinate node. E.g. the fact that a canary is yellow, can sing etc. are defining properties of "canary" and are thus stored directly with the "canary" node. The fact that it can also fly, has wings, has a beak etc. are properties that canaries share with other birds and are stored at the "bird" node. We can infer the fact that a canary has these properties from the knowledge that a canary is a bird.

Although this system of storage is efficient in terms of space it is less efficient in terms of time. Inferences require searches through the network and these processes take time. Collins and Quillian (1969) have presented evidence which is consistent with this concept of the organization of human semantic memory although an alternative explanation of their results has been suggested by Landauer and Meyer (1972).

Search processes are assumed to operate through an "activation" process that starts at a particular node and traces along all of the pointers to the other nodes in the network. The meaning of any given concept is defined in terms of the other concepts to which it is connected. E.g. from the node "canary" the first information retrieved would be that it is a bird, it sings, is yellow and so on. From the node "bird" the search would retrieve the fact that a canary is also an animal, that it flies and has feathers etc. An important assumption is that each node reached will be "activated". In the computer model the node is tagged indicating that it has recently been passed through and also which node led to this tag. This is important for language comprehension. Each word in a sentence starts a search from its node that expands outwards. When the searches intersect, i.e. where one search reaches a node that has already been activated by another search, this will indicate that a path has been found between two nodes, in other words that they are related in some way. This relationship must then be checked against the relationship in the sentence to see that the found relationship is permissible.

Collins and Quillian (1970b) suggested that if such processes do occur in human language comprehension then there is a possible implication for pairs of sentences presented in succession in a sentence verification task. They put forward the hypothesis that there would be a facilitation effect leading to a shorter RT for the second sentence if verification of the second sentence involved using the same fact (following the same path through the network) as the first sentence. E.g. verifying the sentence "A canary is a bird" should reduce RT for "A canary can fly" more than it should reduce RT for "A canary can sing" since the fact that canaries fly is assumed to involve first inferring that a canary is a bird whereas the fact that canaries can sing is not supposed to involve this inferential step. Similarly RT for "A canary is a bird" should be more reduced by verifying "A canary can fly" than by verifying "A canary can sing". Collins and Quillian derived 12 such predictions and found support for 8 of them. They conclude that their results further support the notions that human memory has the same kind of hierarchical organization as the semantic memory in Quillian's Teachable Language Comprehender (1969).

Collins and Quillian propose two possible models to explain how a previous inference might facilitate later retrieval. One model they call the Subway Map model, based on the electric maps in the Paris Metro. They describe this model as follows: "To use the subway maps of Paris a person presses a button for the station that he wants to travel to, and the shortest path, the path of least electrical resistance lights up. In an analogous way, one can imagine that a path in semantic memory lights up connecting the word concepts referred to in the sentence and that facilitation occurs for the future use of that path.", (1970 p 312).

The other model they call the Spreading Activation model. This is similar to Pavlov's (1927) spread of activation theory. In this case excitation spreads from the word-concepts specified in the sentence. The spreading activation theory assumes that facilitation occurs for each surrounding node reached in the intersection process, though facilitation may be greater for the starting nodes themselves. Thus the sentence "A canary is a bird" may facilitate "A canary can fly" more than "A canary can sing" since "bird" is connected directly to "fly" and "canary" but only indirectly to "sing". The spreading activation model predicts facilitation for nodes both on and off the path between the two starting nodes. The subway map model predicts facilitation only for nodes on the path itself. Neither model specifies whether it is the accessibility of the nodes that is facilitated or the transit time between the nodes that is facilitated or both.

Collins and Quillian examine four predictions that serve to differentiate the two models. Three of the four predictions are in the right direction for the spreading activation model but none reach the 0.05 significance level. Thus what tenuous evidence is available supports the spreading activation model.

Further evidence of facilitation of retrieval of information was found by Brenker (1973). Brenker required subjects to verify sentences such as "A horse has a tail" and preceded the sentence by a "primer" for the first concept i.e. the word "horse". Such priming leads to a consistent decrease in verification time. Synonyms used as primers also produced similar effects.

The present experiment was designed to test the two models, proposed by Collins and Quillian, of how facilitation operates. As they indicate, the critical difference between the two models is that the spreading excitation model predicts facilitation for nodes both on and off the path between the two nodes. The subway map model, on the other hand, predicts facilitation only for nodes on the path itself. This experiment used a technique similar to that of Brenker but instead of priming by using the subject of the sentence, the sentence was cued with the superordinate of the subject. E.g. the sentence "A cat has a tail" would be preceded by the cue "Animal". By using sentences in which the property or attribute of the subject was assumed to be stored directly with the subject, according to Quillian's model, it is possible to derive different predictions from the two models. Since to verify the sentences required no knowledge about the subject's superordinate the subway map model would predict that a superordinate cue should produce no facilitation. However, the spreading excitation model would predict facilitation since there is a connection between a word and its superordinate along which excitation could travel.

There are a great many factors which might reasonably be expected to affect verification times for sentences and which may also interact with any effects produced by the superordinate cue. Anderson and Bower (1974) describe the problem as follows: "Such research is fraught with experimental dangers due to the confounding of experimental manipulations with inherent characteristics of the materials. The experimenter is not totally free to choose his experimental materials. He must select from what has been provided by the whims and quirks of natural language and culture. When the experimenter assigns material to conditions on the basis of some semantic criterion he is also probably producing differences between conditions on the basis of word frequency, conjoint propositional frequency and recency, concretences or some other dimension. It thus becomes very difficult to assess the significance of a difference in RT between the conditions. Is it due to the specified change in the semantic variable or is it some unspecified variable that happens to correlate with the semantic variable?" (p 379). Clark (1973) has discussed some of the statistical implications of these problems. By using a suitable experimental design and the appropriate statistics some of these problems may be avoided. In the present experiment several variables might be assumed to play some role will be subjected to post hoc examination. While certainly not a complete list of such possible variables it is hoped that they are some of the more important ones. These variables are frequency of occurrence in the language of both the cue and the subject of the sentence, the initial difficulty of the sentence and the size of the category from which the item was drawn.

Method:

Equipment: The equipment used in this experiment was a specially designed display system. In this equipment were inserted cards with the sentences typed on them. The sentences were covered by a shutter. When E pressed the "start" switch the shutter was lowered displaying the sentence underneath and simultaneously started a stop-clock. The shutter remained down until S pressed one of two response keys marked TRUE and FALSE which also stopped the clock. The tops of the cards were visible above the shutter and on these were typed the particular superordinate cues or the words "NO CUE".

Materials: 48 sentences were used. These were all in forms of simple propositions. As far as possible the relational terms were restricted to "is", "has" and "can". The subjects of the sentences were selected from different categories of the Battig and Montague (1969) and the Brown (1972) category norms.

All subjects and sentences were from the six most frequent instance of their category. The superordinate cues were the names given to the categories by Battig and Montague and Brown. Bearing in mind the distinction made by Collins and Quillian between properties that would be stored with the exemplar node and those properties that would be stored with the superset node all the true sentences contained properties that were assumed to be specific to the particular instance and would not be stored with the superset.

Half the sentences were true and half were false. Half the sentences were cued and half were not cued. These two factors were combined so that there were 12 cued true sentences, 12 cued false, 12 not cued true and 12 not cued false.

Examples of the sentences are shown in Table 1.

Table 1. Examples of sentences

TRUE SENTENCE	CUE	FALSE SENTENCE	CUE
Cars have wheels	Vehicle	Apples are blue	Fruit
Dogs can bark	Animal	Vets cure people	Profession

Subjects: 32 first year undergraduates acted as subjects. These were divided into two groups of 16. The 24 sentences which were cued and the 24 sentences which were not cued were reversed for the two groups so that each sentence occurred an equal number of times in the cued and not cued condition. The subjects participated to fulfil a course requirement.

Procedure: S sat facing the tachistoscope with a finger from each hand on two buttons. The right hand button was marked TRUE and the left hand button was marked FALSE. In each trial S read aloud the cue of the words NO CUE from the top of the card. E then displayed the sentence on the card and S responded by pressing true or false key. That card was then removed revealing the next cue on the top of the next card. The order of presentation of the cards was random. Each experimental session lasted approximately 15 minutes. 12 practice trials were carried out before the experimental trials.

Results: The mean RT for each condition is shown in Figure 3 The results were analyzed by calculating quasi F ratios in which both subjects and materials are treated as random variables. (see Clark, 1973). Superordinate cues produced a mean reduction in verification time of 66msecs (min. $F'(1,57) = 7.8$, $p < 0.01$). True sentences were on the average 175 msecs faster than false sentences (min. $F' = (1,57) = 12.6$, $p < 0.01$). The interaction between cueing and true/false was not significant by the quasi F test, although it was significant by the less conservative analysis by subjects ($F(1,31) = 6.13$, $p < 0.025$).

	Mean	S.D.
Cued - True	1366	207.5
Not Cued - True	1582	217.8
Cued - False	1473	211.4
Not Cued - False	1607	219.3

Mean and standard deviations for Figure 3.

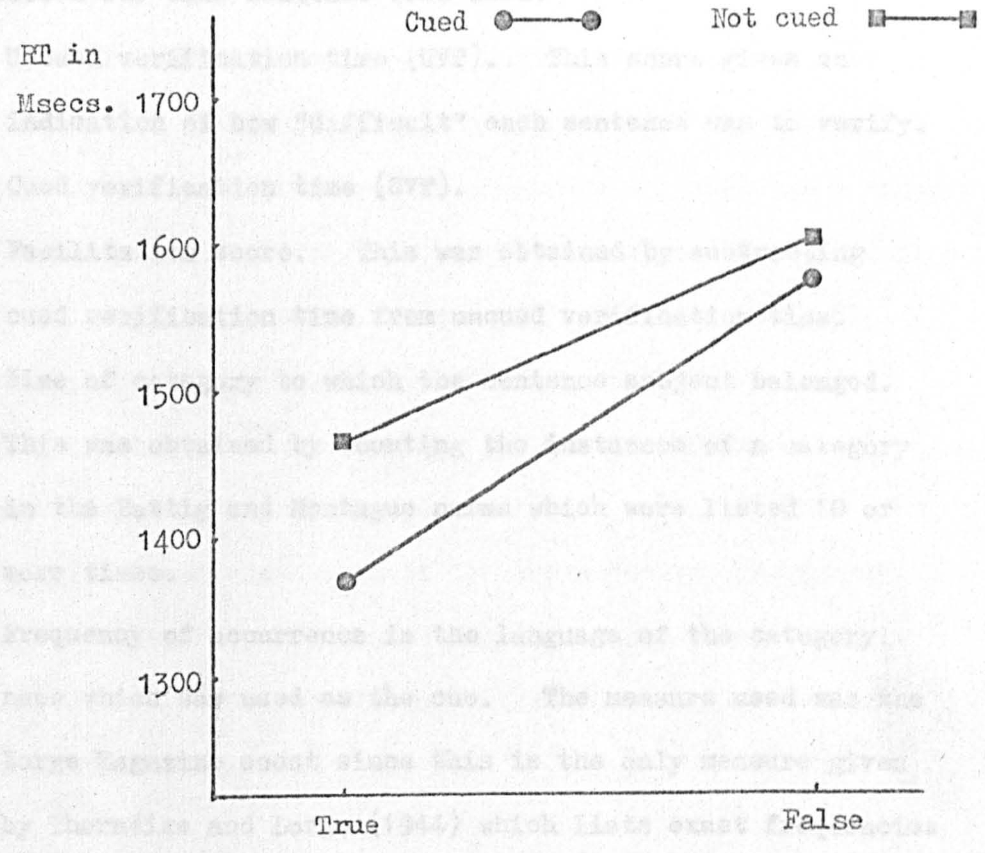

Cued
Not cued
RT in
Msecs. 1700
1600
1500
1400
1300
True
False
Mean verification time.

Figure 3.

Overall error rate was 11% but there was no significant difference in error rates between conditions.

<u>Analysis of Extra-experimental variables</u>:

Six scores for each sentence were used:

1 Uncued verification time (UVT). This score gives an indication of how "difficult" each sentence was to verify.

2 Cued verification time (CVT).

3 Facilitation score. This was obtained by subtracting cued verification time from uncued verification time.

4 Size of category to which the sentence subject belonged. This was obtained by counting the instances of a category in the Battig and Montague norms which were listed 10 or more times.

5 Frequency of occurrence in the language of the category name which was used as the cue. The measure used was the Lorge Magazine count since this is the only measure given by Thorndike and Lorge (1944) which lists exact frequencies for all words.

6 Frequency of the subject of the sentence. (Also the Lorge count).

The first three scores can be regarded as "experimental" in that they derive from the present experiment. The last three can be regarded as extra-experimental scores. True and

False sentences were analysed separately. Unfortunately the frequency data was not available for one of the true sentences reducing the number of true sentences to 23. Spearman's Rank Correlation Coefficients were calculated for each pair of scores. See Table 3 for true sentence correlations.

Kendall Rank Correlations and Partial Correlations were also calculated for each pair except for cued VT which only correlated with uncued VT, and frequency of the sentence - subject which did not correlate with any of the other scores. Kendall's (τ) for the remaining pairs is shown in Table 4. Table 5 shows the partial correlation coefficients holding size of category constant. Table 6 shows partial correlations holding frequency of the category name constant.

For false sentences none of the three "extra-experimental" scores correlated significantly with any of the three "experimental" scores. The correlation matrix for the three experimental scores is shown in Table 7.

	UVT	CVT	Facil	Size of Cat.	F of Cat. Name	F of Subject of sentence.
UVT		0.696**	0.45*	0.22	0.34	-0.06
CVT			0.0	-0.12	0.13	-0.02
Facil				0.63**	0.37*	-0.05
Size of Cat.					0.31	-0.1
F of Cat. Name						0.05
F of sent. subj.						

* = $p < 0.05$ ** = $p < 0.01$

True sentences. Spearman's Rank Correlations for all pairs.

TABLE 3

	UVT	Facil	Size of Cat.	F of Cat. Name.
UVT		0.38	0.21	0.31
Facil			0.46	0.26
Size of Cat.				0.29
F of Cat. Name				

True Sentences. Kendall's Rank Correlations.

TABLE 4

	UVT	Facil	F. of Cat. Name
UVT		0.32	0.26
Facil			0.15
F. of Cat. Name			

True Sentences. Kendall Partial Correlations Holding Size of Category Constant.

TABLE 5

	UVT	Facil	Size of Cat.
UVT		0.32	0.13
Facil			0.42
Size of Cat.			

True Sentences. Kendall Partial Correlations Holding Frequency of Category Name Constant.

TABLE 6.

	UVT	CVT	Facil
UVT		0.698	0.42
CVT			-0.342
Facil			

False sentences. Spearman Rank Correlations.

TABLE 7.

Discussion: These results provide further evidence of facilitation effects in retrieval from memory. Furthermore they reject the subway map model proposed by Collins and Quillian. Superordinates are not on the path between a member of a category and its characteristic properties and therefore, according to the subway map model, cannot produce any facilitation. Out of the models proposed by Collins and Quillian this leaves the spreading excitation model as the only one consistent with these results. There is, however, a third model, not discussed by Collins and Quillian, that could also predict these results. Meyer, Schvaneveldt and Ruddy (1972) have suggested a model of memory retrieval which makes the same predictions for this experiment as the spreading excitation model. Meyer et al. call this model the shifting location model. According to this model memory is seen as similar to a reel of magnetic tape or a magnetic disk. Information is retrieved by means of a fixed "reading-head". The time taken to retrieve a piece of information depends on how far the tape or disk has to move so that the information is under the reading-head. Information on the tape or disk is organized so that related topics are found in the same area. Given a cue word (e.g. "animal") the tape /disk can move until the appropriate area is under the reading mechanism. When presented with "A cat has a tail" the tape/disk has less far to move to find the relevant information that enables verification of the proposition than if it had no cue.

The results of the present experiment are inadequate to distinguish between the spreading activation model and the shifting location model. Meyer et al. present evidence which gives more support to the spreading activation model than the shifting location model. The problem of distinguishing between these models will be dealt with more fully in the next chapter.

The present experiment is also unable to answer the question posed by Collins and Quillian as to whether it is the accessibility of the nodes that is facilitated or the transit time to move between nodes, or both. This problem is dealt with in greater detail in the next section.

What can be concluded about sentence differences from the post hoc correlations? As noted earlier a number of significant post hoc correlations were found for true sentences but not for false sentences. False sentences tended to be slower, show considerably less facilitation (25 msecs versus 107 msecs) and to have much higher variance than true sentences. Also false responses were made with the left (i.e. largely non-dominant) hand. These factors may account for the lack of significant correlations in the post hoc analyses. This discussion will concentrate on true sentences.

Consider first uncued verification time (UVT). Not surprisingly UVT correlates highly with cued verification time (CVT). The correlation between UVT and facilitation reveals a possible "floor" effect. The slower a sentence is when uncued the more the facilitation it receives from the cue. This may simply be that faster sentences have less room for improvement or it may reflect some interaction between "difficulty" of the sentence (as measured by verification time) and the effectiveness of the cue. It is not easy to identify a priori what constitutes an "easy" or a "difficult" sentence. For instance, one might expect the frequency of the constituent words to be an important factor. However, UVT does not correlate with frequency of the subject of the sentence. This is rather surprising in view of the large body of evidence showing an inverse relationship between word frequency and recognition time (e.g. Broadbent, 1967). This suggests that recognition of the words in the sentence is a relatively minor part of the verification task. It seems that it is the relationship between the subject and its property that is more important.

UVT correlates positively but not significantly with the size of the category to which the sentence-subject belongs. The partial correlation (see Table 6) indicates that much of this correlation can be attributed to the frequency of the category name (larger categories tending to have more frequent names). The correlation between UVT and frequency of the category name is at first sight rather puzzling. Why should VT be related to the frequency of the category name when that word is not present? One hypothesis, consistent with the spreading activation model, is that the more frequent the category name the more likely it is to be activated when a member of the category is presented. (An assumption of Morton's logogen Model is that the threshold for any word is related to its frequency). E.g. flower names may be more strongly associated to their superordinate "flower" than to any of their properties. The most salient fact about a rose may be that it is a flower rather than that it has thorns. When presented with "A rose has thorns" the fact that a rose is a flower may be retrieved before the fact that it has thorns. The availability of this fact may hinder the accessing of other facts, leading to a slower VT. It is difficult to derive a comparable hypothesis from the location shifting model.

Facilitation scores also correlate with size of category and frequency of the category name. Here, however, it seems that it is the size of the category which is the more important factor. The instances from bigger categories have higher facilitation scores than those from smaller categories. This appears to rule out any conscious guessing strategy on the part of the subjects. Given a cue as to which category the following sentence will belong to, a guessing strategy should be more likely to produce the actual instance the smaller the category. This is the opposite to the result obtained here. It should be noted that all the categories used contain at least 20 common members. Different results may be obtained if much smaller categories were used (e.g. months of the year). Presumably the larger the category the more useful it is to have a label for the class as a whole. This correlation may reflect that the bigger categories tend to be "better" labelled and that these labels tend to be of greater help in accessing a category. This is similar to the suggestions of Sapir and Whorf that classes which are important to a culture will be more differentiated, i.e. members of the culture will be able to distinguish more instances of important classes. This would be reflected in normative data on category size. Thus bigger categories will be those that are more important to a culture and more likely to receive a well defined class name. This may be related to the effectiveness of a superordinate name as a cue.

<u>Conclusions:</u> Superordinate cues facilitate verification of sentences, even when verification requires no knowledge of the category to which the instance belongs. Out of two models proposed by Collins and Quillian only the spreading excitation model is consistent with these results. There is, however, a third model, the shifting location model, which makes the same predictions as the spreading excitation model.

The correlation between uncued verification time and frequency of the category name is more easily explained in terms of a spreading activation model of memory search than a shifting location model.

Sentences concerning instances of larger categories tend to receive more benefit from a superordinate cue. This may reflect the fact that larger categories are "better" labelled.

Superordinate and subordinate cues in a lexical decision task:

EXPERIMENT TWO

Introduction:

Exp. 1. provided evidence that superordinate cues can facilitate sentence verification. One of the major questions left unanswered is whether the superordinate cue produced its facilitation by increasing the accessibility of the nodes (the words' "locations" in memory) or by speeding up the actual verification part of the task. It was argued that the lack of any correlation between frequency of occurrence in the language and uncued sentence verification time suggested that word recognition was a minor part of the verification task since there is considerable evidence that word frequency is related to recognition time. On the other hand there is a large body of evidence that context influences word perception. Rubenstein and Pollack (1963) regarded verbal context as a constraint on the probability of a given word's occurrence and showed that intelligibility is a simple power function of probability of occurrence. Similarly Miller, Heise and Lichten (1951) found that words drawn from a set of 2 alternatives required a signal-to-noise ratio of 24db less than the same words selected from a set of 1000 alternatives to achieve the same level of intelligibility. It may be that the effect of the cue in Exp. 1 was to reduce the number of possible alternatives for the subject of the sentence, lowering its recognition threshold. Tulving and Gold (1963) describe the situation as follows: "It is reasonable to assume that different sources of information are complementary to one another in the sense that if one source provides much information then less information is needed from other sources." Thus given a cue indicating the possible set of alternatives the subject is prepared to accept the occurrence of one of the set on the basis of less evidence than if he had no cue. Requiring less information to make his decision will presumably mean the subject can make his decision more quickly. This description implies a conscious strategy on the part of the subject. Morton (1969) has described such an interaction more formally in his logogen model of word recognition. In this model the interactions of information from different sources occurs automatically in a hypothetical word-recognizing device Morton calls a logogen.

The verification task used in Ex. 1, although producing a facilitation effect, is not very suitable for a more thorough investigation of how context affects word recognition. It inevitably involves recognition of several words and verification itself, is a complex process that is poorly understood at present. A more appropriate task is provided by the lexical decision task, which has recently been used by a number of investigators. (e.g. Landauer and Freedman, 1968; Meyer and Ellis, 1970; Rubenstein, Garfield and Millikan, 1970). In the lexical decision task the subject has to decide whether a string of letters forms a real word or not.

Meyer and his co-workers (e.g. Meyer and Schvaneveldt, 1972) have produced evidence of facilitation in recognizing pairs of words. Subjects were quicker to decide that BUTTER was a word if they had previously made a decision about an associated word such as BREAD than if the word was unassociated e.g. NURSE.

It is assumed here that these effects reflect the underlying organization of the lexical memory which contains the information a person has stored about the words he knows. The model assumed here is the same as that assumed by Meyer and Schaneveldt. The model includes two assumptions made by a number of other investigators (e.g. Norman, 1968; Morton, 1969; Collins and Quillian, 1969; Meyer, 1970; Rumelhart, Lindsay and Norman, 1972). Meyer, Schaneveldt and Ruddy (1973) describe these assumptions as follows: "The first assumption is that words are stored at distinct "locations" in lexical memory and the memory is organized semantically, so that in some sense, associated words like BREAD and BUTTER are relatively close together, whereas unassociated words like NURSE and BUTTER are further apart. The second assumption is that accessing information from a given memory location produces residual neural activity that spreads to other nearby locations. This temporary increase of excitation then produces the faster recognition of associated words." As was noted in Exp. 1. the shifting location model makes a number of similar predictions as the spreading excitation assumptions.

Meyer et al. have concentrated mainly on identifying where in the word-recognition process contextual effects operate. They have used materials drawn from association norms and have not specifically investigated the effects obtained with different kinds of associations. Recent models of semantic memory imply that certain kinds of relations may be basic to the organization of semantic memory. Collins and Quillian's model relies heavily upon the superset relation as a means of efficiently storing information. Similarly Rumelhart, Lindsay and Norman make frequent use of the ISA relation, although they place less emphasis on the hierarchical structure of memory than do Collins and Quillian. (see also Sanford and Seymour, 1974).

The aim of the present experiment was to see whether different kinds of associations would produce different facilitation effects. Simple associated pairs from association norms would be compared with superordinate-subordinate and subordinate-superordinate pairs. According to a hierarchical model of memory these latter two types of pairs should be stored close together and thus produce facilitation.

A subsidiary aim was to examine the effects of a cue that was phonemically similar but neither graphemically nor semantically similar to the test word (e.g. CALF-LAUGH). Meyer, Schvaneveldt and Ruddy (1973) examined the facilitation effects between words that were either graphemically and phonemically similar (e.g. BRIBE-TRIBE) or phonemically dissimilar but graphemically similar (e.g. COUCH-TOUCH). If the two words were both graphemically and phonemically similar then recognition of the second word was facilitated. If, however, the two words were graphemically similar but phonemically dissimilar then recognition was inhibited. Meyer et al. propose a model of visual word recognition to explain these results. The model assumes that there is an initial encoding stage where grapheme-phoneme correspondence rules are applied to form a phonological representation of the word. A lexical decision is made by accessing memory to determine whether or not the representation has been stored there previously. If it is not found and if the string has more than one possible representation then the encoding and decision operations are repeated. The results obtained by Meyer et al. can be explained in terms of encoding biasses. If the graphemic encoding stage detects similarities between the first and the second word then the phonemic encoding stage is biassed towards applying the same rules to the second word as it used for the first. Thus graphemic and phonemic similarity would facilitate recognition of the second word since less time need be spent applying the grapheme-phoneme conversion rules. This would be particularly important where the second word has more than one possible pronunciation. Inhibition would occur for graphemically similar but phonemically dissimilar pairs, since the second word would receive the wrong phonological representation. When the search fails to find this representation in lexical memory the word is recoded phonemically and the search repeated. This coding and search will add to recognition time.

As the model stands it is not possible to predict the effects of phonemic similarity and graphemic dissimilarity. It is necessary to make one of two assumptions, called here the weak and the strong assumption. The weak assumption is that the phonological coding bias only occurs when the graphemic encoding stage detects some similarity between the two words. In this case there will be no effect of phonemic similarity and graphemic dissimilarity. The strong assumption is suggested by Meyer et al. It assumes that differences in graphemic structure could bias the phonemic encoding stage to form dissimilar representations of the two words. This would lead to inhibition in the same way that graphemic similarity and phonemic dissimilarity does.

<u>Method</u>.

<u>Equipment</u>: The equipment used in this experiment was a one channel tachistoscope, a millisecond stop-clock and a voice key.

Materials: The materials used were 60 real English words and 60 pseudowords. Each word and pseudo word was matched with a cue word. The real words were divided into 5 groups of 12 words, according to the relationship between the test word and its cue. These relationships were:

1 Cue superordinate to test word. The cue word was the most frequently given superordinate in the Loftus and Scheff categorization norms (1971). E.g. BIRD - ROBIN.

2 Cue subordinate to test word. Cue word was one of the six most frequently listed instances of the category in the Battig and Montague category norms (1969). E.g. VIPER - SNAKE.

3 Cue was a frequent associate of test word according to Palermo and Jenkins word association norms (1964), but was neither a superordinate or a subordinate of the test word. E.g. BREAD - BUTTER.

4 Cue word was phonemically similar but graphemically dissimilar to test word. Also cue word was not listed as an associate of the test word. E.g. AIR - CARE.

5 Cue word was not an associate of the test word, i.e. not listed in the word association norms as a frequent associate. E.g. PATH - QUEEN. These were also phonemically and graphemically dissimilar.

As far as possible test words in each condition were matched for length and frequency.

The 60 pseudowords were formed by replacing one letter (vowel with vowel, consonant with consonant) of frequent English words (A or AA words in Thorndike-Lorge Count, 1944). Changes were made so that the resulting pseudowords conformed to English phonological and spelling rules. Each pseudoword was paired with a real English cue word, drawn from the same sources as the cue words for the real English test words.

The 120 word and pseudoword pairs were randomized. The cue words were then printed in a small booklet with one cue on each page.

Subjects: 12 undergraduates participated in the experiment as subjects to fulfil a course requirement.

Procedure: The subject sat facing the tachistoscope with the booklet of cue words on the table in front of him. The voice key was also positioned on the table in front of the tachistoscope. On each trial the subject read aloud a cue word out of the booklet. As soon as he had read the word-cue he looked into the tachistoscope and the experimenter pressed a button which displayed the test word /nonword and started the clock The subject was instructed to say "yes" if the item was a real English word and "no" if it was a pseudoword. He was instructed to respond as quickly as possible but without making mistakes. The subject's verbal response stopped the clock and terminated the display. The subject then turned to the next page of the booklet and read the next cue word. The experimenter recorded the reaction time and the response.

Ten practice trials were given. Each session lasted approximately 20 minutes.

<u>Results</u>: The mean RT for each condition is shown in Figure 4.

An analysis of variance was carried out using only data from real English test words. Quasi F ratios were calculated treating both subjects and materials as random effects. (see Winer, 1970; Clark, 1973).

$$\min F' \ (4,93) = 5.6 \ p < 0.01.$$

(N.B. Degrees of freedom depend on size of mean square errors).

Overall error rate was 13% but there was no significant difference between conditions. Only data from correct responses were used in computing statistics. Comparisons of each pair of conditions are shown in Table 8.

	Super-ordinate	Sub-ordinate	Associate	Non-associate	Phonemically similar
Superordinate		N.S.	N.S.	0.01	0.01
Subordinate			N.S.	0.01	0.01
Associate				0.01	0.01
Non-associate					N.S.
Phonemically Similar					

Pairwise Comparisons (Newman-Keuls)

<u>TABLE 8</u>

Relation of cue word to test word.	Mean	S.D.
Superordinate	693	103.9
Subordinate	692	102.2
Associate	697	106.7
Phonemically similar	762	114.1
Non-associate	757	108.5
Pseudoword	781	93.8

Means and standard deviations for Fig.4.

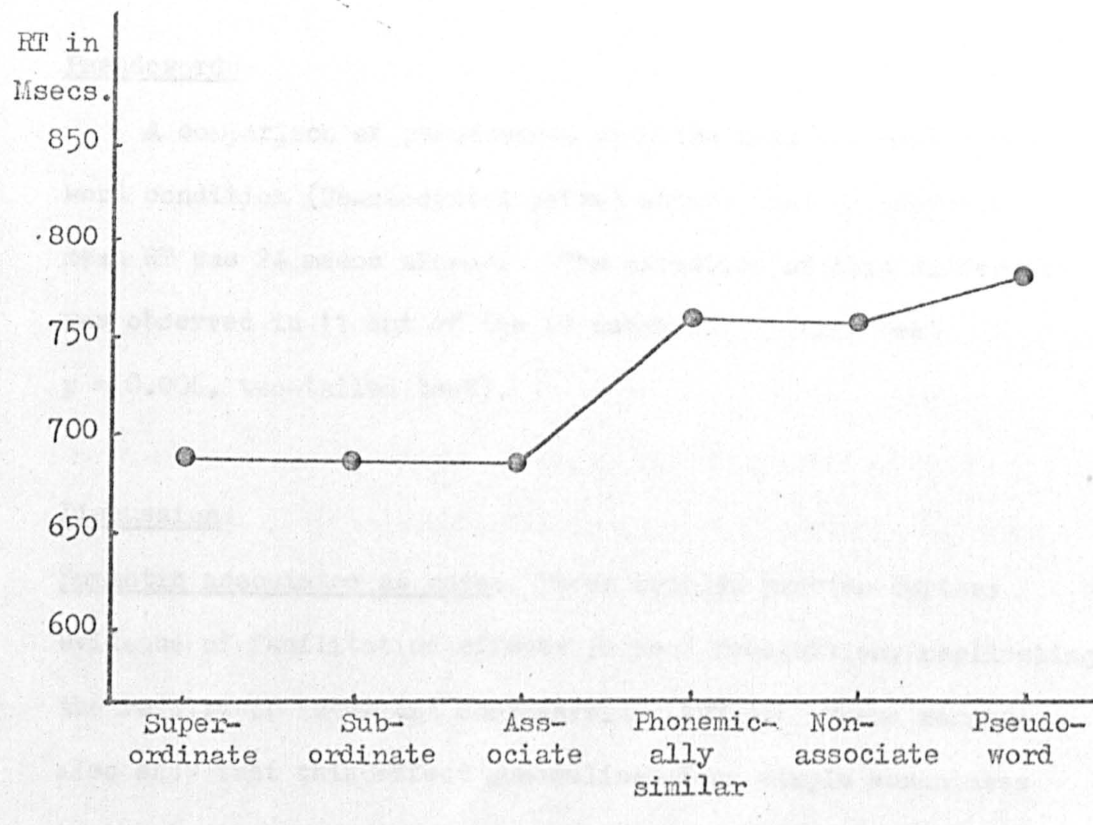
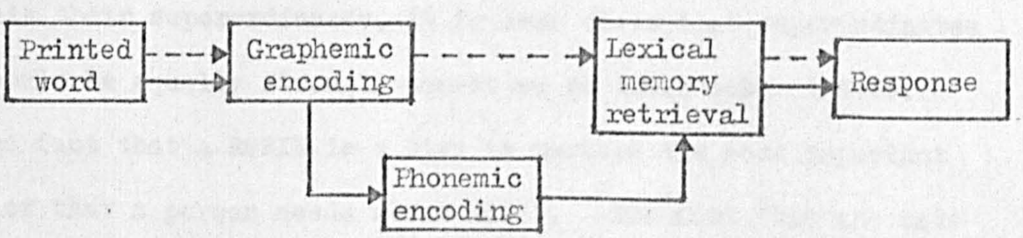


Figure 4.



Dual encoding model of word recognition.
(From Meyer et al., 1974).

Figure 5.

Pseudowords:

A comparison of pseudowords with the most comparable real word condition (Unassociated pairs) showed that pseudowords mean RT was 24 msecs slower. The direction of this difference was observed in 11 out of the 12 subjects. (sign test $p = 0.006$, two-tailed test).

Discussion:

Semantic associates as cues: These results provide further evidence of facilitation effects in word recognition, replicating the results of Meyer and Schvaneveldt (1972). These results also show that this effect generalizes from simple associates to the logical relationships of subordinate and superordinate, although these terms may not be highly frequent associates of the test word as indicated by association norms. All three related conditions of superordinate, subordinate and associated pairs were significantly faster than the unassociated pairs. It is interesting to note that the three related conditions did not differ from each other. While it is to be expected that subordinates would be relatively strongly associated with their superordinates, it is less clear that superordinates should be equally strongly connected to their subordinates. The fact that a ROBIN is a bird is perhaps the most important fact that a person needs about ROBIN. The fact that the category BIRD includes ROBIN is less salient. It should be pointed out that the members of the categories used were all highly "representative" members in the sense used Heider (1973) e.g. ROBIN is judged as being a better representative of BIRD than is CHICKEN. (See also Smith, Shoben and Rips, 1974).

These results suggest then that grouping into classes is an important principle in the organization of semantic memory.

As far as the problem from Exp. 1 as to whether the superordinate cues facilitated access to the subordinate nodes or decreased the transit time between the nodes, is concerned, the results of the present experiment are suggestive but not conclusive. In this experiment words primed with their superordinates were 64 msecs. faster than words preceded by unassociated words. This compares with the 66 msecs. overall facilitation effect of superordinate cues in Exp. 1. Since in Exp. 2. the only possible effect is on word recognition time this suggests that the facilitation in Exp. 1. was similarly caused by increased accessibility of the individual word nodes.

<u>Phonemically similar words as cues</u>: These results support the weaker of the two hypotheses put forward as possible predictions from Meyer et al's. visual word recognition model. That is, the phonemic encoding bias only occurs when the graphemic encoding stage detects some similarity between the priming word and the test word. There was no evidence that different gra-

phemic codes biasses the phonemic encoding system to produce different (in this case "wrong") phonological representations.

These arguments are assuming that visual word recognition involves a grapheme-phoneme recoding stage, as in Meyer et al's model. A number of investigators (e.g. Bower, 1970; Kolers, 1970; Baron, 1973) have argued that visual word recognition can occur directly from the graphemic code, without any intervening phonological encoding. The present result is also consistent with this model since any phonemic similarity between the words would be irrelevant for deciding if they were words or not. Meyer et al. (1974) have proposed a parallel race model, where both a graphemic and phonemic code are used in word recognition. (see Fig. 5). Both codes can find a word in lexical memory and sometimes one and sometimes the other will be quicker. Since the phonemic code involves an extra stage i.e. the grapheme-phoneme conversion which will presumably take time (c.f. Sternberg, 1969), such a model implies that a phonological representation is a more efficient code for lexical memory retrieval. Otherwise the graphemic-code-based search would always win the "race" and hence a phonological code would be useless. If it is true that a phonological representation is advantageous for finding a word in lexical memory it is reasonable to expect subjects to adopt this strategy (assuming they have some control over the process). One is left with the original conclusion that there is a phonemic encoding bias only if there is graphemic similarity. Meyer et al (1974) point out that the codes which help in recognizing printed words may depend upon the type of task involved. There remains therefore the possibility that subjects may have only used graphemic encoding, in which case phonemic similarity would have no effect, as was found.

As can be seen from the preceding discussion, a dual encoding model makes it difficult to reach any strong conclusions, using the lexical decision task. It is possible that the situation may be clarified by using a task where the subject has to make a grapheme-phoneme conversion.

This is examined in Exp. 3.

Pseudowords: Pseudowords were significantly slower to be rejected as words than real words were to be accepted. This finding has been reported widely in the literature (e.g. Meyer and Ellis, 1970; Rubenstein, Lewis and Rubenstein, 1971) although the size of the difference reported here is less than is often reported by other investigators. This finding is consistent with an exhaustive serial scan model of lexical memory. If memory search operates by examining in a serial fashion all locations to find a match for the letter string then pseudowords would take longer than real words since all locations would have to be examined. This suggests extremely fast search rates. According to an estimate by Oldfield (1966) somewhere between 55,000 and 70,000 locations would have to be examined in approximately $\frac{1}{2}$ second. This seems implausibly high but it is difficult to answer the question of when serial search rates become "too fast".

A number of investigators (e.g. Oldfield, 1966; Swanson and Wickens, 1970; Rubenstein, Garfield and Millikan, 1970) have explained frequency effects in word recognition by assuming a serial scan in which frequently accessed locations are examined first. According to this view pseudowords are treated as highly infrequent words.

An alternative model is suggested in Meyer and Ellis (1970). As a result of their investigations Meyer and Ellis propose that recognizing a string of letters as a word does not depend on searching "all or even a significant part of the words stored in memory." According to this model visual and/or acoustic features of a string of letters are used to compute an "address" in memory (c.f. Norman, 1969; Atkinson and Schiffrin, 1968). This address may then be used to check a location in memory to see if the letter string has occurred in the past. The system is able to compute an address for a pseudoword but examination of the corresponding memory location will find it "empty" or "meaningless" in the literal sense. According to this model frequency effects are assumed to be dependent upon the time taken to compute the address. Computation time for the address rather than memory search are assumed to be related to frequency of the word, frequent words addresses being computed more quickly than infrequent words. Pseudowords being highly infrequent take a long time in the address computation stage. (See also Herriot, 1974).

This view is similar to Morton's Logogen Model (1969), if it assumed that logogen is an address computing device. The logogen model, however, in its simplest form is incompatible with the view expressed above that the system can compute an address in lexical memory, even for pseudowords. It would be necessary to assume an infinite number of logogens to deal with an infinite number of possible pseudowords. The logogen model can account for the slower decision times for pseudowords if it is assumed that the system waits a certain length of time after the input of a letter string for an output from a logogen. If after this set interval there has been no output from a logogen the system responds "non-word". This view differs from that of Meyer and Ellis by assuming a failure of address computation for nonwords rather than the discovery of an "empty" location in lexical memory. This view must be an oversimplification of the system since there must be a capacity for entering new words into lexical memory. The logogen model as it stands at the moment makes no provision for handling new information.

One way to test whether pseudowords' longer decision times are due to some early address computing stage or some memory search stage is to examine them in a task which does not specifically demand an access of lexical memory. (See Exp. 3.).

General Discussion:

The present experiment makes no test of the two models proposed in Exp. 1. to account for the associative priming effects. Both the spreading activation and the location shifting models make the same predictions. The problem of deciding between these models will be dealt with in a later section. (See Exp. 4.).

Neither does the present experiment give any indication as to where in the visual word recognition process these facilitation effects are operating. It is possible that context could influence either an encoding stage or a search stage. It is possible to examine this more closely by using a task involving no explicit memory search. This is dealt with in Exp. 3.

It is worth noting that the error rate is high compared to that reported by investigators using a two-button choice technique (e.g. See Meyer and Schvaneveldt, 1972). It may be that the use of a voice key and verbal response is somehow more "artificial" than the button pressing task. However, the use of a verbal response is useful in that it enables a more direct comparison between the results of Exp. 2. and those of Exp. 3.

It must also be pointed out that there are a number of methodological and technical problems with this experiment.

These will be discussed more fully in the next section since Exp. 3. used a similar design and largely the same materials.

Conclusions:

This experiment provides evidence that contextual priming effects occur not only for simple associates but also for superordinate and subordinate terms. This is interpreted as support for models of semantic memory which stress the importance of categorization.

Evidence also indicates that phonemic similarity is insufficient to produce facilitation without graphemic similarity.

The relative slowness of rejection of pseudowords is interpreted as reflecting either slowness in computing an address in memory for pseudowords or failure to compute an address at all (logogen model) rather than time taken in searching all possible memory locations.

Superordinate and subordinate cues in a pronunciation task:

EXPERIMENT 3.

Introduction:

In Exp. 2. three results were discussed. Various semantic relations and associations presented as cues were shown to facilitate word recognition. Phonemically similar but graphemically dissimilar words were found not to produce any facilitation of word recognition. Pseudowords were found to take longer to be rejected than real words were to be accepted.

Each of these findings left at least one question unanswered. In the case of semantic associates as cues the problem remained as to which stage in the proposed model was influenced by the cue. Either an early encoding stage or the lexical memory retrieval stage or both these stages could be affected.

In the case of the phonemically similar cues the dual-encoding model allows for the possibility of subjects recognizing words directly from their graphemic representation in which case phonemic similarity between cue and test word would have no effect. Alternatively subjects could be using a phonemic code but no bias occurs for phonemic encoding without graphemic similarity.

It is unclear whether the slowness of rejecting pseudowords is due to as exhaustive serial scan of lexical memory or to slowness in computing an address in lexical memory.

All these problems can be investigated using a task where the subject simply has to pronounce the strings of letters rather than deciding if they form a real word or not. According to Meyer, Schvaneveldt and Ruddy (1974), the pronunciation task and the lexical decision task share a common grapheme-to-phoneme stage but that they differ in terms of the other processes involved. In particular the lexical decision task necessitates retrieval from lexical memory whereas the pronunciation task is assumed not to need access to lexical memory. If one can find similar facilitation effects for semantic associations in both tasks then it can be assumed that in both cases it is the grapheme-to-phoneme stage that is being influenced and not any retrieval from lexical memory stage.

The pronunciation task explicitly demands that the subject makes a grapheme-to-phoneme conversion therefore it is reasonable to expect that any phonemic encoding bias as a result of phonemic similarity would show up here, although as suggested in Exp. 2. any bias may only result if there is graphemic similarity as well.

If one assumes that the hypothesized address computation discussed above largely consists of producing an internal graphemic and phonological representation of the letter string, then the address computation stage of the lexical decision task will also be present in the pronunciation task. However, since in the pronunciation task there is no need to access lexical memory any slowness in pronouncing pseudowords cannot be attributed to any search of items in lexical memory.

Method:

Equipment: The equipment used in Exp. 3. was the same as used in Exp. 2. A one channel tachistoscope, a voice key, and a millisecond stop-clock.

Materials: These were the same 60 pairs of English words and cues as used in Exp. 2. divided into the same 5 groups of 12, according to the relationship between the cue and the test word.

To recap these were:

1 cue superordinate to test word.

2 cue subordinate to test word.

3 cue simple associate of test word.

4 cue phonemically similar but graphemically dissimilar to test word.

5 cue not associated to test word - also both graphemically and phonemically dissimilar

12 of the pairs of pseudowords and cues used in Exp. 2. were also included.

Subjects: 10 undergraduates participated as subjects to fulfil a course requirement. They were from the same pool as the subjects used in Exp. 2. but none of the subjects in Exp. 3. had participated in Exp. 2.

Procedure: As in Exp. 2. the subject sat facing the tachistoscope and voice key with a booklet containing the cue words in front of him. After reading a cue word out loud the subject looked into the tachistoscope and the test word/pseudoword was displayed. Simultaneously the clock was started. The subject was instructed to pronounce the letter string as quickly as possible but to make sure that he used the "correct" pronunciation. The subject's verbal response stopped the clock and terminated the display. Ten warm-up trials were given. Each session lasted approximately 15 minutes.

Results: A flexible criterion of "correctness" was used. As long as the pronunciation conformed with a possible application of English phonological rules, the data were included. In practice subjects had little trouble in pronouncing either the real words or the pseudowords.

The mean RTs for each condition are shown in Figure 6. The results of Exp. 2. are also included for comparison.

An analysis of variance was carried out for the real word data. Analysis over subjects $F(4,36) = 5.6$ $p < 0.01$. Analysis over materials $F(4,55) = 2.81$ $p < 0.05$. The quasi F ratio was not significant min F' $(4,90) = 1.9$ $p > 0.05$.

Relation of cue word to test word.	Mean	S.D.
Superordinate	629	130.8
Subordinate	633	137.4
Associate	636	180.3
Phonemically similar	680	151.7
Non-associate	683	170.4
Pseudoword	827	254.4

Means and standard deviations for Fig.6.

PRONUNCIATION TASK.

(FOR LEXICAL DECISION TASK SEE FIG. 4).

		Mean	S.D.
Lexical decision:	Real words	757	108.5
	Pseudowords	781	93.8
Pronunciation:	Real words	683	170.4
	Pseudowords	827	254.4

Means and standard deviations for Fig.7.

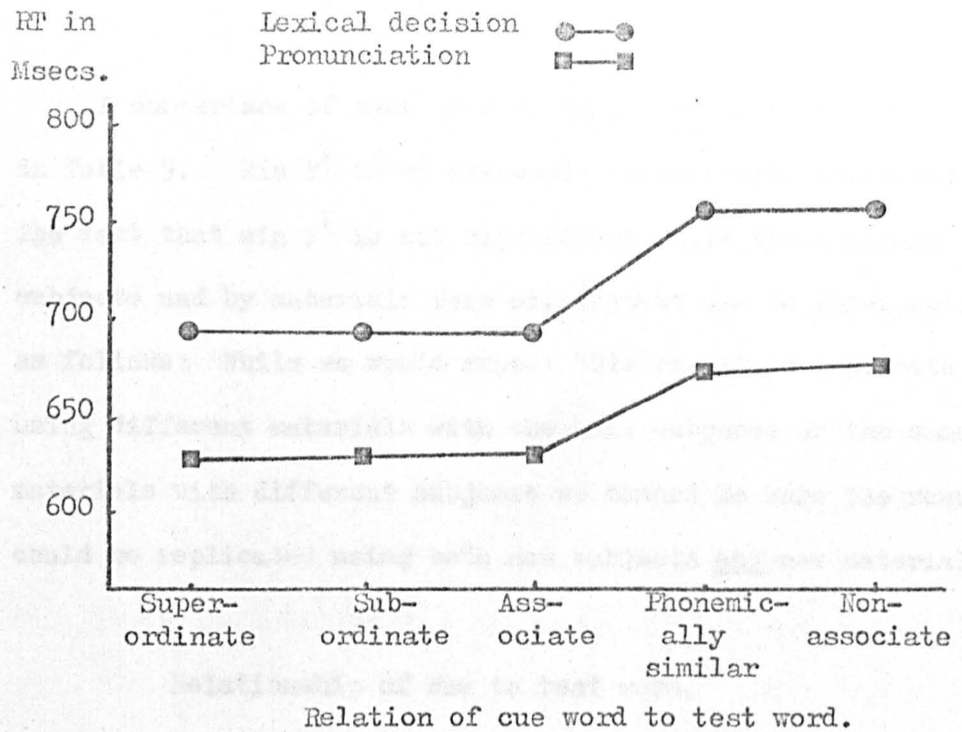
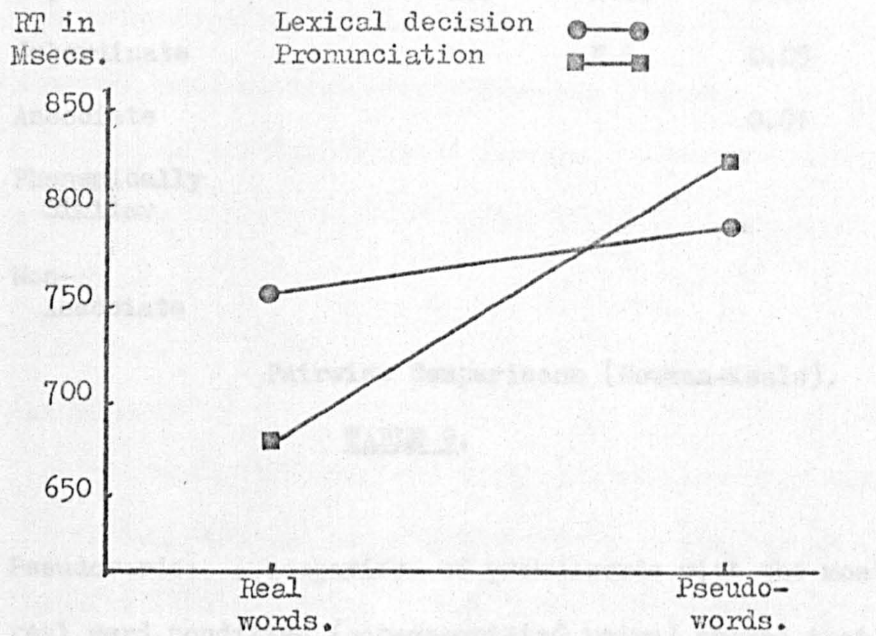


Figure 6.



Comparison of real and pseudo-words in the lexical decision and pronunciation tasks.

Figure 7.

A comparison of each pair of real word conditions are shown in Table 9. Min F' is an extremely conservative statistic. The fact that min F' is not significant while the analyses by subjects and by materials were significant can be interpreted as follows: While we would expect this result to replicate using different materials with the same subjects or the same materials with different subjects we cannot be sure the results could be replicated using both new subjects <u>and</u> new materials.

Relationship of cue to test word.

	Super-ordinate	sub-ordinate	Associate	Phonemically Similar	Non-Associate
Superordinate		N.S.	N.S.	0.05	0.05
Subordinate			N.S.	0.05	0.05
Associate				0.01	0.05
Phonemically Similar					N.S.
Non-Associate					

Pairwise Comparisons (Newman-Keuls).

TABLE 9.

Pseudowords: A comparison of pseudowords with the most comparable real word condition (non-associated pairs) showed that pseudo-words mean RT was 144 msecs slower. This difference was in this direction for all 10 subjects (by sign test p = 0.002, two-tailed test).

Comparison of pronunciation task with lexical decision task:
It is possible to compare the results from the pronunciation task with the results from the lexical decision task.

Comparing mean RT for real words only, the pronunciation task was on the average 68 millisecs. faster. This difference occurred for 54 of the 60 real words (by sign test $Z = 6.1$ $p < 0.0001$).

Comparing mean RTs for the 12 pseudowords used in Exp. 3. with their RTs for the lexical decision task (Exp. 2.) the pseudowords were 17 millisecs. faster in the lexical decision task. Figure 7 shows the mean RT for the 12 pseudowords in the pronunciation and the lexical decision tasks compared with the real word control (non-associated pairs).

The interaction between pseudo/real words and type of task is significant ($F(1,44) = 7.01$ $p < 0.025$, analysis by materials). See figure 7.

Discussion:

The most striking fact about the results from the pronunciation task is that the pattern is so similar to the results from the lexical decision task (except for pseudowords which will be discussed in detail below). For real words the same pattern of significant differences was found but mean RT was faster in the pronunciation task than in the lexical decision task. Meyer et al. (1974) have argued that the fact that the pronunciation task is faster than the lexical decision task supports the assumption that the pronunciation task involves one less stage (i.e. lexical memory retrieval). An alternative argument is that producing a pronunciation of a string of letters is a far more practised task than deciding if the string of letters is a word or not. It may be the practice variable that leads to RT differences rather than any difference in the number of stages involved.

Semantic associates as cues:

Exactly the same results were found in this experiment for semantically associated cues as in Exp. 2. Superordinate, subordinate and simple associate cues all produced significantly faster pronunciation time compared to unassociated cues. Superordinate, subordinate and simple associate pairs did not differ significantly from each other. These results confirm the conclusions drawn in Exp. 2. concerning the plausibility of hierarchical structure as a principle of organization of semantic memory, as suggested by Collins and Quillian (1969).

It is necessary to note that there are a number of experiments which produce evidence that conflicts with this notion of a formal hierarchical structure (e.g. Schaeffer and Wallace, 1970; Conrad, 1972; Landauer and Meyer, 1972). It is worth noting an alternative way of describing these results proposed by Herriot (1974), (See also Craik and Lockhart, 1972). Herriot prefers to discuss the processes involved in memory rather than the structure and the content. In particular he is concerned with the coding of input in terms of attributes. Thus a word will be coded in terms of graphemic, phonemic and semantic attributes. A word will presumably share some of the attributes by which it is coded with its superordinates and its subordinates. It is these shared attributes which facilitate the coding process rather than some underlying structure of lexical memory. Such an approach makes it clearer why superordinate and subordinate terms are equally effective as cues. As discussed in Exp. 2. a superordinate has many subordinates to which it may be connected in some kind of associative network but a subordinate has far fewer possible superordinates. If the "activation" which produces facilitation (assuming for the moment a spreading activation model) spreads over all possible connections and assuming there is only a limited "amount" of excitation to be spread around (possibly in some kind of probabilistic way as inKiss' model, (1972), then a superordinate concept should spread its excitation over a greater number of subordinate terms and thus be less effective as a cue than a subordinate which will be connected to fewer concepts. However, according to the view that emphasises coding by attributes it is more reasonable to expect superordinates and subordinates to produce equal facilitation. If facilitation depends on the number of shared attributes which are used in coding then these attributes-in-common will be the same, regardless of whether the superordinate or the subordinate is the cue.

A similar explanation holds for the simple associates. Associated pairs tend to be words which share a number of common attributes. This, of course, only applies to paradigmatic associations and not to syntagmatic associations. However, paradigmatic associations form the vast majority of free associations (see Deese, 1965; Clark, 1970).

Phonemic Similarity as a cue:

As in Exp. 2. phonemically similar but graphemically dissimilar cues did not produce any facilitation. In pronouncing the string the subject was forced to code it phonemically where as in Exp. 2. the subject may have been able to decide the string was a word without forming a phonemic representation. Yet even in this present situation phonemic similarity alone does not facilitate pronunciation. Neither was there any evidence that graphemic dissimilarity caused any problems by biassing subjects to produce phonemically dissimilar representations. However, in English the correspondence between graphemic and phonemic representations is not one-to-one. English is notorious for its disparity between spelling and pronunciation. English speaking subjects may thus be flexible in their approach to pronouncing letter strings.

It may be that subjects who speak a language with a closer correspondence between spelling and pronunciation (e.g. Russian or Turkish) would be inhibited by phonemic similarity and graphemic dissimilarity as Meyer et al. (1974) suggest.

The results obtained by Meyer et al. (1974) and this experiment can be described in a similar way to the coding-by-attributes approach used in the discussion of semantically associated cues. In the case of graphemic similarity subjects are biassed to use the same attributes to code the two words. For phonemic coding there is only a bias to use the same attributes if the subject has previously detected graphemic similarity.

Graphemic, Phonemic and Semantic Coding:

What is the respective status of graphemic, phonemic and semantic coding? Graphemic coding is operating most closely with the stimulus as presented (the nominal stimulus in Herriot's terminology). The attributes used for coding here are directly obtained from the stimulus itself. Facilitation due to graphemic similarity may be due to activity in very "low order" stimulus analysing mechanisms (see Introduction).

Phonemic coding is dependent upon the operation of phonological conversion rules upon the graphemically coded representation. Presumably these conversion rules are stored in some memory store. If a phonological representation can be formed only through the use of these rules then facilitation through phonemic similarity may occur by biassing the selection of rules as Meyer et al. (1974) suggest. An alternative explanation is discussed in detail in the next section, which assumes that knowledge stored about the word may enable the production of a pronunciation independently of the rule system.

Similarly semantic coding depends upon gaining access to information in a long-term store (lexical memory). Herriot would object to the use of the term "store". He regards semantic memory as a process not a structure. However, it seems necessary to assume that past information is retained somehow. It is assumed here that the results of both the graphemic and phonemic coding operations are capable of being used in the semantic coding process. Which code will be used depends upon the task. If one assumes a logogen-like device which receives information from the graphemic and phonemic coding systems and uses this information to decide whether a given word has occurred, then after one word has been accepted as having occurred for a time afterwards the logogen system will accept that a related word has occurred on the basis of less evidence.

This view is consistent with Meyer et al's. (1974) finding that association effects were larger for visually degraded words. Given a complete listing of the physical features of the stimulus the decision making system can easily decide which word has occurred. The less physical description the system has available the greater the relative importance of context. This last point is also relevant to the question of why it is necessary to hypothesize an evidence weighing mechanism at all. It could be argued that a graphemic or phonemic representation of a word should be sufficient to say that it has occurred and to retrieve its meaning. However, in most normal word recognition situations, such as reading or listening to a discourse, it is probable that the sensory information available for constructing a coded representation is far from complete. Given that the sensory evidence is only fragmentary, deciding which word has occurred will be a probability problem. Context, mediated by past experience, helps the system to make the "best bet". Mechanisms like the logogen describe this interaction of evidence in word recognition (see also Norman, 1968). An advantage of such a mechanism is that the process is speeded up since analysis of the potentially available information from the senses is reduced to a minimum. A disadvantage is that since the mechanism works probabilistically it will sometimes make mistakes and decide that the wrong word has occurred.

An account of such a model is given in more detail in the next section.

Pseudowords:

As in the lexical decision task responses the pseudowords were slower than to real words. Since it is assumed that the pronunciation task involves no lexical memory search it must be assumed that pseudowords are slower in some other stage, either in an encoding or a response stage or both. It is interesting to note the interaction between word/pseudoword and lexical decision/pronunciation task. (See Fig. 7). One hypothesis to explain this interaction is to assume that there are two sources of difficulty for pseudowords in the pronunciation task but only one source in the lexical decision task. The coding difficulty is common to both tasks but the pronunciation task has an extra difficulty in the response stage. Even when coded internally the actual motor plan for pronouncing the pseudoword is completely unpractised and therefore slower. This is discussed more fully in the next section.

General Discussion:

As mentioned in Exp. 2. there a number of methodological and technological problems in both Exp. 2. and Exp. 3. The major problem is that of comparing the results of the groups of words in the experimental conditions with those of the words in the control condition (the unassociated pairs) and comparing experimental group among themselves. In this design comparing difficult conditions involves a Between-Words comparison. There is the danger that differences between conditions may have been confounded with differences between different groups of words. For instance in the pronunciation task different words may have taken different times to activate the voice key. However, these effects are likely to be small and will probably have averaged out over words. It is unlikely that any such effects would be important to the overall facilitation effect which appears to be quite robust. They may, however, affect results where the differences may be smaller, possibly in the comparison of the effectiveness of different kinds of cues. E.g. although frequency was controlled for as far as possible there was a non-significant trend for superordinate words to be more frequent than subordinate words.

On the technical side it was felt that having the subjects read the cues from the booklet was not ideal, in that the time of exposure to each cue was not subject to close control. This is unlikely to have affected the results seriously since all conditions were subject to the same variations.

It is necessary to bear in mind the problems discussed above when drawing any conclusions from Exps. 2. and 3. However, while encouraging caution it is unlikely that these problems invalidate the basic findings.

Conclusions:

It is concluded from this experiment that associative priming effects influence a stage common to both the lexical decision and the pronunciation tasks. It is argued that this stage involves coding of the stimulus and that information from visual, acoustic and semantic sources interact in this stage. Phonemic similarity without graphemic similarity did not produce facilitation of pronunciation. This confirms the finding of Exp. 2. It is argued that pseudowords are difficult to process in both an encoding and a response stage.

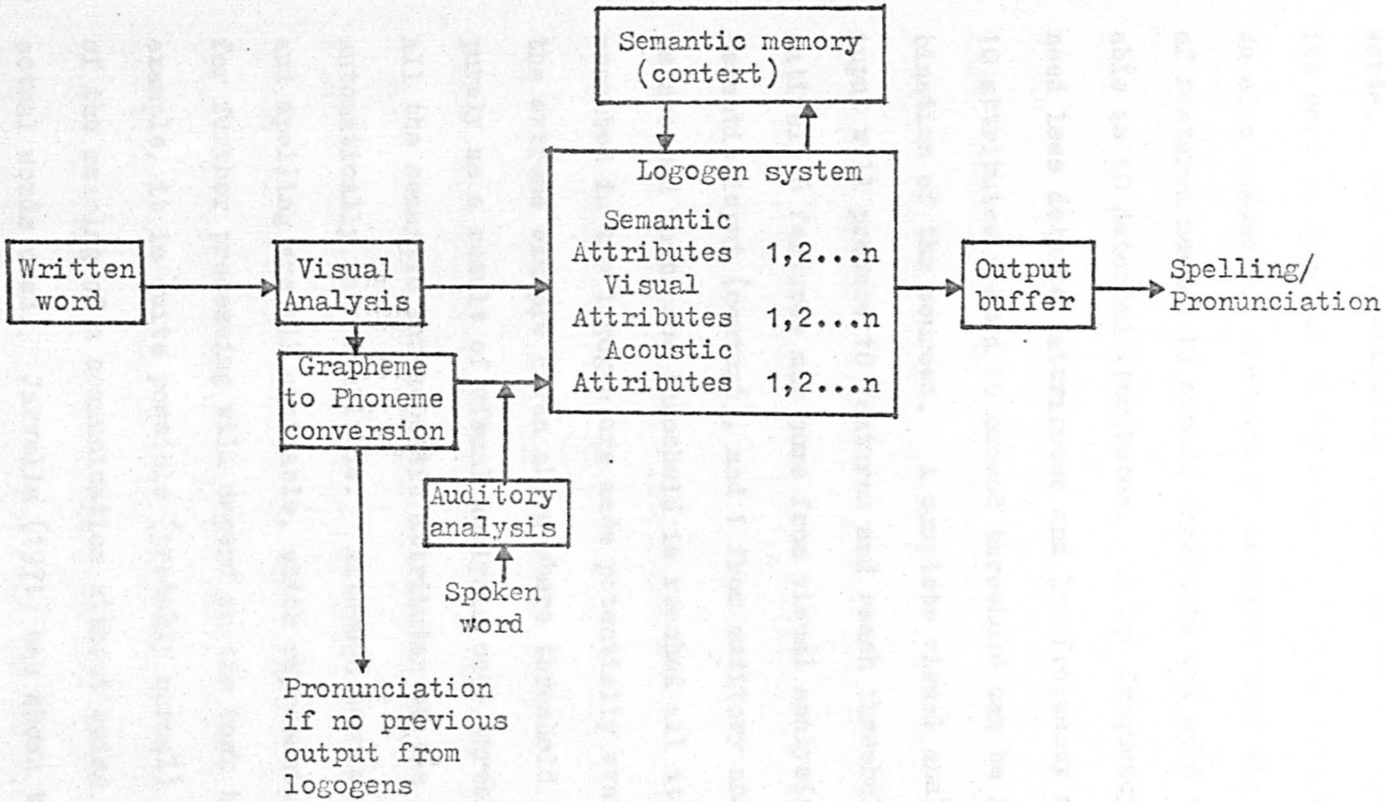
<u>A proposed model of word recognition:</u>

In this section a model is outlined which can describe most of the existing data on word recognition and pronunciation. It is essentially Morton's logogen system but with an additional feature to explain the handling of non-words and some other results which present problems for the logogen model as it stands at the moment. The emphasis here is on visual word recognition although auditory word recognition is also considered. The model is portrayed in Figure 8.

The logogen system is concerned exclusively with recognizing words. At the same time it can make use of all the available information and knowledge (essentially abstract and non-verbal) stored in semantic memory. Each word is represented in the logogen system as three arrays of attributes. These arrays consist of semantic, graphemic and acoustic features. Each logogen thus represents its word by a unique combination of attributes. Each logogen monitors the input from the sensory system and counts the number of its attributes it detects in the stimulus. If the count exceeds some threshold the form of either an articulatory plan which can become vocalized or covertly rehearsed, or the output can be input into the semantics system. Both options can of course occur together.

Let us consider in more detail the working of the logogen, in particular the interaction of attributes from the three different sources. First we shall consider visually presented

Figure 8.



A proposed model of word recognition.

words. Let us arbitrarily assume that each logogen describes its word in terms of 30 attributes, 10 for each kind of attribute. We also assume for the sake of argument that the average number of features needed to exceed threshold and make the word available is 10 detected attributes. H igh frequency words will need less detected attributes and low frequency more. These 10 attributes needed to exceed threshold can be from any combination of the sources. A complete visual analysis of the input will produce 10 features and reach threshold. Alternatively 6 features may come from visual analysis, 3 from semantic input (context), and 1 from auditory analysis. It is assumed that once threshold is reached all attributes represented in the logogen are made potentially available. In the extreme example given above where threshold was reached purely as a result of visual analysis once threshold was reached all the semantic and acoustic attributes of the word were automatically made available. Although meaning, pronunciation and spelling are all available, which representation is selected for further processing will depend on the task involved. For example, it is quite possible (probably normal) to be aware of the meaning of a communication without being aware of the actual words used. Jarvella (1971) has shown that subjects "code discourse in terms of its meaning, retaining the exact form of words only in the sentence currently being coded." (Herriot, 1974, p 73). Similarly it is quite possible to read a passage out loud without being aware of its meaning.

Sources of input to the logogen.

Semantic Attributes: Each logogen has an array of semantic attributes defining the meaning of its word. In Fig. 8 semantic attributes are pictured as being contained in the logogen system, separate from semantic memory. This separation is without doubt artificial, since attributes defining words must be part of semantic memory. However, it is convenient, for the moment, to distinguish words as words from words as labels for abstract concepts.

Input to the semantic attributes array of the logogen comes from the semantic system as a result of previous output from the logogen system into the semantic system. (The semantic system may well make spontaneous output to the logogen system). A given attribute will be shared by a number of different logogens. If one of these logogens makes available the semantic system its semantic attributes, the semantic system will "activate" the shared attribute in the other logogens. (Morton (1970) has explicitly stated that logogens are not directly connected with each other). E.g. if the attribute FOOD was made available as a result of one logogen reaching threshold all other logogens containing the attribute FOOD as one of their defining set would detect its occurrence via the semantic system and increase their count of members of their set which have occurred.

The semantic system also feeds into the logogen system information from non-verbal context. Presumably the attribute detectors only remain "activated" for a limited period of time. It is unclear at the moment whether they are simply "on" or "off" or whether they decay over time.

Graphemic Attributes: Seymour (1973) says that "conversion of stimulus to a visual representation (Vi) is an obligatory operation, which is perhaps analogous to the formation of an icon. It is less clear whether accumulation of members of Vi by units in the logogen system is also obligatory or whether it is an optional operation which corresponds to the encoding of information in the icon, and permits the type of spatial selectivity which has been demonstrated by Sperling and others". It is assumed here that in normal reading the visual representation of the stimulus actually used by the system (the functional stimulus, in Herriot's terminology) is a far from complete representation. Visual attributes represented in the logogen's array of visual analysers may be fairly crude, e.g. first letter, length, overall shape etc. Recognition of a word in context means that some of the semantic attributes of the logogen will already have contributed to the count of features, so that detection of some, rather than all, visual attributes will be sufficient for the logogen to reach threshold. Only in situations where there is no context is a complete letter-by-letter analysis likely to be necessary.

Acoustic Attributes: The model presented here assumes that as soon as a functional visual representation has been formed grapheme-to-phoneme conversion rules can operate. (See e.g. Gibson, Pick, Osser and Hammond, 1962; Bradshaw, 1975). The other reason is the considerable body of evidence (Rubenstein, Lewis and Rubenstein, 1971; Snodgrass and Jarvella, 1972; Stanners, Forback and Headley, 1971) showing that phonological representations influence word recognition. Thus some information concerning the acoustic attributes of a letter string is available as input to the logogen system.

Sometimes the phonemic conversion using the rules system will provide sufficient information to produce a pronunciation of a word before that word's logogen has reached threshold. This is most likely to happen in the case of rare words with a high threshold, and will also of course occur for non-words which have no representation in the logogen system. A fundamental distinction is drawn here between a subject being able to pronounce a word because he "knows" the word and being able to compute its pronunciation by applying the rules of pronunciation he has learned.

Auditorily presented words:

The operation of the logogen system is basically the same for auditorily presented stimuli as for visually presented stimuli. Each logogen inspects the input from the auditory analysis to see if any of its defining acoustic attributes are present. The effect of the context is through activating semantic attributes in the logogen in exactly the same way as for visual stimuli. Once a logogen reaches threshold as a result of semantic and acoustic input all of its attributes are made potentially available, including the visual attributes.

Presumably people possess rules which enable them to produce some kind of graphemic code from a phonemic coding of new words and pseudowords. However, such a possibility is not included explicitly in the model since it seems unlikely that such a conversion occurs in normal word recognition independently of the logogen system.

Word production: The view of the logogen system taken here is that word production operates in fundamentally the same way as word recognition. The semantic system inputs to the logogen system a number of semantic attributes and the logogen with the corresponding attributes in its defining set reaches threshold, making either the visual or acoustic attributes or both available for output depending on the task.

Sentence production and comprehension: The model at present does not deal in detail with sentences. It is assumed that semantic memory contains the rules for combining words according to the grammar of the language (see Tulving's definition of semantic memory in the Introduction). These rules will interact with the input to and output from the logogen system. No attempt is made here to examine this interaction.

Summary of the model: The logogen system functions as a three-way interchange for different kinds of codings of words. Given a semantic, acoustic or visual representation of a word as input, a semantic, acoustic or visual representation may be output. Each logogen represents its word by a unique combination of semantic, acoustic and visual feature detectors. Each kind of detector examines input from its own information source. The detectors interact so that if one set of detectors recognizes a number of its attributes from its own source, less information is needed from the other sources for the logogen to reach threshold.

The present model differs from Morton (1969, 1970) by assuming that grapheme-to-phoneme rules can operate on the visual representation of a stimulus independently of the logogen system. The results of this conversion are available as input to the logogen system.

EXPERIMENT 4

Experiment on Associative Priming effects: their time course, differences over kinds of associative and the effect of intervening items.

Introduction:

In the model of word recognition proposed in the last section it was suggested that associative priming effects occur by activating semantic attribute detectors in some hypothetical evidence-weighing mechanism named a logogen (after Morton, 1969). If a logogen reaches threshold (i.e. detects a sufficient number of its defining features) it makes available a full visual, acoustic or semantic description of its word. If the list of semantic attributes are input to the semantic system there is a feedback from the semantic system to the logogen system in the form of information about which semantic attributes have recently been used. These attributes are now "activated" in all logogens which contain them as part of their defining set. The result of this activation is that these logogens will reach threshold on the basis of less information from other sources. This lowered threshold appears as reduced recognition times.

Activation of the feature detectors can only last for a limited period of time. Otherwise as input to the semantic system increases so will the activation of detectors in the logogen system and logogens will be reaching threshold although their words have not been presented. Morton (1969) suggests that activation will rapidly decay and will have disappeared completely after one second. Meyer, Schvaneveldt and Ruddy (1972) have reported that associative priming effects do decay over time, being greatest with a zero time interval between words. However, they report that there was still considerable facilitation after four seconds. Morton's figure of one second may still be correct for normal reading and listening situations where there is a continuous input. In the Meyer et al. experiment subjects may have been able to maintain excitation by some kind of rehearsal loop. The time intervals used by Meyer et al. were 0, 1500, and 4000 msecs. It may be possible to clarify the situation by using other time intervals. In particular, a time interval between 0 and 500 msecs. may be long enough to demonstrate any decay of activation over time but be short enough to discourage any active rehearsal strategy.

In Exp. 1. two models were proposed to explain the associative priming effects. These were the spreading activation model and the shifting location model. The model of word recognition put forward in the previous section is more consistent with the former model since the concept of activation of feature detectors is central to its working. As yet, though, no evidence has been presented which allows a choice to be made between the two models empirically. Meyer et al. (1972) have tested the two models by separating two associated words by an unassociated word (e.g. BREAD-DOCTOR-BUTTER). According to the location shifting model the presence of the unassociated word alone should prevent any facilitation of the third word by the first. On the other hand the spreading excitation model predicts that the unassociated word will not prevent facilitation provided that the time interval between the two associated words is not too long. (See the Discussion of Exp. 1. for the rationale behind these predictions). Meyer et al. found that separating the associated words with an unassociated word did not eliminate the associative priming effect. Thus the shifting location model can be rejected.

One curious result reported by Meyer et al. (1972) was that inserting a non-word between the associated words (e.g. BREAD-SATH-BUTTER) completely eliminated any facilitation. They suggest that "such an effect indicates that processing a non-word may "reset" the memory system to a neutral state." It is hard to think of any reason why this should be. It was proposed in this experiment to subject this finding to a severe test by separating associated words by two non-words. If there was any evidence of priming effects in this situation then doubts must be cast on Meyer et al's. finding.

One clear prediction from the model proposed in the previous section that is tested in the present experiment, is that facilitation of recognition of a word will be dependent on the number of attributes it shares with the priming word. The argument here is restricted to shared semantic attributes. Synonyms should produce most facilitation since they have identical lists of attributes.[1] Next most effective should be antonyms since they differ from each other only on one feature. As shown in Exp. 2. and 3. subordinates are effective at producing facilitation. It is to be expected that subordinates would be less effective than synonyms and antonyms having relatively fewer attributes in common with their superordinates. The type of associates Underwood (1974) calls parallel associates (mainly coordinates, e.g. BREAD-BUTTER, ARMY-NAVY, but also such pairs as SPIDER-WEB) should also produce facilitation since they share a number of attributes. It is expected that they will be less effective than synonyms or antonyms. It is not easy to decide how parallel associates will compare with superordinates since there will be considerable item variation but on the average the number of shared attributes will be approximately the same for both kinds of associations.

The present approach has a number of similarities with the account of word associations given by Clark (1970), if it is

1.
This prediction is complicated by the fact that there are very few "true" synonyms, in the sense of two words being completely interchangeable. Although two words may have the same referent there are usually connotative or stylistic differences.

assumed that the processes underlying word association production overlaps to a large extent with the priming effects. Clark follows Katz and Fodor (1963) by assuming that a word can be represented by an ordered list of abstract features that completely characterize the "surface realization". E.g. MAN could be represented as (+Noun, +Det__, +Count, +Animate, +Human, +Adult, +Male). Clark states that free associations follow a "simplicity-of-production" rule which can be summarized as "perform the least change on the lowest feature with the restriction that the result must correspond to an English word". Clark describes rules to produce associations which are tried in the order of "simplest first". The first rule is named the "minimal contrast rule" which produces antonyms. The second and third rules are feature detection and addition rules which respectively produce superordinate and subordinate associations. Feature addition may also produce near-synonyms. Other rules include idiom-completion and selectional feature realization rules. It is not clear why exact synonyms are not given preference over the minimal contrast rule that produces antonyms since they require no feature changes at all. Clark seems to assume that the feature list must be changed. This is a difference between predictions from the word recognition model and Clark's theory. Clark's theory, by assuming a feature must be changed predicts antonyms will be more probable associates than synonyms and by inference will produce more facilitation.

The proposed word recognition model assumes that the only factor will be the number of shared attributes and hence synonyms are predicted to produce more facilitation than antonyms.

There are thus three aims for this experiment:

1 To examine the effects on associative priming effects of time intervals.

2 To examine the effects of nonwords intervening between two associated words.

3 To investigate associative priming effects over different kinds of association.

Method:

Equipment: The experiment was carried out using a GT40 display screen (Digital Equipment Corporation) under on line control of a PDP 11/45 computer. The computer controlled display time and response-stimulus interval (RSI), measured reaction time and recorded subject's responses.

Materials: Two groups of three lists were made up, the two groups consisting of paired "A" and "B" forms. Each list contained 10 practice items and 94 test items. Four different kinds of associations were used. Each list contained four pairs of antonyms (e.g. BLACK-WHITE), four pairs of conceptual associates (subordinate-superordinate e.g. CANARY-BIRD), four pairs of parallel associates (e.g. BREAD-BUTTER) and four pairs of synonyms (e.g. TINY-SMALL). Most associated pairs were drawn from the materials used by Underwood (1974) but some "Americanisms" were replaced with items from association norms. (Palermo and Jenkins, 1964).

There were two conditions of intervening items between associated pairs: no intervening items and two intervening items. Half the pairs of each kind of association followed each other in the list and half the pairs were separated by two non-words. Also in each list were 16 non-associated words that were from the same pool as the associated words. Each word in "A" list that was paired with an associate appeared in a "B" list without an associate; similarly each word in "B" list that was paired with an associate appeared in an "A" list without an associate. This design meant that each critical item (i.e. the second word in an associated pair) appeared in the lists an equal number of times primed and unprimed.

The remaining 46 items in each list were non-words produced by changing one letter in AA words from the Thorndike-Lorge Frequency count, so that they no longer formed real English words but still conformed to English spelling and pronunciation rules.

Procedure: The subject sat facing the GT40 screen with one finger from each hand on a micro-switch. He was instructed that he would be presented with strings of letters and he must decide as quickly as possible whether the string of letters was a real English word or not. If he decided it was a real word he pressed the right hand switch, if he decided it was not real he pressed the left hand switch. Subjects' response terminated the display of the letter string which was then followed by a set interval before the next letter string was presented. If the subject did not respond the display was terminated automatically after two seconds and the interval before the next item followed. Each subject received three lists, either all "A" or all "B" from lists. Each list constituted one block and for each block there was a different response-stimulus interval (i.e. the time between the subject pressing the button and the next item appearing). These RSIs were 300, 1000 and 2000 msecs. Half the subjects received the A lists and half the B lists. For the A and B lists subjects, lists RSIs, and order of presentation were combined in a Graeco-Latin Square as shown in Table 10.

Thus each critical word appeared an equal number of times primed and unprimed in each RSI condition and each order of presentation of blocks.

Subjects were instructed in what order they would receive the different RSIs. They performed all 3 blocks in one session with a two-minute interval between blocks. Each session lasted approximately 25 minutes.

	Order 1	Order 2	Order 3
Group 1	List 1/RSI 1	List 3/RSI 3	List 2/RSI 2
Group 2	List 2/RSI 3	List 1/RSI 2	List 3/RSI 1
Group 3	List 3/RSI 2	List 2/RSI 1	List 1/RSI 3

Orders of Presentation of Materials

TABLE 10.

Subjects: 18 subjects were used. They were all psychology undergraduates of Stirling University participating as a course requirement. N one of them had participated in Exps. 1, 2. or 3.

Results and Discussion:

The results presented here include only the results of critical items (second members of each associated pair). The overall error rate for these items was 2.9%. There were no significant differences in error rate between conditions. Only data from correct responses was included in the analyses.

Analyses of variance were carried out treating both subjects and materials as random effects (see Clark, 1973). There was a significant main effect comparing 0 intervening items versus 2 intervening items (by subjects: $F(1,17) = 10.06$ $p < 0.01$; by materials: $F(1,94) = 4.24$, $p < 0.05$; min. $F^1(1,96) = 2.98$ $p < 0.1$). See Figure 9.

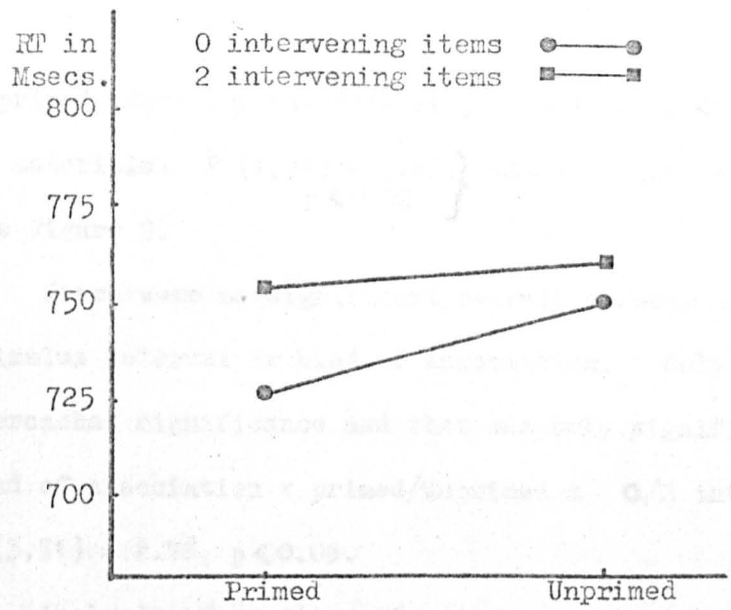
There was also a significant main effect for primed versus

		Mean	S.D.
Primed:	0 intervening items	726	143.5
	2 '' ''	754	113.8
Unprimed:	0 '' ''	751	192.9
	2 '' ''	761	139.3

Means and standard deviations for Fig.9.

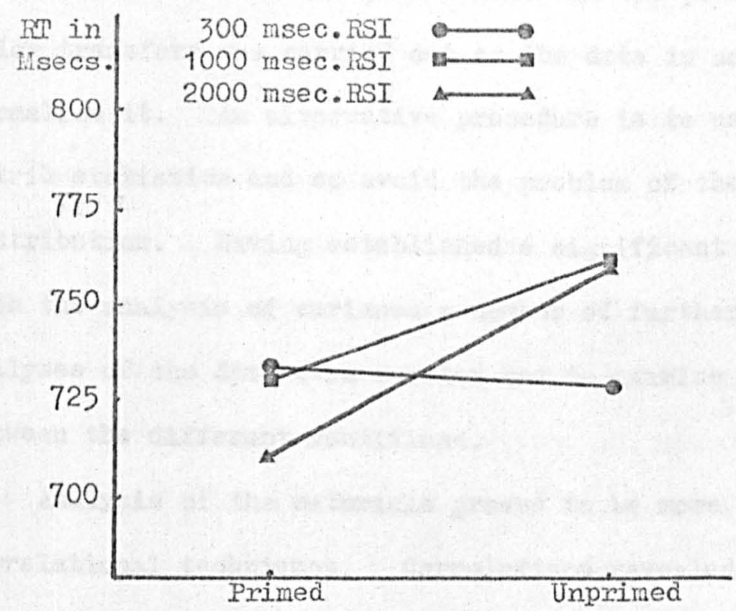
		Mean	S.D.
Primed:	300ms.RSI	730	175.2
	1000ms.RSI	733	141.8
	2000ms.RSI	710	114.3
Unprimed:	300ms.RSI	732	110.7
	1000ms.RSI	762	148.3
	2000ms.RSI	761	132.4

Means and standard deviations for Fig.10.



Mean RT for 0/2 intervening items conditions, primed and unprimed.

Figure 9.



Mean RT for primed and unprimed conditions at each RSI with 0 intervening items.

Figure 10.

unprimed items (by subjects $F(1,17) = 11.06$ $p < 0.01$; by materials: $F(1,94) = 6.63$; $p < 0.02$ } min $F^1 (1,80) = 4.14$ $p < 0.05$). See Figure 9.

There were no significant overall effects of response-stimulus interval or kind of association. Only one interaction approached significance and that was only significant by subjects: kind of association x primed/unprimed x 0/2 intervening items $F(3,51) = 2.76$, $p < 0.05$.

Analysis of reaction time data presents a problem because of the non-normal distribution of the data. Typically reaction time data is skewed to the left. (see appendix). A number of suggestions have been made by different experimenters for the analysis of reaction time data. In the present analysis a log transform was carried out on the data in an attempt to normalize it. An alternative procedure is to use non-parametric statistics and so avoid the problem of the non-normal distribution. Having established a significant priming effect with the analysis of variance a number of further non-parametric analyses of the data were carried out to examine differences between the different conditions.

Analysis of the materials proved to be more amenable to correlational techniques. Correlations revealed a number of subtle relationships which were not readily apparent in the results of the analysis of variance.

The effects of length of RSI - 0 intervening items.

The mean RTs for the primed and unprimed conditions at each RSI with intervening items are shown in Figure 10.

The differences between unprimed and primed RT at each RSI are: 300 - 4 msecs, not significant; 1000 - 29 msecs, not significant, 2000 - 51 msecs, sign test over subjects, x=3 p = 0.004.

None of these differences were significant over materials. As will be seen later different items tended to produce different results at different RSIs.

A comparison can be made between these results and those of Meyer, Schvaneveldt and Ruddy (1972). A facilitation score is used, which is obtained by subtracting the primed RT from the unprimed RT. This score can of course be negative as well as positive. The comparison is shown in Figure 11.

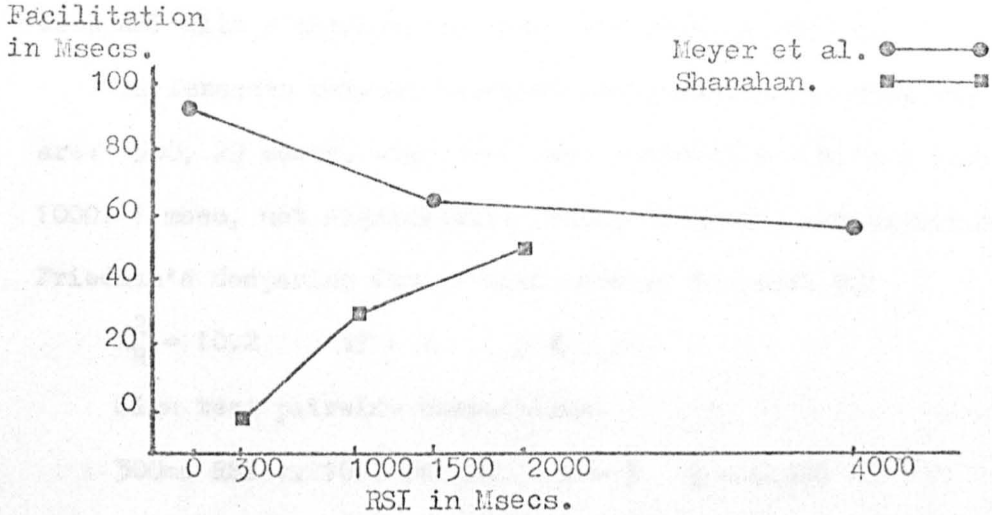
The results obtained in the present experiment do not show a significant difference between the three RSIs but the trend is clearly in the opposite direction to that of Meyer et al's. The reason for these different findings may lie in the nature of the materials used. Later analyses and discussion will show that it is insufficient to describe the time course of excitation independently of the nature of the materials used.

The effects of length of RSI - 2 intervening items:

The mean RTs for the primed and unprimed conditions at

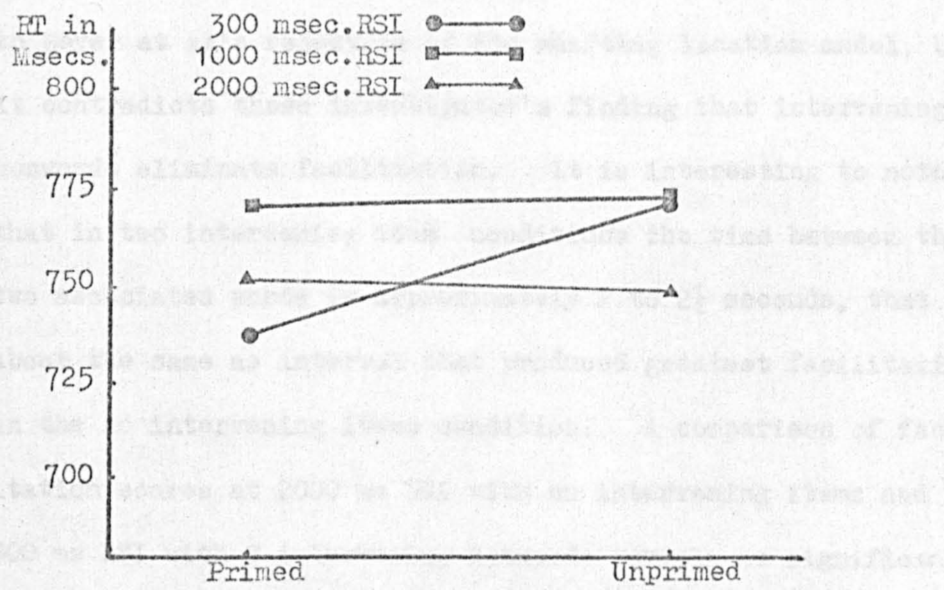
		Mean	S.D.
Primed:	300ms.RSI	739	116.1
	1000ms.RSI	768	125.7
	2000ms.RSI	753	103.5
Unprimed:	300ms.RSI	768	141.0
	1000ms.RSI	769	159.8
	2000ms.RSI	746	120.1

Means and standard deviations for Fig.12.



Comparison of priming effects obtained by
Meyer et al. (1972) and Shanahan (1975).

Figure 11.



Mean RT for primed and unprimed conditions
at each RSI with 2 intervening items.

Figure 12.

each RSI with 2 intervening items are shown in Figure 12.

Differences between unprimed and primed RT at each RSI are: 300, 29 msecs, sign test over subjects x = 4, p = 0.015; 1000, 1 msec, not significant; 2000, -7 msecs, not significant. Friedman's Comparing facilitation scores for each RSI $X_R^2 = 10.2$ df = 2 $p < 0.01$.

Sign test pairwise comparisons.

300ms RSI v. 1000 ms RSI. x = 3 p = 0.004

300 ms RSI v. 2000 ms RSI. x = 4 p = 0.015

1000 ms RSI v. 2000 ms RSI not significant.

At a short time interval between items (300 ms) even inserting two nonwords between two associated words does not eliminate facilitation. This result gives further support to Meyer at al's rejection of the shifting location model, but it contradicts these investigator's finding that intervening nonwords eliminate facilitation. It is interesting to note that in two intervening item conditions the time between the two associated words is approximately 2 to $2\frac{1}{2}$ seconds, that is about the same as interval that produced greatest facilitation in the no intervening items condition. A comparison of facilitation scores at 2000 ms RSI with no intervening items and 300 ms RSI with 2 intervening nonwords reveals no significant difference. (by subjects - sign test x = 7, p = 0.4). This suggests that time is the more important variable than intervening items.

Differences in facilitation between kinds of association:

As stated above there was an indication of an interaction between kind of association x primed/unprimed x 0/2 intervening items. This interaction is depicted in Figure 13. For the sake of simplicity facilitation scores (unprimed RT - primed RT) have been plotted.

Analysis of facilitation scores of each kind of association with two intervening items revealed no significant differences. The rest of this discussion will concentrate on the results from the no intervening items condition.

Analysis of the different kinds of associations facilitation scores by Kruskall-Wallis showed an overall significant difference ($H = 7.4$ d.f. = 3 $p < 0.05$. N.B. Only significant over materials, not over subjects). The results of pairwise comparisons by Mann-Whitney are shown in Table 11.

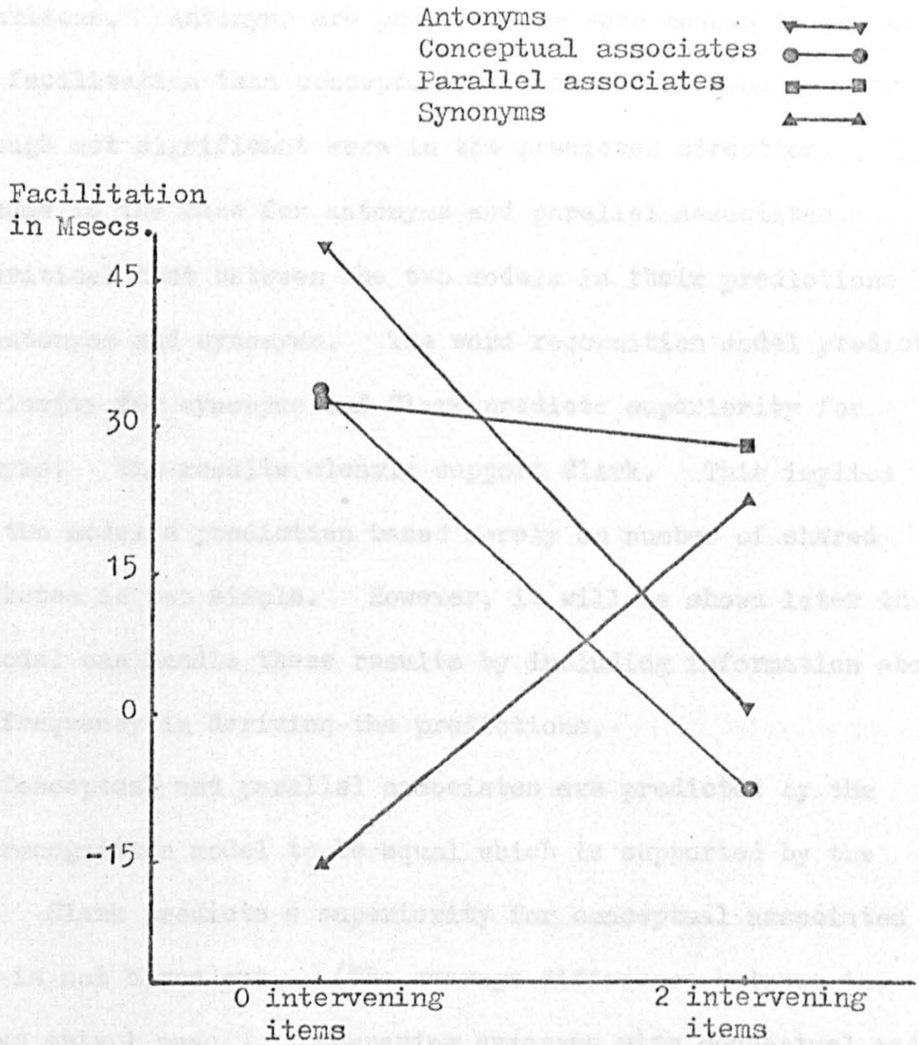
	Antonyms	Conceptual Associates	Parallel Associates	Synonyms
Antonyms		N.S.	N.S.	0.01
Conceptual Associates			N.S.	0.025
Parallel Associates				N.S.
Synonyms				

Pairwise Comparisons (Mann - Whitney)

TABLE 11.

	Primed.	Mean	S.D.	Unprimed.	Mean	S.D.
0 intervening items:	Antonyms	720	43.3		769	74.3
	C.A.s	708	40.5		740	64.0
	P.A.s	728	67.5		762	44.7
	Synonyms	749	55.6		743	37.7
2 intervening items:	Antonyms	737	65.6		739	42.7
	C.A.s	753	47.1		746	40.0
	P.A.s	762	83.3		772	68.5
	Synonyms	763	71.8		787	56.3

Means and standard deviations for Fig.13.



Priming effects for different kinds of associations with 0 or 2 intervening items.

Figure 13.

What are the implications of these results for the predictions made by the model of word recognition and Clark's theory of word associations? Consider each pair of possible comparisons. Antonyms are predicted by both models to produce more facilitation than conceptual associates. These results although not significant were in the predicted direction. The same is the case for antonyms and parallel associates. The critical test between the two models is their predictions for antonyms and synonyms. The word recognition model predicts superiority for synonyms and Clark predicts superiority for antonyms. The results clearly support Clark. This implies that the model's prediction based merely on number of shared attributes is too simple. However, it will be shown later that the model can handle these results by including information about word frequency in deriving the predictions.

Conceptual and parallel associates are predicted by the word recognition model to be equal which is supported by the data. Clark predicts a superiority for conceptual associates which is not borne out. (The average difference between the two was only 1 msec.). Comparing synonyms with conceptual and parallel associates again the data supports Clark. The word recognition model predicts greater facilitation for synonyms whereas Clark predicts less facilitation.

Overall then both theories do reasonably well at predicting the results of antonyms, conceptual associates and parallel associates. However Clark's theory is better able to handle synonyms than the simple predictions from the word recognition model based only on number of shared attributes.

Item differences:

Clark (1973) has stressed the importance of including language materials as a random and not a fixed effect in the calculation of statistics. It was readily apparent in analysing these results that there were large differences between items in the facilitation scores they had for each RSI and at the same time there were differences within items over RSIs.

In order to try to discover what were the factors determining these item differences a number of post-hoc analyses were carried out. Since the priming effects were much less reliable in the two intervening item condition only words used in the no intervening item condition were included in these analyses. Four facilitation scores were used in these analysis. These were each words facilitation score for each RSI and a mean facilitation score for all RSIs.

Table 12 shows the intercorrelations between unprimed RT, primed RT and facilitation.

These results follow the same pattern as the results of Experiment 1 on cued sentence verification. The most interesting finding is the positive correlation between unprimed RT and facilitation. That is, items with a slow unprimed RT tend to have high facilitation scores and items with fast unprimed RTs have small facilitation scores.

	Unprimed RT	Primed RT	Facilitation
Unprimed RT		$+0.42^{**}$	$+0.57^{***}$
Primed RT			-0.48^{**}
Facilitation			

$^{**}p < 0.01$

$^{***}p < 0.001$

Spearman Rank Correlations.

TABLE 12.

This is presumably the result of a ceiling effect. The easier an item is to recognize unprimed the less "room" there is for facilitation.

As discussed in Exp. 1. one factor which is known to affect both associations and word recognition is the frequency of occurrence of a word in the language.

The model of word recognition proposed here specifically includes frequency as a factor in word recognition. The relationship between unprimed RT and facilitation may arise because they are both related to word frequency. Low frequency words have high thresholds in that they need to detect a large number of their attributes in order to accept that the word has occurred. This leads to slow recognition time. On the other hand high frequency words are accepted as having occurred on the basis of detecting relatively few attributes. Thus activation of semantic attributes by presenting a related word has more "room" to facilitate a low frequency word which is far from threshold before priming than a high frequency word which is near threshold before priming.

This hypothesis was tested by correlating frequency with unprimed RT and facilitation. Frequency was measured by Thorndike-Lorge, where two words tied ranking was determined by referring to the Thorndike count. Table 13 shows the correlations.

Correlation of frequency and unprimed RT	-0.47***
Correlation of frequency and facilitation	+0.05

*** $p < 0.001$

Spearman Rank Correlation

TABLE 13.

The negative correlation between frequency and unprimed RT is standard result. The higher the frequency the faster the RT and vice versa. However, there was no correlation at all between frequency and facilitation. Thus there is no support at present for the hypothesis that words with slow RTs unprimed have high facilitation scores because of a related word frequency variable.

As mentioned earlier it was apparent that individual words' facilitation scores varied considerably between the three RSIs. To see if any pattern could be identified the facilitation scores at each RSI for the 48 items were correlated. See Table 14.

	300 ms RSI	1000 ms RSI	2000 ms RSI
300 ms RSI		-0.53***	-0.51***
1000 ms RSI			0
2000 ms RSI			

*** $p < 0.001$

Spearman Rank Correlations.

TABLE 14.

These correlations suggest that there may be several different groups of words all receiving different amounts of facilitation at different intervals. The existence of different groups of words differing over RSI may explain why no correlation was found between frequency and facilitation. Information may have been lost by averaging over all RSIs. To test this frequency was correlated with facilitations scores at each RSI. No significant correlations were found.

Facilitation is a complex interaction between two words as has been shown in the differences between different kinds of associations. Since two words are involved it may be inadequate to consider only the frequency of the second word. Accordingly some analysis is required which takes into account both the frequency of the priming word and the frequency of the associated word.

In an effort to gain a clearer understanding of what role, if any, frequency plays in associative priming effects, the 48 pairs of words were divided into those pairs with a high frequency primer and high frequency associate, low primer and high associate, high primer and low associate, and low primer and low associate. These were then analyzed according to whether the frequency of the primer and associate were the same or different. Fig. 14 shows the average facilitation for each kind of pair averaged over all RSIs. Figures 15, 16 and 17 show the average facilitation for each kind of pair for each RSI.

Over all RSIs. (see Fig. 14) although there **is** a tendency for same-frequency pairs to have a higher facilitation score than different-frequency pairs the difference is not significant. The interaction is not significant. For the 300 ms. RSI (see Fig. 15) the difference between same and different-frequency pairs is significant. ($F(1,44) = 4.4$ $p < 0.05$). For the 1000 ms and 2000 ms. RSIs (Figs. 16 and 17) there are no significant effects although for 1000 ms RSI there is a tendency for different-frequency pairs to be superior to same-frequency pairs ($F(1,44) = 2.5$ $p < 0.1$).

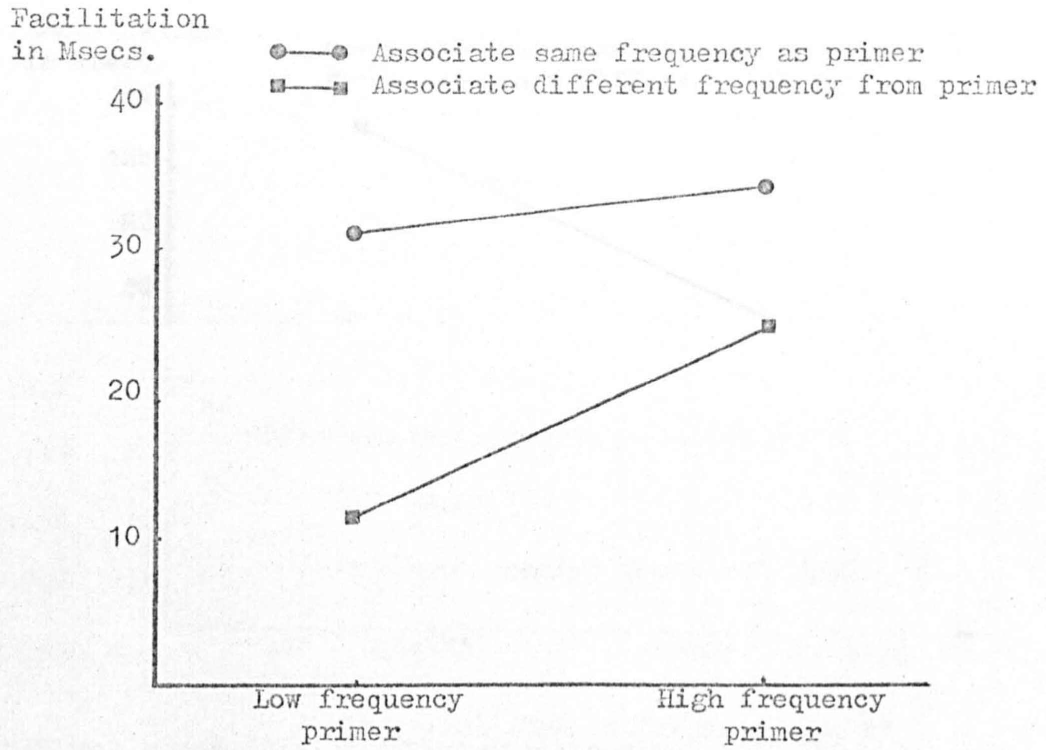
What can be concluded from these analyses? At the 2000 ms. RSI there is a large facilitation effect that is independent of the frequency of the words in the associated pairs. At the 300 ms and 1000 ms RSIs facilitation is related to word frequency. The nature of the relationship is reversed for the two RSIs. At 300 ms RSI same-frequency pairs are superior

		Mean	S.D.
Low frequency primer:	low associate	31.1	73.2
	high ''	11.0	65.3
High frequency primer:	low ''	25.1	50.2
	high ''	34.7	64.9

Means and standard deviations for Fig.14.

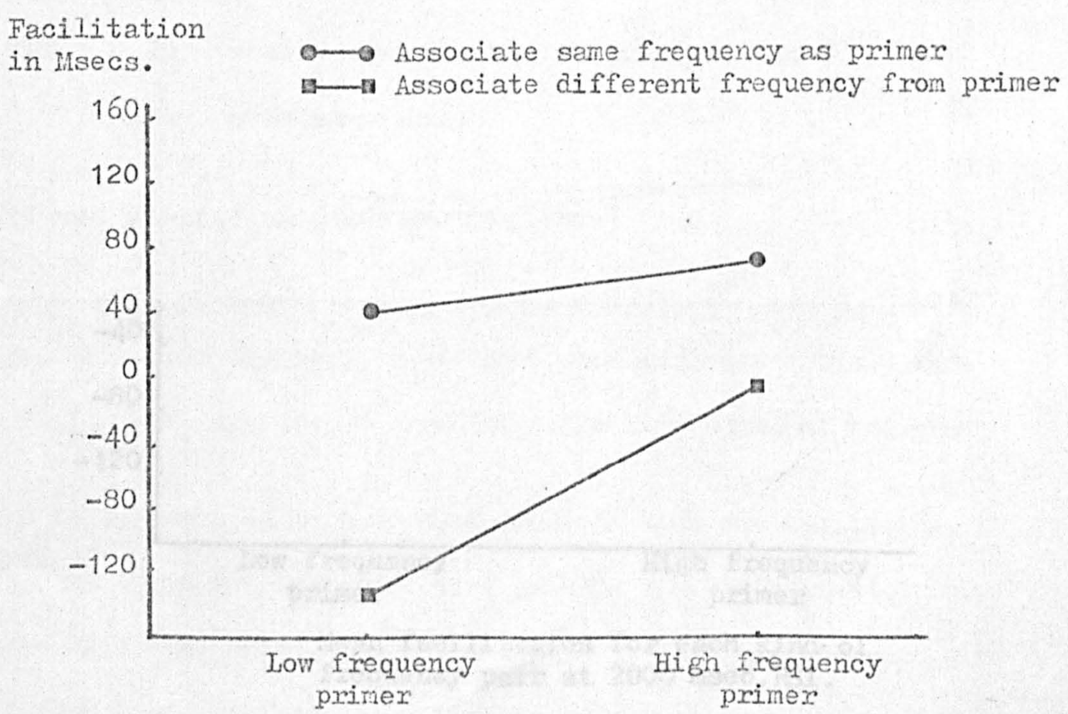
		Mean	S.D.
Low frequency primer:	low associate	42.1	253.3
	high ''	-134.5	218.0
High frequency primer:	low ''	-1.8	189.8
	high ''	76.3	167.1

Means and standard deviations for Fig.15.



Mean facilitation for each kind of frequency pair averaged over all RSIs.

Figure 14.



Mean facilitation for each kind of frequency pair at 300 msec.RSI.

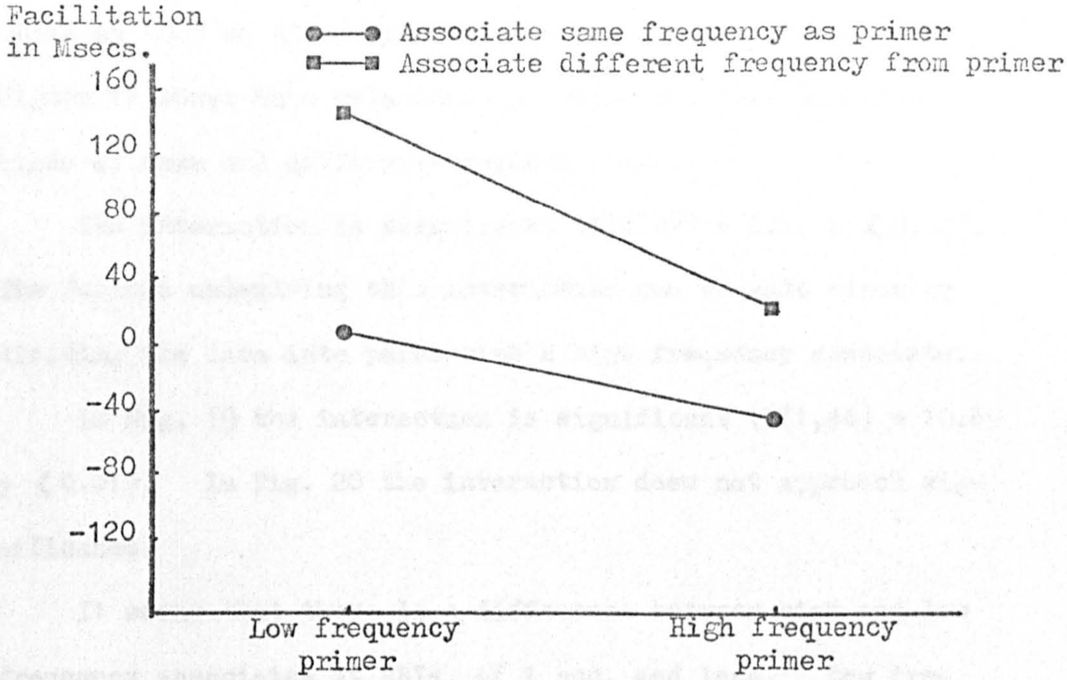
Figure 15.

	Mean	S.D.
Low frequency primer: low associate	7.1	211.3
high "	138.3	222.6
High frequency primer: low "	14.5	169.8
high "	-41.0	213.8

Means and standard deviations for Fig.16.

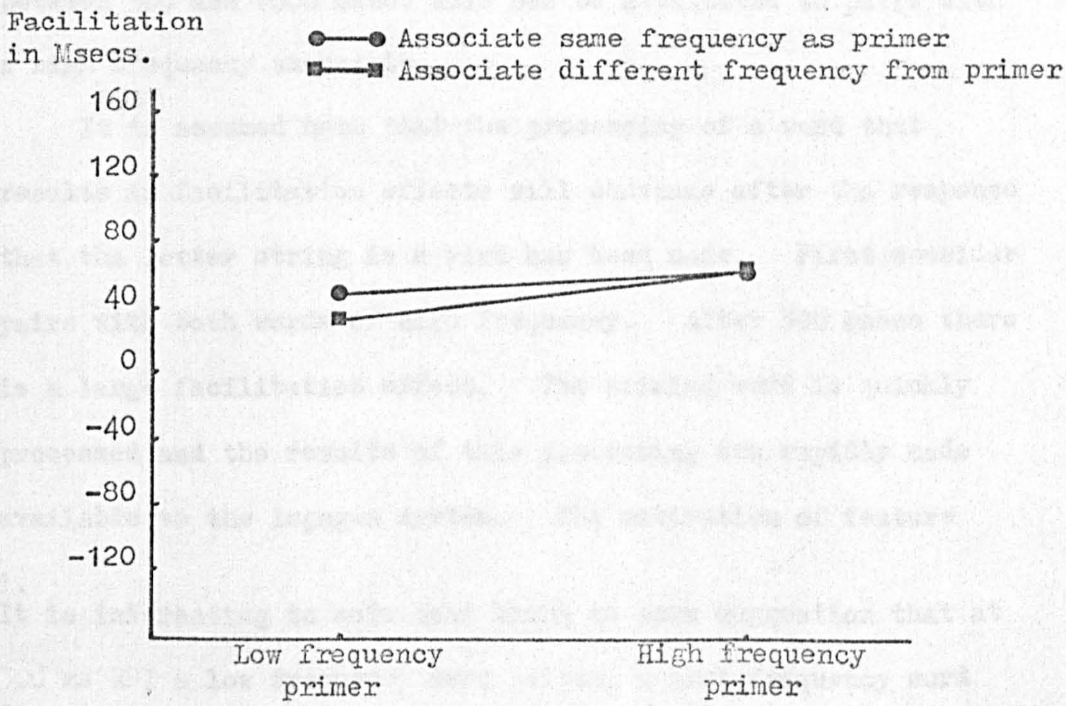
	Mean	S.D.
Low frequency primer: low associate	43.7	165.2
high associate	33.9	150.3
High frequency primer: low associate	61.9	212.6
high associate	61.5	210.8

Means and standard deviations for Fig.17.



Mean facilitation for each kind of frequency pair at 1000 msec. RSI.

Figure 16.



Mean facilitation for each kind of frequency pair at 2000 msec. RSI.

Figure 17.

while at 1000 ms RSI. different-frequency pairs are superior.[1] Figure 18 shows this relationship, collapsed over different kinds of same and different-frequency pairs.

The interaction is significant ($F(1,92) = 6.7$, $p < 0.02$). The factors underlying this interaction can be made clear by dividing the data into pairs with a high frequency associate.

In Fig. 19 the interaction is significant ($F(1,44) = 10.69$ $p < 0.01$). In Fig. 20 the interaction does not approach significance.

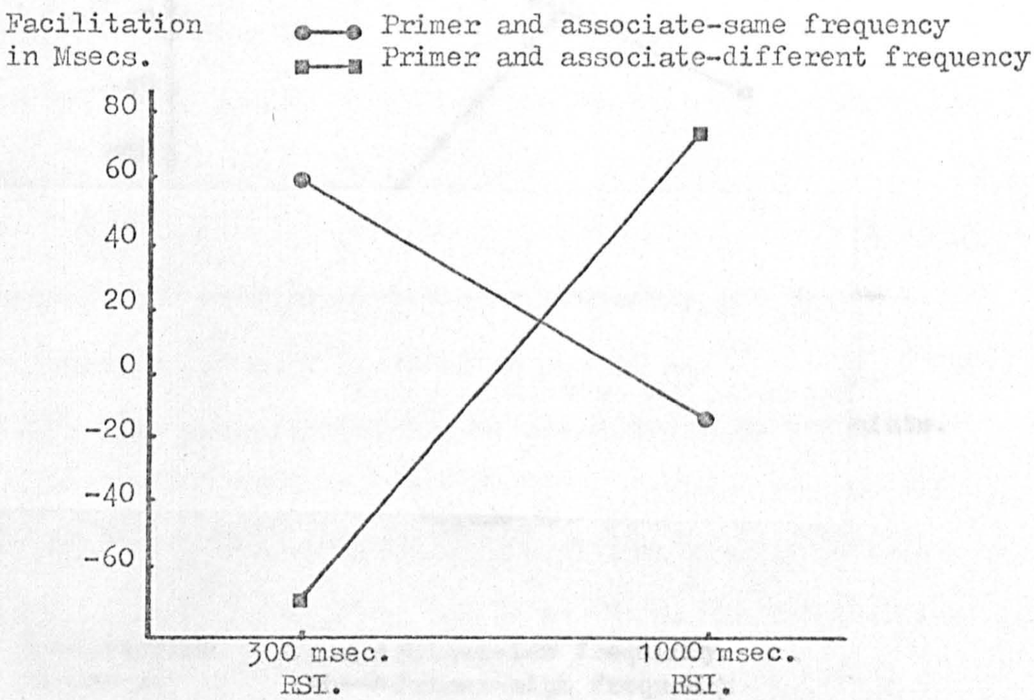
It seems that there is a difference between high and low frequency associates at RSIs. of 1 sec. and less. Low frequency associates appear to be relatively unaffected by either high or low frequency primers (except possibly low-low pairs - average facilitation = 42.1 msecs). Most of the difference between 300 and 1000 msec. RSIs can be attributed to pairs with a high frequency associate.

It is assumed here that the processing of a word that results in facilitation effects will continue after the response that the letter string is a word has been made. First consider pairs with both words of high frequency. After 300 msecs there is a large facilitation effect. The priming word is quickly processed and the results of this processing are rapidly made available to the logogen system. The activation of feature

1. It is interesting to note that there is some suggestion that at 300 ms RSI a low frequency word priming a high frequency word actually "inhibits" recognition. 10 of the 12 words were slower primed than unprimed. (Sign test $p = 0.038$).

		Mean	S.D.
300ms.RSI:	same frequency	59.3	210.5
	different frequency	-68.1	211.1
1000ms.RSI:	same frequency	-17.0	209.3
	different frequency	76.4	203.7

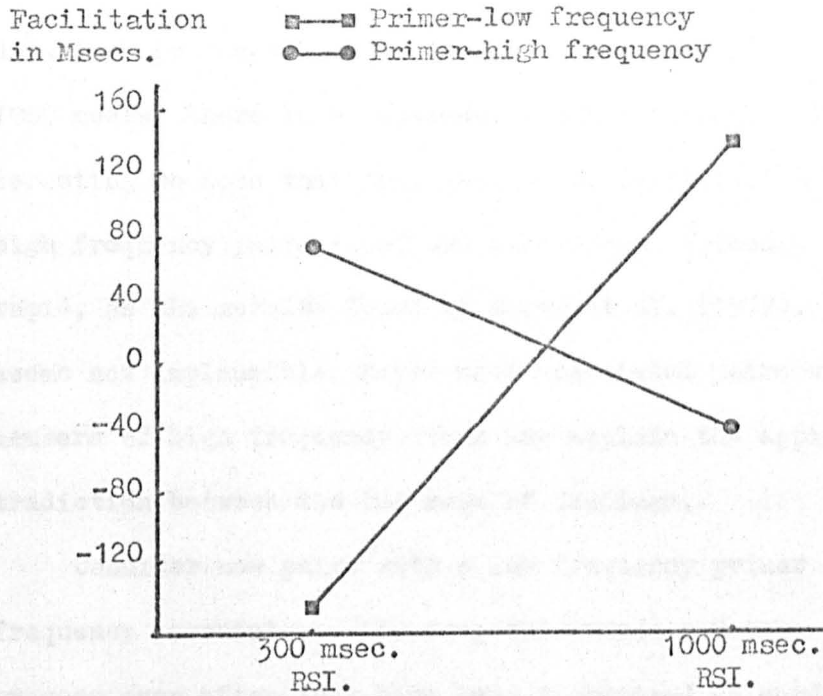
Means and standard deviations for Fig.18.



Priming effects for pairs of words of same or different frequency.

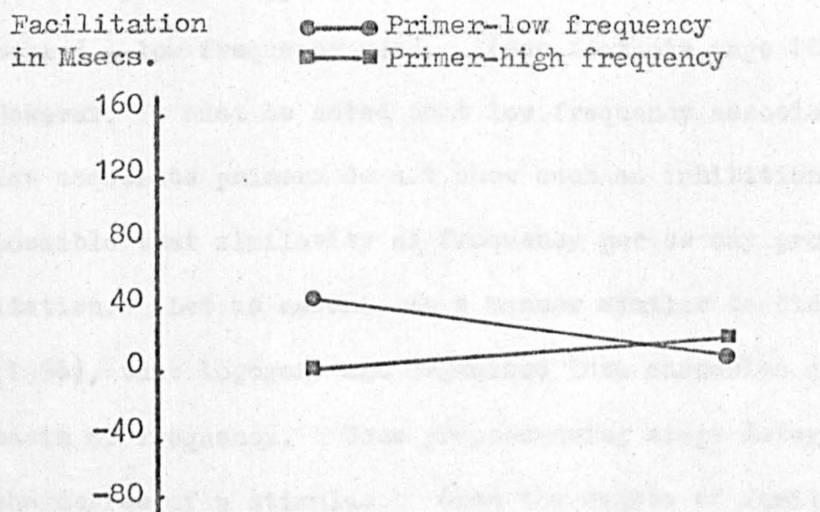
Figure 18.

The means and standard deviations for Figures 19 and 20 can be obtained from the tables facing Figures 15 and 16, following page 103.



Mean facilitation for pairs of words with a high frequency associate.

Figure 19.



Mean facilitation for pairs of words with a low frequency associate.

Figure 20.

detectors produced by this feedback decays over time and at 1000 msecs. there is an absense of facilitation. It is interesting to note that this decline of facilitation for high-high frequency pairs is of the same shape, although much more rapid, as the results found by Meyer et al. (1972). If, as seems not implausible, Meyer used associated pairs with both members of high frequency, this may explain the apparent contradiction between the two sets of findings.

Consider now pairs with a low frequency primer and a high frequency associate. Low frequency words may take longer to process even after they have been recognized as words. This processing may still be carrying on 300 msecs. after the response and this may interfere with the processing of the next word. Such an interference could account for the fact that high frequency words appear to be inhibited when following closely behind a low frequency word. (See footnote page 104). However, it must be noted that low frequency associates following low associate primers do not show such an inhibition. It is possible that similarity of frequency per se may produce facilitation. Let us assume, in a manner similar to Oldfield (1966), that logogens are organized into ensembles on the basis of frequency. Some preprocessing stage determines the degree of a stimulus. Once the degree of familiarity is established the information is input into the appropriate ensemble of logogens. As already described in the model, threshold of a logogen is related to its frequency.

For a time after accessing an ensemble (possibly up to $\frac{1}{2}$ second) it is easier to access the same logogen.

Why should facilitation at 2 seconds RSI be independent of frequency? It is assumed that where facilitation occurs at 300 and 1000 msec. RSIs the processes are largely automatic and are an integral part of normal language comprehension mechanisms. With a 2 second interval between words the subject can code and process the word completely, and may still have time to carry out some conscious operations leading to facilitation.

Another problem is why low frequency associates should be largely impervious to associative priming effects at the shorter RSIs. This finding suggests that the predictions from both the word recognition model and Clark's theory of word associations are inadequate if the only factor considered is shared features. Attributes-in-common may be a necessary condition for two words to be associated but it is not a sufficient condition. Frequency must be included as a factor. The word recognition model can easily account for frequency effects since it already incorporates frequency as a determinant of threshold. When a word is processed by the semantic system it feeds back to the logogen system information about the semantic attributes recently used. If we assume two logogens share an equal number of attributes with the word just processed the one with the higher frequency will be closer to threshold, i.e. more likely to be produced in a word association situation or will be more facilitated in a word recognition situation.

These assumptions can also account for the superiority of antonyms to synonyms in the present experiment. It is possible that in the comparison of synonyms and antonyms number of features-in-common was confounded with frequency. If a word has an antonym of higher frequency than its synonyms, then the difference in the number of shared attributes may be outweighed by the difference in threshold. For example, given the word KING the word MONARCH has the same list of semantic attributes. However, the word QUEEN has a similar, but not identical, list of semantic attributes and being of higher frequency than MONARCH is more likely to be given as an associate to KING. This problem of probability of a word being given as an associate and frequency of the word is followed up in the next experiment.

Conclusions:

The conclusions of the present experiment are that time is a more important variable in associative priming effects than intervening items. At short time intervals (approximately 1 second and less) facilitation is the result of automatic processes. These processes are closely involved with the frequency of occurrence of the words used. At longer intervals facilitation is not dependent on frequency, possibly because the longer time interval allows the use of conscious strategies. Number of shared features also determine amount of facilitation.

It is suggested that this factor also interacts with word frequency.

These results suggest that a model of word recognition which includes frequency of occurrence as a variable is to be preferred to Clark's theory of word associations which does not include frequency as a factor in either word associations, or by inference in associative priming effects.

Introduction:

The results of Experiment 4 suggest that associative priming effects (also possibly word associations) are the result of an interaction between number of shared semantic attributes and frequency of the associate. Let us borrow the "hydraulic analogy" of Broadbent (1967). Broadbent describes the analogy as follows: "Let us suppose a vast array of test tubes, each partly full of water, and each corresponding to a word in the language. The choice of one tube corresponds to perception of a word, and the probability of choice of any tube is greater when the level of water in it is higher." The more frequent the word the higher the level of water in its test tube. The effect of detecting one of a word's defining set of attributes is to add a drop of water to its test tube. Thus according to the word recognition model the effect of recognizing one word is that other words sharing some of its semantic attributes have the water level raised in their test tubes. The rise in the water level will be dependent upon the number of shared

EXPERIMENT 4A

The relationship between Rated Association Value, Facilitation and Frequency.

Introduction:

The results of Experiment 4 suggest that associative priming effects (also possibly word associations) are the result of an interaction between number of shared semantic attributes and frequency of the associate. Let us borrow the "hydraulic analogy" of Broadbent (1967). Broadbent describes the analogy as follows: "Let us suppose a vast array of test tubes, each partly full of water, and each corresponding to a word in the language. The choice of one tube corresponds to perception of a word, and the probability of choice of any tube is greater when the level of water in it is higher." The more frequent the word the higher the level of water in its test tube. The effect of detecting one of a word's defining set of attributes is to add a drop of water to its test tube. Thus according to the word recognition model the effect of recognizing one word is that other words sharing some of its semantic attributes have the water level raised in their test tubes. The rise in the water level will be dependent upon the number of shared attributes. Thus the probability that the test tube of the word with most shared attributes will have the highest water level in the array is increased. However, the test tubes all started with different water levels and so it is possible that some other word although having fewer shared attributes being more frequent and consequently starting with a higher water level may still be higher than the word with most shared attributes.

The aim of this experiment was to gain further evidence for the plausibility of this two-factor theory of associative priming effects. By using a task with some similarity to normal word associations it was hoped to examine more closely the assumption made in Exp. 4 that the processes underlying word associations overlap to a large extent with the processes underlying associative priming effects. The task used in the present experiment required subjects to rate the probability of a given word being given as an associate to another word. If there was considerable agreement between pairs of words rated association value and their facilitation scores in Exp. 4 then the assumption of overlap of processes would be supported.

Again predictions can be made for this experiment concerning different kinds of associations. The predictions from Clark's theory (1970) for rated association values are the same as for Exp. 4:

Antonyms > Conceptual Associates > Parallel Associates > Synonyms.

The predictions for the word recognition model are less clear cut since there is the possibility of frequency outweighing number of shared attributes, but other things being equal they are:

Synonyms > Antonyms > Conceptual Associates ≏ Parallel Associates.

Method:

Materials: The 48 pairs of associated words used in no intervening items condition of Exp. 4. were used. There were 12 pairs of antonyms, 12 conceptual associates, 12 parallel associates and 12 synonyms. Use of these materials unfortunately precludes varying kind of association and frequency independently but they do enable a comparison between R.A.V. and facilitation scores.

Subjects: 20 postgraduates and lecturers in the Department of Psychology of the University of Stirling acted as subjects. Relatively "sophisticated" subjects were used to facilitate understanding of the instructions.

Instructions: Subjects were asked to make an estimate of the probability that the second word of a pair would be given as an associate to the first word in a free association test. Subjects were told to give a rating of 7 to pairs where they thought the first word would always elicit the second and a rating of 1 to pairs where they thought the 1st word would never elicit the second.

Results: For each pair of words their rated association value was calculated as a percentage of their total possible score. The degree of agreement between subjects in their rating was calculated by Kendall's Coefficient of Concordance.

$W = 0.64 \quad X^2 = 597.9 \quad d.f. = 47 \quad p < 0.001$

There is a highly significant correlation between subjects indicating the reliability of the results.

The mean rated association value for each kind of association is shown in Figure 21.

Analysis showed an overall significant difference between the kinds of associations. ($F(3,33) = 26 \; p < 0.001$).

Pairwise comparisons by Newman-Keuls are shown in Table 14.

	Antonyms	Parallel Assocs.	Conceptual Assocs.	Synonyms
Antonyms		0.05	0.01	0.01
Parallel Assocs.			0.01	0.01
Conceptual Assocs.				N.S.
Synonyms				

Pairwise Comparisons (Newman-Keuls)

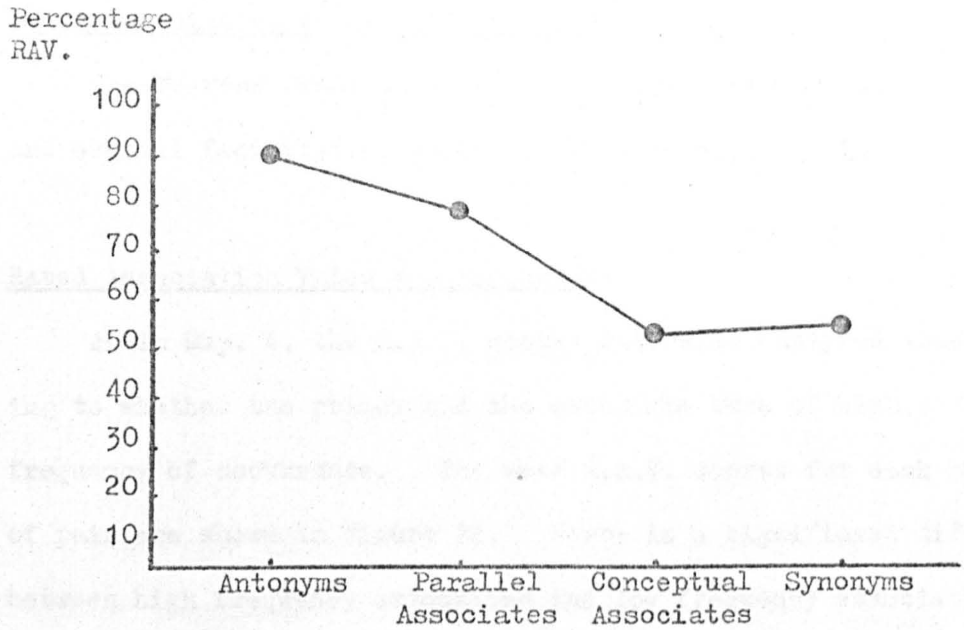
TABLE 14.

	Mean	S.D.
Antonyms	90.3	8.7
C.A.s	53.4	8.1
P.A.s	79.7	18.1
Synonyms	56.1	11.2

Means and standard deviations for Fig.21.

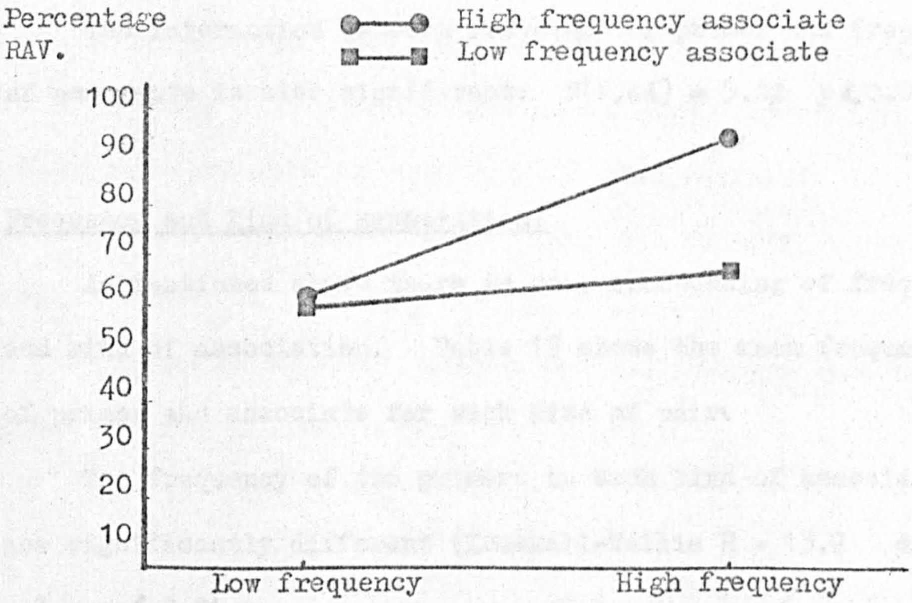
	Mean	S.D.
Low frequency primer: low associate	60.5	16.3
high "	61.9	10.7
High frequency primer: low "	67.3	23.0
high "	90.3	10.5

Means and standard deviations for Fig.22.



Mean rated association value for each kind of association.

Figure 21.



Mean rated association value according to frequency of primer and associate.

Figure 22.

Rated Association Value and Associative Priming Effects:

The Pearson Product moment correlation between R.A.V. and overall facilitation scores in Exp. is +0.39 ($p < 0.01$).

Rated Association Value and Frequency:

As in Exp. 4. the R.A.V. scores were also analysed according to whether the primer and the associate were of high or low frequency of occurrence. The mean R.A.V. scores for each kind of pair are shown in Figure 22. There is a significant difference between high frequency associates and low frequency associates: $F(1,44) = 7.14$, $p < 0.01$.

There is a significant difference between high frequency primers and low frequency primers: $F(1,44) = 14.99$ $p < 0.01$.

The interaction between frequency of primer and frequency of associate is also significant: $F(1,44) = 5.22$ $p < 0.05$.

Frequency and Kind of association:

As mentioned above there is some confounding of frequency and kind of association. Table 15 shows the mean frequency of primer and associate for each kind of pair.

The frequency of the primers in each kind of association are significantly different (Kruskall-Wallis $H = 13.2$ d.f. $= 3$ $p < 0.01$.

The frequency of the associates are also significantly different (Kruskall-Wallis $H = 7.9$, d.f. $= 3$ $p < 0.05$).

	Antonyms	Conceptual Associates	Parallel Associates	Synonyms
Primer	1526	325	740	226
Associate	1573	336	642	1125

Mean frequency per $4\frac{1}{2}$ million words (Lorge Count).

TABLE 15.

This means it is possible that the differences between kinds of associations on R.A.V. scores are due to frequency effects or conversely that frequency effects are due to different kinds of associations.

It is still possible to examine the effects of frequency by correlating R.A.V. scores with frequency scores within each kind of association, thus holding kind of association constant. Table 16 shows R.A.V. scores correlated with frequency of both primer and associate for each kind of association.

Correlation of R.A.V. scores over all kinds of associations with frequency of primer $r_s = +0.43$ $p < 0.01$

Correlation of R.A.V. scores over all kinds of association with frequency of associate $r_s = 0.31$ $p < 0.05$.

	Antonyms	Conceptual Associates	Parallel Associates	Synonyms
Frequency Primer	0.6*	-0.52*	0.57*	0.06
Associate	0.39	0.52*	-0.18	+0.35

* $p < 0.05$.

Spearman Rank Correlations of Rated Association Value and Frequency.

TABLE 16.

The confounding of kind of association and frequency:

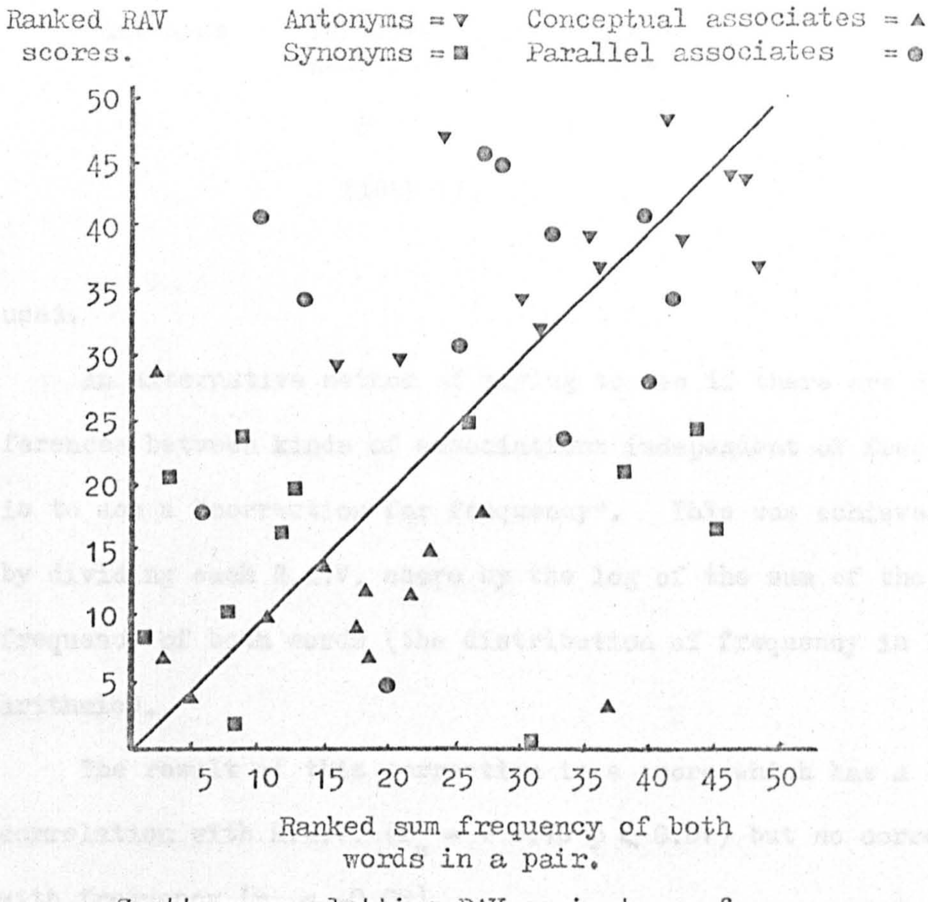
As has been noted it is not clear whether the differences shown in Figure 21 are caused by the kinds of associations or the differences in frequency. In an attempt to disentangle these variables a scattergram (see Figure 23) was plotted, plotting ranked R.A.V. scores against the ranked sum of the frequency of both words for each pair. The correlation between these variables (Spearman's r_s) = +0.502, $p < 0.01$.

A line of unit slope was drawn through the origin. It was argued that if there were no differences between kinds of associations scores on R.A.V. then for any given kind of association there should be an equal number of points lying above and below the lines. Table 17 shows the number of points for each kind of association lying above the line.

This distribution does not differ significantly from chance. There is however some evidence of a trend in the data. Since this test is rather weak another method of analysis was

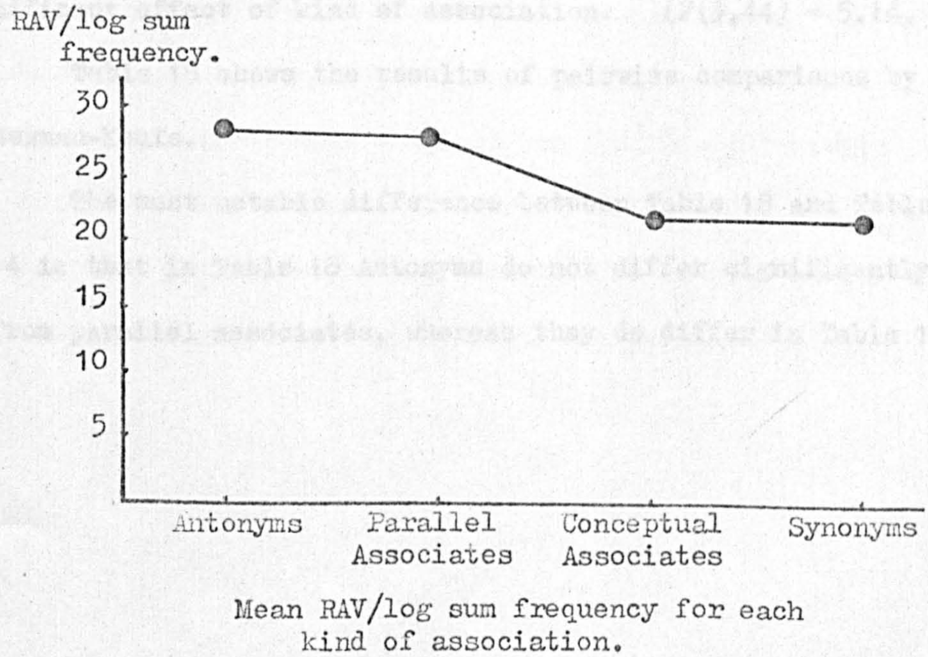
	Mean	S.D.
Antonyms	27.5	2.7
C.A.s	20.6	6.1
P.A.s	26.6	6.4
Synonyms	20.9	6.0

Means and standard deviations for Fig.24.



Scattergram plotting RAV against sum frequency.

Figure 23.



Mean RAV/log sum frequency for each kind of association.

Figure 24.

Antonyms	Parallel Associates	Conceptual Associates	Synonyms
8	8	6	2

TABLE 17.

used.

An alternative method of trying to see if there are differences between kinds of associations independent of frequency is to use a "correction for frequency". This was achieved by dividing each R.A.V. score by the log of the sum of the frequency of both words (the distribution of frequency is logarithmic).

The result of this correction is a score which has a high correlation with R.A.V. ($r_s = +0.76$ $p < 0.01$) but no correlation with frequency ($r_s = -0.08$).

Figure 24 shows the mean corrected R.A.V. scores for each kind of association. An analysis of variance produced a significant effect of kind of association. ($F(3,44) = 5.14$, $p < 0.01$).

Table 18 shows the results of pairwise comparisons by Newman-Keuls.

The most notable difference between Table 18 and Table 14 is that in Table 18 Antonyms do not differ significantly from parallel associates, whereas they do differ in Table 14.

	Antonyms	Parallel Assocs.	Conceptual Assocs.	Synonyms
Antonyms	.	N.S.	0.05	0.05
Parallel Assocs.			0.05	0.05
Conceptual Assocs.				N.S.
Synonyms.				

Significance levels of pairwise comparisons. (Newman-Keuls).

TABLE 18.

Discussion:

Associative Priming Effects and Word Associations:

The significant correlation between R.A.V. and facilitation supports the assumption that association processes and associative priming effects share some underlying mechanisms. While the R.A.V. technique may not necessarily produce exactly the same results as collecting association data in the usual way, it is assumed that the results obtained are sufficiently similar to make the comparisons valid. Both word associations and associative priming effects are assumed to be by-products of normal processes of word recognition and language use.

Kind of Association and Frequency:

These findings have shown that R.A.V., like facilitation, is dependent on both kind of association and frequency of the words in the pairs. It has become apparent in the analysis of these results that any model of word recognition and associations must include both kind of association and word frequency as factors. Clark (1970) is correct in trying to relate word associations to normal speech processes but such a feature-based model is inadequate if it ignores frequency as a variable.

As can be seen from Figure 22 associated pairs with both words of high frequency differ dramatically from the other three kinds of pairs, which have relatively similar scores. Not only do members of these high frequency pairs share many attributes but they tend to form common idioms. An idiom is assumed here to be a phrase of two or three words which occurs frequently in the language. The words may or may not be connected by a function word such as 'and' (e.g. BLACK-WHITE, LOVE-HATE, GIVE-TAKE, BREAD-BUTTER). Clark does include an idiom completion rule in his list of rules for association production but its priority relative to the other rules is uncertain. It is difficult to decide whether pairs such as LOVE-HATE are products of the minimal contrast rule or the idiom completion rule. It is in the nature of language use in the real world that idioms tend to be contrastive in nature (e.g. LOST-FOUND) or reflect objects that naturally occur together (e.g. LOCK-KEY). These pairs are consistent with both a feature-change rule and the idiom-completion rule. (Of course, there are some pairs e.g. COTTAGE-CHEESE that are only consistent with the idiom-completion rule). Generally speaking synonyms and subordinate-superordinate pairs do not frequently occur in idioms. One approach to this finding is to assume that if a word frequently occurs in an idiom then membership of this idiom will be included in its defining set of semantic attributes. So far the model of word recognition has made no assumption about ordering of attributes (cf. Clark's theory which is based on ordered lists of features). It may be that attributes are ordered in terms of level of saliency, i.e. some attributes carry more weight in that they contribute more to the logogen reaching threshold than other attributes. Thus the attribute "IS A NOUN" may contribute little to the count of detected attributes whereas "FREQUENTLY OCCURS WITH BREAD" may contribute greatly. Possibly the saliency of an attribute is inversely related to the number of logogens sharing that attribute. It may also be the case that the semantic system feeds back into the logogen system information about the last word processed in the order of the word's most salient attributes first. Attributes concerning membership of idioms are likely to be highly salient and this information is likely to be the first feedback to the logogen system from the semantic system. This is consistent with the finding in Exp 4. that the pairs with both words of high frequency (i.e. the idioms) showed a high degree of facilitation at the short interval (300 ms.). It is reasonable that a mechanism should exist for rapid feedback of information about idiom membership since in normal language use the words in idioms would follow each other almost immediately. Conversely in normal speech if after 1 second the idiom has not been completed it is probable that the word is not being used in its idiomatic context. After 1 second idioms tended to be inhibited rather than facilitated (see Exp. 4, Fig. 19). As mentioned in Exp. 4, after a 2 second interval other strategies may operate independently of the normal word recognition processes.

What about word pairs that do not form idioms? The assumption of the word recognition model that number of shared attributes will be the dominant variable (if frequency is controlled) seems to be over-simple. In particular there is some indication that antonyms are still rated higher than similar frequency synonyms even after idiomatic pairs have been removed. No firm conclusions can be drawn since the number of pairs suitable for comparison is very small. But antonyms such as DIRTY-CLEAN are still rated higher than similar frequency synonym pairs such as STARVED-HUNGRY.

Deese (1965) points out that most frequent adjectives tend to form pairs defining some dimension. As such the definition of an adjective is very closely related to that of its opposite and information about its antonym is likely to be more salient than information about words of similar meaning. The same is true for nouns that can be represented as polar opposites.

It is interesting to note in Table 16 that for conceptual associates frequency of the instance is negatively correlated with R.A.V. This suggests that category membership is a more salient attribute for low frequency words than for high frequency. It is more important to be able to classify a rare word (e.g. MINNOW) which may have few stored attributes than to classify a common word (e.g. CAR) which has many stored attributes. As Collins and Quillian (1969) have stated it is possible to gain much more information than may be available from the instance alone by classifying it.

Conclusions:

The present experiment suggests that word associations and associative priming effects share the same underlying processes and that these processes are part of normal word recognition and language use. Exps. 4 and 4A. indicate that the following may be a plausible account of word association production and associative priming effects:

When a word is recognized (i.e. its logogen reaches threshold) the logogen system makes available to the semantic system the attributes that characterize the word. Attributes are made available according to their saliency-attributes with a high information content being made available first. The semantic system feeds back to the logogen system information about the most salient attribute which activates this attribute detector in logogens containing it as part of their defining set. If this feedback is sufficient to cause a logogen to reach threshold then this word is produced as an associate. It is this part of the process which is influenced by frequency of occurrence. High frequency words are more likely to reach threshold as a result of the first input of feedback. If no logogen has reached threshold after the first feedback, information about the less salient attributes is input to the logogen system until a logogen reaches threshold.

EXPERIMENT 5

Context and Ambiguity.

Introduction:

Ambiguity has long held an important place in the study of the psychology of language, dating back to Hughlings Jackson in the last century. The major problem is that although ambiguity is very frequent in the English language we are rarely troubled by it.

Ambiguity in a language can arise from a number of sources. One source can be called lexical ambiguity. This arises from the fact that many English words have more than one meaning and we may be uncertain which one is intended. E.g. "He has lost his bat." Which meaning of "bat" the listener assigns may depend on whether "he" refers to a zoo-keeper or a cricketer.

Another source of ambiguity may be called phrase structure ambiguity. In this case we are uncertain what phrase structure fits a particular sentence. E.g. "They are visiting sailors". It is not possible to decide if it is the sailors who are doing the visiting or whether they are being visited. Which meaning is intended will depend on some wider context.

Another source of ambiguity may not be resolved by either knowing the meaning of the lexical items or knowing the appropriate phrase structure. This can be called, after Chomsky, deep structure ambiguity . E.g. "The shooting of the hunters was terrible." This sentence has only one phrase structure but two deep structures. Again which one we assign will depend on what we know of the context in which the sentence occurs.

The present experiment deals only with the first kind of ambiguity, that is, ambiguity due to words having more than one meaning. Much of the previous work on ambiguity has been concerned with ambiguity arising from other sources but whatever the source a problem common to all is whether only one meaning or all possible meanings of a word or sentence are computed.

One of the important functions of the word recognition model proposed earlier is to integrate information from different sources so that the correct word in any situation is perceived. Each logogen defines its words by semantic, graphemic and acoustic feature lists. Different words may have lists in common. Two words may have the same meaning, i.e. have similar semantic feature lists (e.g. TINY-SMALL). Two words may have identical acoustic features (e.g. SEA-SEE). In this experiment we are concerned with words that have identical graphemic and acoustic features but have different semantic feature lists (e.g. BAT, FILE, SOLEetc.)

Before considering how the word recognition model decides which is the intended meaning of an ambiguous word let us first examine the previous approaches to ambiguity.

One of the earliest approaches was by Lashley (1951) who proposed a hypothesis which has since been called the Garden Path Hypothesis. Lashley took as evidence sentences of the sort "Rapid / rajtiŋ / with his uninjured hand saved from loss the contents of his capsized canoe." Lashley argued that people process only one meaning at a time so that if one meaning becomes inappropriate we have to then process the other meaning. So that when the first interpretation of / rajtiŋ / is seen to be wrong when the word "capsized" is heard the listener has to go back to the beginning and reprocess the sentence in the light of the new meaning. In this theory context determines which meaning of an ambiguous word we access, or if no context is available we choose the meaning with the higher a priori probability occurrence. While this view is consistent with our introspections that we are only aware of one meaning, as Mackay (1970) points out this does not exclude the possibility that all meanings are accessed at some subconscious level.

Mackay (1970) presents a model which he calls the Exhaustive Computation Hypothesis. This model conflicts with the Garden Path Hypothesis by assuming that all possible meanings of an ambiguous item are computed at an early stage of processing and at some later point a single reading is selected for attention. Mackay has also stated a stronger version of this hypothesis which he calls the Perceptual Suppression Theory. This theory proposes that all meanings of an ambiguous item are processed.

in parallel at an unconscious level. In order to perceive one meaning the other meanings must be suppressed. The time for suppression depends on the a priori likelihood of the suppressed meaning, independent of context. Context does have an effect however, by strengthening the activation of the appropriate meaning. Mackay quotes evidence from ambiguous visual figures (e.g. the Necker cube), physiological evidence for suppression (Mountcastle, 1961) and extensive experimental evidence.

In summary the Garden Path Hypothesis suggests that context and a priori probability determine which <u>one</u> of the possible meanings will be accessed. The Exhaustive Computation Hypothesis and the Perceptual Suppression Theory suggest that all possible meanings are accessed at an early stage and only later does context and a priori probability combine to determine which meaning will be selected.

The evidence for these theories is itself ambiguous. Evidence for the Garden Path Hypothesis has been found by Carey, Mehler and Bever, (1970), Foss, Bever and Silver (1968), Foss and Jenkins (1973) and Foss (1970). Evidence supporting the Exhaustive Computation Hypothesis has been produced by Fodor, Garrett and Bever (1968), Lackner and Garrett (1973), Mackay (1966, 1970) and Conrad (1974). I shall return later to possible reasons for these contradictory results.

A related problem was raised by Miller (1970). The problem he posed is "Can we show that there exists a subjective lexicon which is an isolable sub-system in memory? The subjective lexicon refers to our intrinsic knowledge of the meanings of words. It is a set of concepts which have been acquired and refined over a relatively long period of time; a set of dictionary entries which define words, specify their selection restrictions and perhaps provide information about their possible syntactic roles." (From Conrad, 1974 p 130).

Miller is suggesting the existence of a "lexical look-up process which is not influenced by context." This is a similar position to the Exhaustive Computation Hypothesis and the Perceptual Suppression Model. It implies that all meanings of a word will be accessed in the lexicon. Conrad (1974) found evidence consistent with this assumption.

The word recognition model assumes that the input to all logogens will be the same. Words with two meanings will be represented by two different logogens. These logogens will have identical graphemic and acoustic feature lists but different semantic feature lists. Given a sensory input which logogen reaches threshold first will depend on both the level of the threshold determined by frequency of occurrence and the activation of semantic feature detectors by the context/semantic system. In the absence of any context the more frequent logogen will reach threshold first. The less frequent logogen may reach threshold first if a sufficient number of its semantic features have been activated by prior context. In assuming some level of accessing of all possible meanings the word recognition model is related to the Exhaustive Computation Model and the Perceptual Suppression Theory. However, the logogen system is not identical with Miller's subjective lexicon. Miller's lexicon operates completely independently of context. In the word recognition model while input to the logogen system is independent of context the output is explicitly affected by context.

The present experiment is an attempt to test these different models. The experiment uses a task where a subject uses a word with one meaning and then has to process either (A) the same meaning of the same word, (B) a different meaning of the same word, or (C) a completely different word. Different predictions can be derived for each model.

It is assumed here that repetition of a word leads to a decrease in its recognition time (e.g. Bertelson (1961) has shown that RTs to a repeated signal are shorter than to 'new' signals). According to the Garden Path Hypothesis if sufficient context is provided to indicate which meaning of a word is intended then only that particular meaning will be accessed. RTs will be shortened if the same meaning of the item is tested. Since different meanings have not been accessed they should be treated as completely new items. The prediction for the Garden Path Hypothesis is:

$$A < B = C.$$

The simplest assumption of the Exhaustive Computation

Model is that all meanings will be accessed and will result in equal facilitation for either meaning. The prediction for the Exhaustive Computation Model (the same as Miller's isolable subjective lexicon model) is:

$$A = B < C.$$

The word recognition model is more specific than the Exhaustive Computation Model. Facilitation occurs through activation of feature detectors in the logogen. The logogen representing the meaning of the word used will have graphemic, acoustic and semantic features activated. The logogen representing the meaning not used will only have graphemic and acoustic features activated. This implies that although both meanings will be facilitated relative to the control condition the meaning actually used will be more facilitated than the meaning not used. The prediction is:

$$A < B < C$$

The Perceptual Suppression Theory although assuming that all meanings will be accessed makes the further assumption that non-selected meanings will be suppressed. This suggests that processing non-primed meanings will take longer than control words. The prediction is:

$$A < C < B$$

Method:

Materials: 60 nouns with a frequency of at least 50 per million were selected from the Thorndike-Lorge Count. All the words had at least two distinct meanings (e.g. PARTY, POKER, BAT). The words were combined with a qualifier to indicate one of the meanings of the word (e.g. IRON POKER, POLITICAL PARTY). These 60 phrases formed the "priming phrases".

60 other two-word phrases were formed for the test phrases. Half the phrases consisted of a noun with an acceptable qualifier (e.g. ELECTRIC IRON) and half consisted of a noun with an unacceptable qualifier (e.g. SAFE CLOUD).

Each priming phrase was paired with a test phrase to produce the following conditions. (+ = acceptable test phrase, - = unacceptable test phrase).

A+ : Test phrase used same noun with same meaning as priming phrase. The qualifier in the test phrase was not an associate of the qualifier in the priming phrase (N=10).

B+ : Test phrase used same noun with a different meaning from the priming phrase. (N = 10).

C+ : Test phrase used completely different noun from priming phrase. (N = 10).

AB- : Test phrase used same noun as priming phrase. Since the qualifiers in the test phrase were not acceptable qualifiers of either meaning of the noun this condition balances the design for both A+ and B+. (N = 20).

C- : Test phrase used different noun from priming phrase (N = 10).

Examples of each condition are given in Table 19.

	Priming Phrase	Test Phrase		
A+	SOCIAL CLUB	YOUTH CLUB	Same Meaning	Acceptable
B+	LOUD REPORT	WRITTEN REPORT	Different Meaning	Acceptable
C+	LEAFY TREE	DIRTY MARK	Different Noun	Acceptable
AB-	COUNCIL RATES	SQUARE RATES	Same Noun	Not Acceptable
C-	LAST WILL	GENEROUS PLAIN	Different Noun	Not Acceptable

Examples of Each Condition.

TABLE 19.

Subjects: 16 psychology undergraduates of Stirling University participated in the experiment to fulfil a course requirement.

Procedure: The subject sat facing a tachistoscope. On the table in front of the subject was a sheet of paper with the 60 priming phrases printed on it. Each trial consisted of two parts - a priming phrase and a test phrase. In the priming part of the trial the subject was instructed to read a cue phrase and write a sentence using this phrase.

In the test part of the trial the paired phrase was presented to the subject on the tachistoscope. The subject pressed a button with his right hand if he thought the phrase was an acceptable combination of words and a left hand button if he thought the combination was not acceptable. The subject's response terminated the display and stopped a millisecond stop-clock which began timing from the start of the display. Acceptability was defined for the subject as excluding any metaphorical or poetic use of the words.

Each subject received 6 practice trials followed by the 60 test trials. Order of presentation was random. Each session lasted approximately 40 minutes.

Equipment: The equipment used in this experiment was the same as that used in Exp. 1. The materials were displayed on the "mousetrap" tachistoscope.
The phrase was typed on cards. The phrase is concealed in the tachistoscope by a shutter. When E presses the "start" button the shutter is lowered displaying the phrase to S. Simultaneously a millisecond stop-clock is started. When S responds by pressing one of two buttons the shutter is raised and the clock is stopped.

Results: For each subject there was a written record of how he had interpreted each priming phrase. This enabled a check to be made to ensure that the subject had interpreted the phrase in the way intended by the Experimenter. Out of a total of 960 trials a difference of interpretation occurred on only 3 trials. Data from these trials were not included in the analysis.

For the test trials the mean error rate was 8.5%. Data from correct responses only were used in the analysis. One item intended as an acceptable combination was responded to as an unacceptable combination by 66% of the subjects. Data from this item was excluded from the analysis.

Mean RT for each condition is shown in Figure 25.

Analysis of "acceptable" phrases:

A v. B v. C. (Reaction times).

By subjects $F(2,30) = 8.2$ $p < 0.01$

By materials $F(2,26) = 4.34$ $p < 0.05$.

Min $F^1(2,49) = 2.84$ $p < 0.1$.

(N.B. Min F^1 is an underestimate of F^1. Computation of F^1 would require calculation of missing data. Some estimate of how conservative Min F^1 is can be obtained by calculating Max F^1.

Max $F^1(2,49) = 3.3$ $p < 0.05$)

Pairwise comparisons are shown in Table 21 .

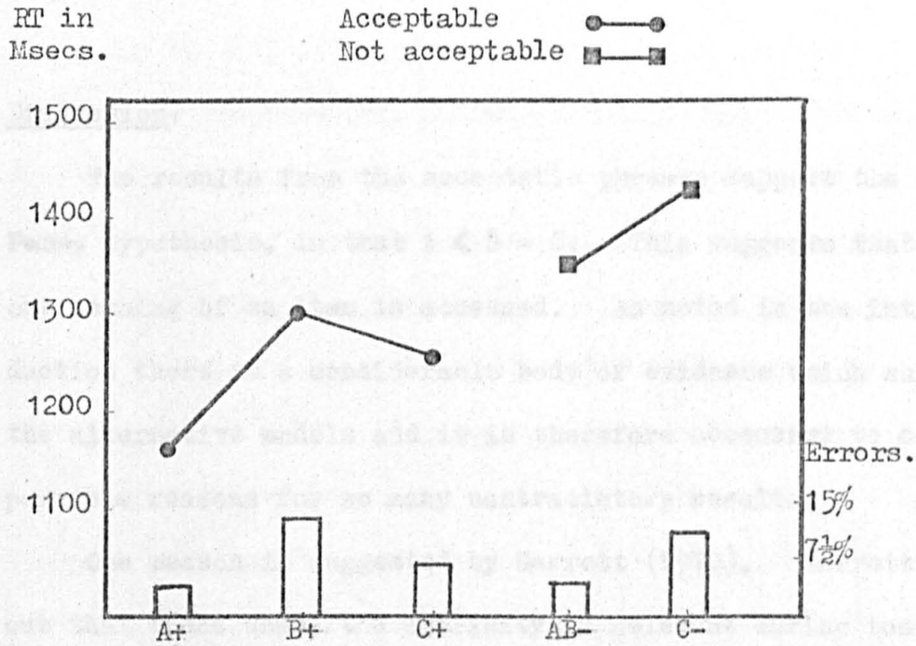
	A+	B+	C+
A+		0.01	0.05
B+			N.S.
C+			

Pairwise Comparisons (Newman-Keuls).

TABLE 20.

	Mean	S.D.	% error
A+	1165	338.1	10
B+	1306	378.7	13.7
C+	1279	303.3	7
AB-	1370	94.0	3
C-	1431	119.9	10

Means, standard deviations and % errors for Fig.25.



Mean RT and percentage error
for each condition.

Figure 25.

A v. B v. C (errors). $X^2 = 6.25$ $df = 2$ $p < 0.05$.

Non-acceptable phrases:

AB- v. C-. AB- (same word repeated) was on the average 69 msecs. faster than C- (different word) $t = 1.86$ $df = 17$ $p < 0.05$ (one-tailed test).

Discussion:

The results from the acceptable phrases support the Garden Path hypothesis, in that $A < B = C$. This suggests that only one meaning of an item is accessed. As noted in the introduction there is a considerable body of evidence which supports the alternative models and it is therefore necessary to consider possible reasons for so many contradictory results.

One reason is suggested by Garrett (1970). Garrett points out that tasks where the ambiguity is relevant during testing tend to support the Exhaustive Computation Model and the Perceptual Suppression Theory, whereas tasks in which the ambiguity only becomes relevant later tend to support the Garden Path Hypothesis. In the present experiment the ambiguity of a word only became relevant in the testing phrase after the priming phrase. This experiment is thus to be grouped with those experiments which support the Garden Path Hypothesis.

The predictions for the present experiment by the word recognition model were based on making the simplest of assumptions about the model. If some more complex assumptions are made then this model is capable of accounting for both sets of contradictory findings.

Inasmuch as different meanings of an ambiguous word have, by definition, identical physical representations then at some low level of sensory processing they must be treated identically. It is the results of some operations (feature extraction?) on this sensory representation that is made available to the logogen system. The logogen system is the meeting place of different forms of coding. We are interested here in how one of two (maybe more) possible meanings are assigned to an ambiguous sensory representation. The two different meanings are represented by two different logogens. We shall assume for the moment that the two meanings are equiprobable and therefore the logogens have the same threshold value. We shall not distinguish graphemic and phonemic codes, preferring to contrast sensory codes with semantic codes.

Given an input of graphemic/phonemic attributes which belong to more than one meaning (e.g. BAT) the acoustic and graphemic attribute detectors of both logogens are activated. Which one reaches threshold first (i.e. the meaning assigned) will depend on how many of their defining sets of semantic attributes have been activated via the context system. E.g. for the word BAT in CRICKET BAT some of the semantic attribute detectors will have been activated by CRICKET for one meaning but not for the other. Thus both meanings are given a graphemic/acoustic code but only one meaning is given a semantic code.

By assuming different time spans for the different codes one can account for Garrett's observation that testing for effects of ambiguity during processing supports the accessing all meanings hypothesis and testing after processing supports the accessing one meaning hypothesis. Baddeley (1968, 1972) produced evidence showing that acoustic coding is important in short term memory and that semantic coding is important in long term memory. While there is some evidence against this generalization (e.g. Wickens, 1972; Schulman, 1972) it is consistent with the commonsense view that the prime function of language is the communication of meaning. It is necessary to retain some physical representation of a sentence while the meaning is encoded. Once the meaning is encoded the exact physical form of the sentence is largely irrelevant and may be allowed to decay (cf Bartlett, 1932).

It is assumed here that for ambiguous words one of the meanings is encoded graphemically, acoustically and semantically while the other only encoded graphemically and acoustically. If one tests for ambiguity after presentation both logogens are activated to some extent. However, the activity in the logogen whose graphemic/acoustic features alone are activated quickly decays, so that the logogen which has produced a semantic representation is the only one still active if the testing is later.

Let us compare the present experiment with the recent experiment of Conrad (1974) which produced evidence of accessing all meanings.

In Conrad's experiment subjects listened to a sentence which they had to hold in memory while they named the colour of a following word. The sentences contained an ambiguous word and one meaning of the word was indicated. E.g. THE GIRL FOUND A NICKEL. The word which followed was either an appropriate or an inappropriate category of the ambiguous word. (For the example above : MONEY or METAL). Conrad found a significant interference on the colour-naming task for both the appropriate and inappropriate category names.

Why should Conrad's results differ from the present experiment? In the present experiment in both the priming and test phrases the selection of one meaning of the word was important. In the priming phrase one meaning is selected and a semantic representation is formed. The non-selected meaning is only represented by an acoustic/graphemic code which has decayed by the time of testing. In Conrad's experiment the meaning of the word is not essential to the memory task and the sentence may only be represented acoustically and not semantically. As stated above in normal language use semantic representations are usually retained not acoustic/graphemic representations but the choice of code is to some extent under the control of the subject and the choice of strategy may depend on the task. This point is made by Herriot (1974, p 172) "...while speakers and listeners have to construct messages from meaning and meaning from messages, memorizers do not necessarily have to do so. They may use non-meaningful forms of coding to memorize, indicating that while the communicative function of language has an effect, this effect can be regulated by conscious control strategies."

Since Conrad required verbatim recall of the sentences her subjects may well have opted for a non-meaningful form of coding. Such a coding would not distinguish between the two meanings, thus both logogens would be activated to some extent. Since, however, the interference of the colour-naming of categories does imply some involvement of meaning, it seems necessary to assume that although neither meaning is given a complete semantic representation the non-semantic codes are still capable of activating some semantic attributes albeit at some subconscious level.

This account offers a description of what happens in sentences such as Lashley's. Only one meaning of /rajtiŋ/ reaches threshold and is used by the context/semantic system. The semantic representation used as information which is fed back to the logogen system and thus the meaning first selected will have an influence on subsequent words in the sentence. The meanings assigned to other words in the sentence may be inappropriate as well as the obviously ambiguous word. When the word "capsized" is reached the alternative meaning of /rajtiŋ/ reaches threshold and has an effect on the assigned meanings of the other words, causing the appropriate meanings to reach threshold. The exact mechanism must be very complex since the semantic representation of the sentence as a whole is completely altered. The way in which the meanings of individual words are combined to form complex semantic structures is not yet clear. (see for example Branford and Franks, 1971; Bransford and Johnson, 1973).

Much of the present discussion has assumed that the different meanings of an ambiguous word are equiprobable. This is clearly not the case (cf. Kausler and Kollasch, 1970). One meaning of a word tends to be more common than the others. Hogaboam and Perfetti (1975) have commented that the relationship between the primary and secondary meanings must be considered in any model of ambiguity resolution. (see also Mackay, 1968). The word recognition model allows for the fact that there will be a bias towards the more common meaning since the threshold of its logogen will be lower, but the present experiment made no systematic examination of the effects of dominance of meaning. Primary and secondary meanings were assigned at random to conditions. By pooling results over primary and secondary meanings some valuable information may have been lost. The next experiment was designed to specifically investigate the effects of primary versus secondary meanings in a controlled manner.

Conclusions:

The results of this experiment appear to support a model that assumes only one meaning of an ambiguous word is accessed. However, it is argued that such a conclusion reflects a general confusion as to what is meant by "accessing". Much of this confusion can be avoided by considering instead how a word may be represented internally. A word can be coded as a set of graphemic, acoustic or semantic features. Each type of code has its own function. Ambiguous words are represented by the same acoustic/graphemic features but by different semantic features.

In normal language use each code has a typical life span. The graphemic code is usually very short, possibly less than $\frac{1}{2}$ second (iconic memory?). The acoustic code has a life span of approximately 2 seconds (STM?) but has the possibility of being retained by articulatory rehearsal. The semantic code has a virtually unlimited time span (LTM?).

The effects of ambiguity will depend on how the different meanings are coded at the time of testing. Testing for a short time after presentation both logogens' acoustic/graphemic attributes will be active leading to the conclusion reached by those experiments that found evidence of access of all meanings. Testing after a longer interval only the meaning which has been given a full semantic representation will be active leading to the conclusion that only one meaning has been accessed.

If coding by acoustic features is more useful for the task and a strategy is adopted to maintain this code then evidence for accessing all meanings may be found even after a longer interval. This is because such a code will be common to both meanings.

EXPERIMENT 6

Context, Ambiguity and Frequency of meaning.

Introduction:

In Exp. 5. it was found that subjects were quicker to respond to an ambiguous word if it was used in the same sense as they had previously used it. If the meaning of the word was changed response times were the same as those for completely new words. This finding is consistent with the hypothesis that only one meaning of an ambiguous word is accessed at any one time. However, Garrett (1970) has pointed out that such evidence is found in experiments where the effect of ambiguity is examined after processing. If the effect of ambiguity is examined during processing then evidence is found for access of all meanings of ambiguous words.

It was argued in Exp. 5. that these time effects can be attributed to different forms of coding of the stimulus and that these forms of coding can be identified to a limited extent with the structural notions of short and long term memory. The problem in processing ambiguous words may lie in the decoding from an acoustic/graphemic code into a semantic code. Testing during decoding may indicate difficulty with ambiguous words but once coded semantically no difficulty may be found.

A similar point has been made by Hogaboam and Perfetti (1975, p 272). "...contrasting the various models in the disambiguation process in effect may be setting up straw men. The models that have been discussed may be best conceptualized as characterizations that might apply to any one, or more, of several levels."

Even allowing for this possibility it must be noted that the conclusions from many experiments (including Exp. 5.) are of limited value because they employed no systematic control of bias for the different meanings. The discussion in Exp. 5. assumed that the different meanings of an ambiguous word were equiprobable in their occurrence but this is clearly not the case (see Kausler and Kollasch, 1970; Perfetti, Lindsey and Carson, 1971). One meaning is usually dominant over the others. Hogaboam and Perfetti (op.cit.) argue that any model of ambiguity resolution must allow for the possibility that there may be differences between the primary and secondary meanings of an ambiguous word. Exp. 5. is subject to the same criticism that Hogaboam and Perfetti make about the experiment of Conrad (1974). In the change of meaning condition the priming phrase may involve the primary sense (S_1) and the test phrase may involve the secondary sense (S_2) or vice versa. There may be differences between the two cases ($S_1 \rightarrow S_2$, $S_2 \rightarrow S_1$). Averaging over both cases may obscure any interaction. It is also possible that there may be differences in the same meaning conditions between ($S_1 \rightarrow S_1$) and ($S_2 \rightarrow S_2$).

The present experiment was designed to investigate the priming effects of different/same meanings for both the primary and secondary senses of an ambiguous word. The task used was a sentence comprehension task (see Haviland and Clark, 1974). Pairs of sentences were presented for comprehension consecutively. An ambiguous word could appear in both sentences with the same meaning (SAME condition), in both sentences with different meanings (DIFFERENT condition) or only in the second sentence (CONTROL condition). This design was repeated for both the primary and the secondary meaning of the word.

We can now consider the predictions for the models discussed in Exp. 5. These predictions refer only to comprehension times for the second sentence in each pair. Table 21 summarizes the predictions for each model. The first two models are the simplest. It is unlikely that any one would seriously propose these models but they are included as representing the extreme positions.

1. The garden path hypothesis stated in its most basic form is that only the meaning indicated by the context will be accessed. The predictions are the same as in Exp.5. and are the same for both primary and secondary meanings.

2. The simplest version of the exhaustive computation hypothesis is that all meanings are accessed in parallel and only

and only at some later stage does context determine which meaning is selected. The predictions are again the same as Exp. 5. and the same for both primary and secondary meanings.

3. Hogaboam and Perfetti (op.cit.) describe an ordered search model which they claim is the most parsimonious description of their results. In this model the primary meaning is always accessed first and then tested against context. If this meaning is inappropriate the secondary meaning is then accessed. As Hogaboam and Perfetti admit their results are also consistent with a model such as the logogen model which assumes parallel processing of both meanings but faster processing of the primary meaning. The ordered search model assuming the primary meaning is always accessed predicts that use of the secondary meaning in the first sentence will facilitate access of the primary meaning in the second but that use of the primary meaning in the first sentence will not facilitate access of the secondary meaning.

4. The logogen model assumes that in the absence of context the primary meaning will reach threshold first. However, sometimes given sufficient context indicating the secondary meaning this meaning will reach threshold first. On the other hand, if the context in the first sentence indicates the primary meaning it is extremely unlikely that the secondary meaning will ever

reach threshold first. Thus ($S_1 \rightarrow S_2$) will never produce facilitation but ($S_2 \rightarrow S_1$) sometimes will but less often than ($S_1 \rightarrow S_1$). There is thus an asymetry between the primary and secondary meanings.

5. Mackay's perceptual suppression theory (1966, 1970) is similar to the logogen model but makes the additional assumption that to perceive one meaning of an ambiguous word the other meanings have to be suppressed. This suggests that comprehension times will be slower following a change in meaning than in the control condition. It is probable that there would be an interaction between suppression and frequency of meaning. Mackay has shown that time for suppression is dependent on saliency of meaning but for this experiment it would be necessary to know the consequences of suppression of a meaning, in particular the time course of suppression but this information is not available at present. The interaction mentioned above may be expected but the exact form it might take cannot be predicted as yet.

MEANING IN SECOND SENTENCE

Model	PRIMARY	SECONDARY
G.P.H.	Same < Different = Control	Same < Different = Control
E.C.H.	Same = Different < Control	Same = Different < Control
Ordered Search	Same = Different < Control	Same < Different = Control
Logogen	Same < Different < Control	Same < Different = Control
P.S.T.	Same < Control < Different	Same < Control < Different

Predictions from each model.

TABLE 21.

Method:

Equipment: The sentences were presented on a Digital Equipment Corporation GT40 display screen, under the control of a PDP 11/45 computer. A micro-switch was in front of the screen.

Materials: 18 ambiguous words were selected from Kausler and Kollasch (1970). This paper indicates the primary and secondary meanings of each word. Words were selected which had one clearly dominant meaning. Words were only selected which had two completely unrelated meanings and also words were only used which had nouns for both meanings (e.g. BAT, FILE, SOLE).

For each word 4 sentences were made up - 2 sentences indicating the primary meaning and two indicating the secondary meaning. These sentences werd then combined to produce the following pairs:

1 Primary - Primary,

2 Secondary - Secondary,

3 Primary - Secondary,

4 Secondary - Primary.

Two further pairs were created by combining the test sentence in each of the above pairs (i.e. the second sentence) with a completely unrelated sentence to produce

5 Unrelated - primary,

6 Unrelated - secondary.

Examples of each type of pair are shown in Table 22.

Type of pair	Priming Sentence	Test Sentence
P → P	The bat was made of wood.	The bat hit the ball.
S → S	The bat had large ears.	The bat flew in the window.
P → S	The bat was made of wood.	The bat flew in the window.
S → P	The bat had large ears.	The bat hit the ball.
U → P	Jim has a bike.	The bat hit the ball.
U → S	Jim has a bike.	The bat flew in the window.

Examples of each pair of sentences.

(P = Primary, S = Secondary, U = unrelated).

TABLE 22.

Six lists of 18 pairs were made up using 3 pairs of each of the 6 types so that each ambiguous word appeared an equal number of times in each meaning and in each condition over all lists.

Subjects: 24 psychology undergraduates of Stirling University participated as subjects to fulfil a course requirement. None had participated in Exp. 5.

Instructions: The instructions used were based on Haviland and Clark (1974), who had reported using the comprehension task successfully with similar instructions:

"In this experiment you will be presented with pairs of sentences. The sentences will appear one after the other. When the first sentence appears read it and press the button as soon as you understand it. The second sentence will immediately follow and again press the button as soon as you understand what it means. There will then be a short interval before the next pair of sentences appear. Try to work as quickly as possible."

If the subject had any questions about the task he was told that "understand" was being used in the "normal, everyday sense" of the word. No further critérion of understanding was offered so to a large extent it was left up to the subject to define "understand" for themselves.

Procedure: Each session started with 18 practice trials. In the practice trials the sentences in each pair were unrelated. No word was repeated. No clearly ambiguous words were used. After the practice trials the subject received one of the six experimental lists. 4 subjects were assigned to each list at random.

On each trial a sentence appeared and the subject responded as soon as he understood the sentence by pressing the microswitch. When the subject responded the first sentence was removed from the screen and the second sentence was immediately presented in the same place on the screen. After the subject responded to this sentence the screen went blank for approximately 2 seconds, until the first sentence of the next pair was presented. Each session lasted approximately 20 minutes.

Results: The analyses were only carried out on comprehension times for the second sentence in each pair. The analysis presented a number of problems. The reaction times were not normally distributed. Reaction times are normally skewed to the left but these results were exceptionally shewed. A further problem was that there was considerable variation both between subjects and within each subject. With only 3 observations per condition per subject this presented a problem of reliability. In the light of the above mentioned problems it was felt that the median was a better estimate than the mean. All scores reported here are based on medians.

The fact that inter-subject variance was so high and subjects varied so much in which sentences they found difficult to comprehend meant that it was impossible to obtain reliable estimates of comprehension times for each sentence so it was not possible to analyze the results over materials. This limits the generalizability of these results to other materials. (see Clark, 1973).

An analysis of variance treating subjects as a random factor produced only one significant result. Comprehension of sentences indicating the primary meaning of the ambiguous words was 88 msecs. faster than comprehension of sentences indicating the secondary meaning. ($F(1,23 = 4.5)$ $p < 0.05$). The mean median for each condition are shown in Figure 26.

Analysis of sentences indicating primary meaning.

Over all conditions Friedman's $X^2_R = 9$ df = 2 $p < 0.02$

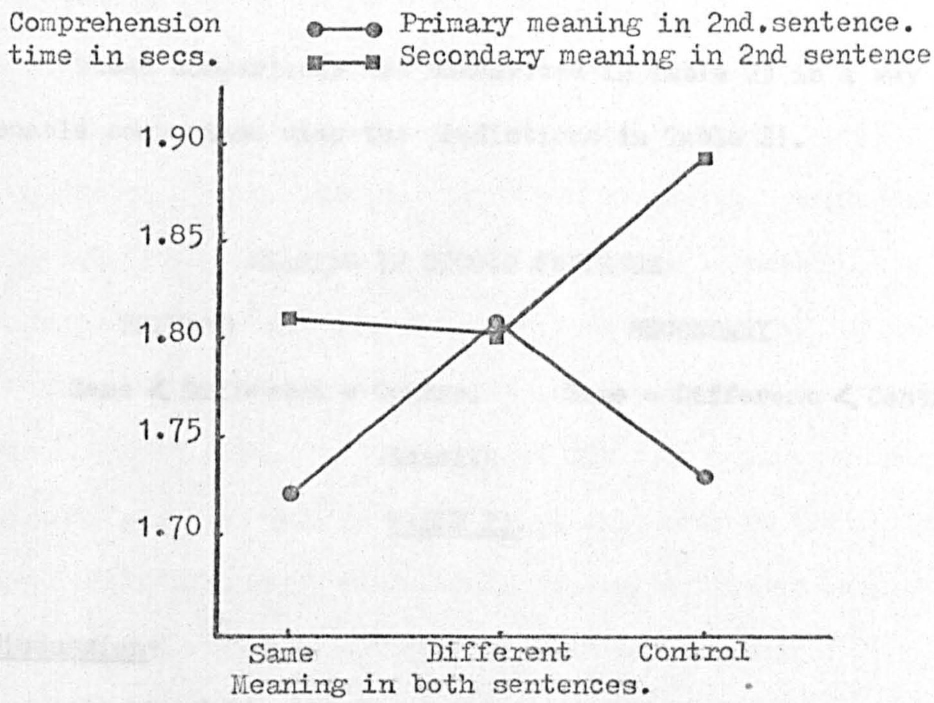
Pairwise comparisons by sign tests ($N = 24$).

SAME	v DIFFERENT	x = 6	p = 0.022
SAME	v CONTROL	x = 7	p = 0.064
DIFFERENT	v CONTROL	x = 12	N.S.

Analysis of sentences indicating secondary meaning.

	Mean	S.D.
<u>Primary meaning in 2nd sentence:</u>		
same meaning in both sentences	1716	943.0
different meaning in both sentences	1805	945.9
control	1728	830.9
<u>Secondary meaning in 2nd sentence:</u>		
same meaning in both sentences	1809	972.6
different meaning in both sentences	1803	833.3
control	1890	990.1

Means and standard deviations for Fig.26.



Median comprehension time for each condition.

Figure 26.

Over all conditions Friedman's $X^2_R = 7.75$ df = 2 $p < 0.05$

Pairwise comparisons by sign tests (N = 24).

SAME v DIFFERENT x = 9 p = 0.154

SAME v CONTROL x = 7 p = 0.064

DIFFERENT v CONTROL x = 6 p = 0.022.

These comparisons are summarized in Table 23 in a way to enable comparison with the predictions in Table 21.

MEANING IN SECOND SENTENCE	
PRIMARY	SECONDARY
Same < Different = Control	Same = Different < Control

Results

TABLE 23.

Discussion:

All the following discussion is based on the assumption that subjects perceived the meaning intended by the experimenter. It must be admitted, however, that the possibility remains that subjects could perceive some non-intended meaning. E.g. it is possible to assign a meaning of some kind to a phrase like "flying bat" even using "bat" in the sense of cricket bat. While possible it was felt that such events would be highly infrequent.

These results do not provide clear-cut support for any one of the models outlined in the introduction. However, the different pattern of results for the primary and secondary meanings do enable us to reject the simplest version of the garden path and Exhaustive computation hypotheses. Both these models predict the same pattern of results for both meanings. The models that predict common meanings should be comprehended more quickly than rare meanings are supported.

It was further assumed three remaining models predicted that repetition of the meaning should facilitate comprehension for both primary and secondary meanings. The results support this prediction although the difference between the SAME and CONTROL conditions does not quite reach significance. The small size of this effect is not particularly surprising. The ambiguous word was only one of four to six words in the sentence and furthermore assigning meaning to the individual lexical items is only one of several stages in comprehension.

The critical comparisons concern the DIFFERENT meaning condition. Let us compare the predictions for each model with the results obtained. The ordered search model predictions. Not one of the three remaining models can account for all the results. The present results give no support to the ordered search model's hypothesis that the primary meaning will always be accessed before the secondary meaning. The results from the secondary meaning are not consistent with the logogen model since it assumes that the primary meaning will never result in priming of the secondary meaning. The results suggest that priming does occur for the ($S_1 \rightarrow S_2$) condition. There was no support for the perceptual suppression theory since there was no evidence that change of meaning made comprehension more difficult than in the control condition. In fact, quite the reverse occurs for the secondary meaning. It must be remembered however that these assumptions about suppression may be over-simple. Mackay (1970) shows that time for suppression depends on the saliency of the suppressed meaning. What has not been established are the consequences for an item of being suppressed, and how long these consequences last. In the present experiment the changed meaning condition requires access to a meaning that has been suppressed in the first sentence. It was assumed that suppression would inhibit access but the important question is how long does an item remain suppressed. The time course of suppression is likely to be closely related to the saliency of the meaning.

It can be assumed that in this experiment there were two conflicting effects. The first assumed effect is that repetition of the same physical stimulus (i.e. repeating the ambiguous word) leads to a faster encoding of the stimulus into some internal representation. The second and conflicting effect is the suppression effect. The effects of suppression can be considered for each meaning relative to the time taken to access that meaning in an unsuppressed state (the CONTROL condition in the present experiment). Accessing the primary meaning causes suppression of the secondary meaning. The <u>relative</u> effect on access time for the secondary meaning will be small since being of low a priori probability it is easily suppressed. Thus an initially slow access time is made only slightly slower. On the other hand accessing the secondary meaning requires suppression of the primary meaning. The primary meaning will require considerable suppression so that an initially fast access time is greatly increased. It must be remembered that the effect is a relative one, so that a suppressed primary meaning may still be more accessible than an unsuppressed secondary meaning. This possibility can be demonstrated as follows: Tp = time to access the primary meaning, Ts = time to access secondary meaning, Tsupp.p = time to access primary meaning when suppressed, Tsupp.s = time to access secondary meaning, Tenc. = time to encode stimulus (stage earlier than accessing meaning).

The whole process of accessing meaning of a stimulus consists of Tenc + Taccess.
Now Tp $<$ Ts but (Tsupp.p-Tp) $>$ (Tsupp.s - Ts). Although this assumption is counter-intuitive it has been made by other psychologists. Campbell, Donaldson and Young (in press) in a different but comparable area of research have made similar assumptions concerning saliency and suppression. Say for example Tp = 8 units, Tsupp.p = 10, Ts = 12, Tsupp.s = 13. However, in the suppressed conditions the repetition of the

stimulus in the present experiment reduces Tenc. by 2 units regardless of meanings then the primary meaning would show no effect of prior use of secondary meaning and the secondary meaning would show that prior use of the primary meaning would lead to a decrease in comprehension time. These results were found in the present experiment.

Thus it can be argued that these results are attributable to conflicting suppression and stimulus repetition effects. This account of the results is certainly not the only one that could be given (e.g. the disambiguation model described by Mackay, 1970, has many similarities to the logogen model but it is quite possible to add assumptions about suppression processes to an ordered-search model). Further the assumptions about the operation of suppression are purely hypothetical at the moment.

This account does, however, point to two possible reasons for the discrepancy between these results and Hogaboam and Perfetti (1975). Hogaboam and Perfetti presented subjects with a sentence indicating one meaning of the final word. Subjects had to decide if the final word had another meaning. They concluded from the experiment that the primary meaning of a word is always accessed first. In the Hogaboam and Perfetti experiment there is nothing comparable to the effects of the stimulus repetition.

To this extent they have avoided the problem of separating the sensory and the semantic elements of the process. (It would be possible using the present task to separate these effects by interposing some task between the presentation of the two sentences so that sensory coding of the stimulus was no longer available, see the Discussion of Exp. 5.). On the other hand the Hogaboam and Perfetti task allows no estimation of the access time of one meaning independent of the other (independent in the sense that the ambiguity of the word is not an intrinsic aspect of the subject's task). It is argued here that such a "base-line" measure is necessary to allow evaluation of the other conditions.

It seems worthwhile to make some comments on the use of the sentence comprehension task in the study of ambiguity. Clark and his coworkers have used this task in a number of studies (not of ambiguity) with apparently little trouble. However, in ambiguity studies the important differences in times are very small and the comprehension task is not exact enough to enable consistent, reliable measurement of these small differences. The main problem is that subjects differ so much in how they interpret the word "understand". It is interesting to compare the present experiment with that of Buhler (1908). Buhler asked famous psychologists to respond as soon as they understood sentences such as "we depreciate everything that can be explained". The mean reaction time was 12 secs. with a range from 5 to 22 secs. The present experiment used simpler sentences and the mean was 1.8 secs with a range from 0.66 to

9.5 secs.

Furthermore there was some suggestion from subjects' reports that a subject might alter his criterion of comprehension from one sentence to the next.

A third disadvantage is that the task does not allow detection of errors. A subject might press the button indicating comprehension when he had in fact interpreted the ambiguous word differently from the meaning indicated by the context. A task permitting detection of these errors might provide another source of information.

In spite of these qualifications this experiment has produced some valuable results. It is interesting to note that the results from the primary meanings are exactly what would be predicted from a simple model assuming access of only one meaning, whereas the results from the secondary meaning are exactly what would be predicted from a model assuming access of all meanings. This supports the argument of Hogaboam and Perfetti (1975), that much of the confusion of earlier investigators arises from their neglect of frequency of meaning as a factor.

This criticism can be levelled at not only those experiments which have been directly concerned with ambiguity (see Exp. 5) but also those experiments which have used ambiguous words as tools in the study of the role of retrieval in recognition and recall (e.g. Light and Carter-Sobell, 1970; Marcel and Steel, 1973).

Conclusions:

The main conclusion of this experiment is that the prevalence of confusion in theories of ambiguity resolution was partly caused by the failure of earlier investigators to systematically control for the effects of frequency of meaning of the ambiguous words they used. The results of the present experiment do not fit readily into any of the existing models of ambiguity resolution. One possible account of the results assumes two conflicting effects, a facilitation effect due to repetition of the ambiguous stimulus and a suppression effect due to change of meaning. The facilitation effect is assumed to be the same for both primary and secondary meanings but the suppression effect is assumed to be relatively more detrimental for the primary meaning than the secondary meaning.

EXPERIMENT 7

The effect of shared features on learning pairs of species names.

Introduction:

It has been suggested in Exps. 4 and 4A, following Clark (1970) that associative priming effects and word associations are both the products of normal language use. It is proposed that both phenomena can be accounted for by a model of lexical storage that assumes that words are represented as bundles of semantic attributes or features. The probability of one word eliciting another as an associate or its effectiveness in priming recognition of that word is a function of the number of attributes that they have in common. In normal language use these will be mainly semantic attributes but acoustic and visual attributes may be involved in certain tasks (see for example Meyer, Schvaneveldt and Ruddy, 1973).

Clark (1970) has described a number of rules governing associations which can be summarized as "change the least possible number of features to produce a new word". Clark assumes that features are ordered and the changes are made according to this order. In the model proposed here the notion of a fixed ordering of features is replaced with the idea that features will vary in their saliency and that the saliency of a feature can to some extent be affected by the context in which they occur.

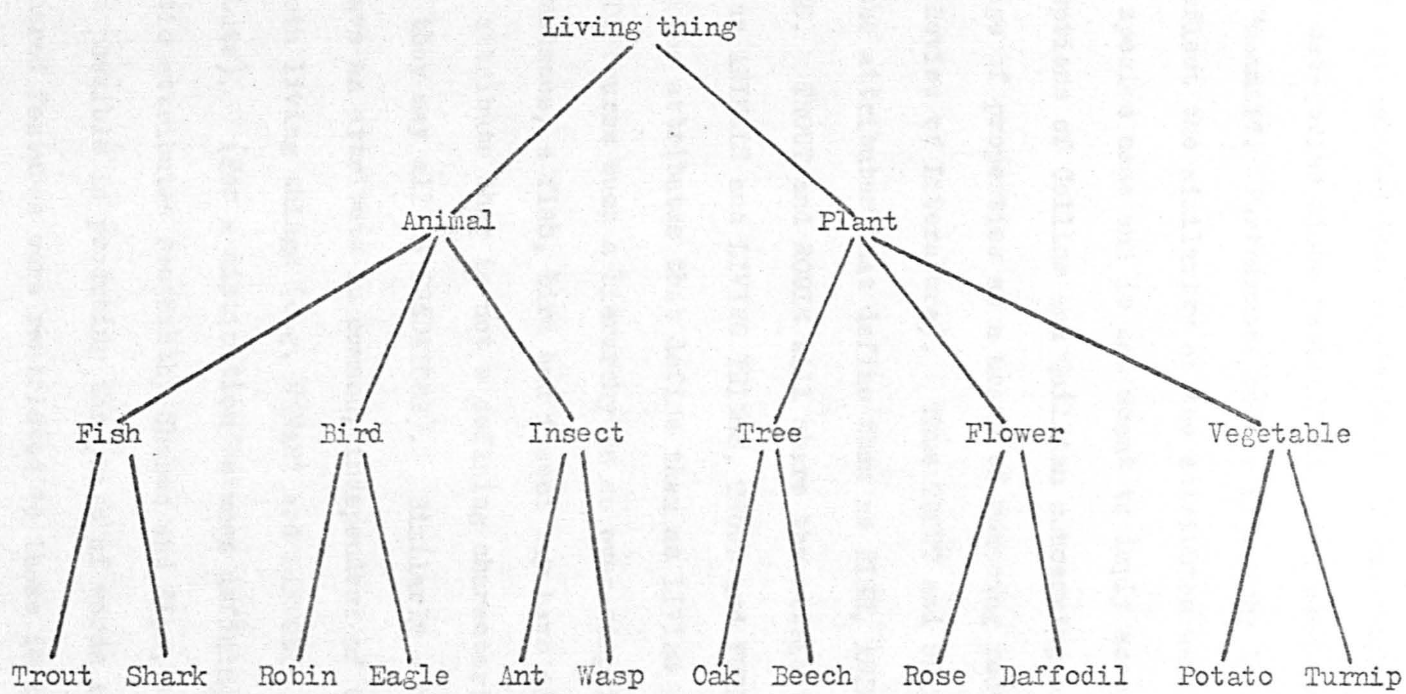
Exps. 4 and 4A gave some support to the hypothesis that the probability of two words being associates or of one priming recognition of the other is related to the number of features that they have in common. Differences were found between pairs of antonyms, synonyms, parallel associates and conceptual associates. These findings were not completely clear-cut since there was some confounding of number of shared features and frequency of occurrence in the language, together with the related problem that a number of the pairs formed idiomatic expressions (e.g. BREAD-BUTTER, LOVE-HATE). Such idioms it was felt may be treated as single lexical items rather than separate words, thus obscuring any effects attributable to shared features alone. (cf. Morton and Broadbent's, 1964 suggestion of the "idiogen" as a higher order "thought unit"). Another shortcoming in Exps. 4 and 4A was that comparison across types of pairs was inevitably confounded with parts of speech. E.g. by their nature antonyms tend to be largely adjectives whereas conceptual associates tend to be nouns).

The present experiment avoids these problems by restricting the materials to nouns and to only one kind of relationship. (items similarity according to membership of different categories and classes). The task required subjects to form associations between pairs of words they were unlikely to have associated into pairs in the past. It was possible to vary the assumed number of shared features independently of past language habits. Habits are assumed to be different from knowledge about words in the sense that BLACK and WHITE are connected through their frequenty occurrence in such phrases as "black and white television", but PURPLE and GREEN could be connected through the knowledge that they are both colours.

The hypothesis tested here is that recall of pairs of words will be better the more features that they have in common. A similar hypothesis was stated by Underwood and Schulz (1960) which they called the associative probability hypothesis. The hypothesis was summarized by Dallett (1964, p 209) as follows: "This hypothesis maintains that subjects generate associations to the items on a verbal list and that the greater the number of such associations the greater the likelihood that one of them will serve as a functional mediator for the pair to be learned." This approach implies that the mediator will be some other word. The approach taken here is that both words will be coded in the form of abstract non-verbal semantic attributes and the more attributes in common the easier will the association be learned.

For this experiment the number of associates two items were assumed to share was derived from a hierarchical structure suggested by the work of Collins and Quillian (1969). See Figure 27.

Figure 27.



Hierarchy of part of the class "Living things".
Based on Collins and Quillian, 1969.

N.B. "Animal" is used here to indicate the quality "animate". It is not assumed that subjects would use "animal" in coding these categories since "animal" is often used as synonymous with "mammal". Furthermore such a hierarchy is only intended to reflect the similarity of the attributes used to represent each species name and is not meant to imply acceptance of the assumptions of Collins and Quillian concerning hierarchical storage of properties as a means of reducing redundancy. (see Review of Literature). Thus TROUT and SHARK will share all the attributes that define them as FISH, ANIMALS and LIVING THINGS. TROUT and ROBIN will share the attributes that define them as ANIMALS and LIVING THINGS, TROUT and TURNIP will only share the attributes that define them as LIVING THINGS.

Of course such a hierarchy is an over-simplification. For instance, a fish, bird and insect may have in common some other attribute that is not a defining characteristic of ANIMAL (e.g. they may all be PREDATORS). Similarly a bird and flower may have an attribute in common, independent of the fact they are both living things (e.g. CANARY and DAFFODIL share a colour attribute). (For a distinction between defining and characteristic attributes see Smith, Shoben and Rips, 1974). As far as possible in producing the pairs of words to be learned the shared features were restricted to those defining attributes that could be derived from the diagram in Figure 1.

It is important to note that the task used involved learning a list of pairs rather than the familiar paired-associate task in that both words had to be remembered since the first word was not provided at recall. We are concerned with the use of abstract semantic representations and there is some suggestion that in the normal paired-associate task subjects tend to learn the list in a "rote" fashion (Jenkins, 1963). It was hoped that having to remember both words would be a more difficult task and that this would encourage the subject to use all his available knowledge to aid recall. For the same reason subjects were instructed to use any strategies to help them remember the pairs.

Method:

Equipment: A memory drum was used which enabled presentation of one pair of words at a time.

Materials: 60 species names were taken from 6 categories in the Battig and Montague (1969) belonging to two classes (ANIMALS and PLANTS). Two lists, each containing 18 pairs, were made up. In one list (referred to as the Animal list) the first word in each pair was an animal name. In the other list (referred to as the Plant list) the first word in each pair was a plant name. Animal names consisted of 6 insects, 6 fish and 6 birds. Plant names consisted of 6 trees, 6 flowers and 6 vegetables.

3 kinds of pairs were made up according to the number of attributes the two species names were assumed to have in common:

1st order pairs: $\frac{1}{3}$ of the names were paired with a species name from the same category (e.g. BIRD with BIRD, FLOWER with FLOWER).

2nd order pairs: $\frac{1}{3}$ of the names were paired with a species name of the same class but not the same category (e.g. FISH with INSECT, FLOWER with VEGETABLE).

3rd order pairs: $\frac{1}{3}$ names were paired with species name from the other class (e.g. INSECT with TREE, VEGETABLE with BIRD).

Each category was represented an equal number of times in each order.

Subjects: 10 1st year undergraduates of Stirling University participated in the Experiment to fulfil a course requirement. None of them had participated in Exps. 4 or 4A.

Procedure: Half the subjects received the animal list and half received the plant list. Subjects were instructed to try to learn the list and to use any knowledge they had of the words in the list to help their recall. Each pair of words were presented for 4 seconds. The pairs were presented in a random order. At the end of the list the subject was instructed to count backwards from 500 in threes for two minutes. After two minutes he was told to write down as many of the pairs of words that he could remember. Three trials (learning list, backward counting and recall) were given. The list was presented in a different random order on each trial. The session lasted 25 minutes.

<u>Results</u>: The effects of three main factors were examined in the analysis of the data, LISTS (plants v. Animals), TRIALS (1st, 2nd 3rd), NUMBER OF SHARED ATTRIBUTES (1st, 2nd, 3rd ORDER). The F values reported were obtained treating SUBJECTS as the random factor. Unfortunately the data was not available to test the effects against MATERIALS as the random factor. The following results were found.

Overall 29% more pairs were recalled from the animal list than the plant list ($F(1,8) = 10.65$ $p < 0.02$). See Figure 28.

There was a significant increase in numbers of pairs recalled over the three trials ($F(2,16) = 72.47$, $p < 0.001$). See Figure 28.

There was a significant interaction between TRIALS and LISTS ($F(2,16) = 10.01$, $p < 0.01$). See Figure 28.

The effect of number of shared attributes was significant. ($F(2,16) = 7.44$, $p < 0.01$). See Figure 29.

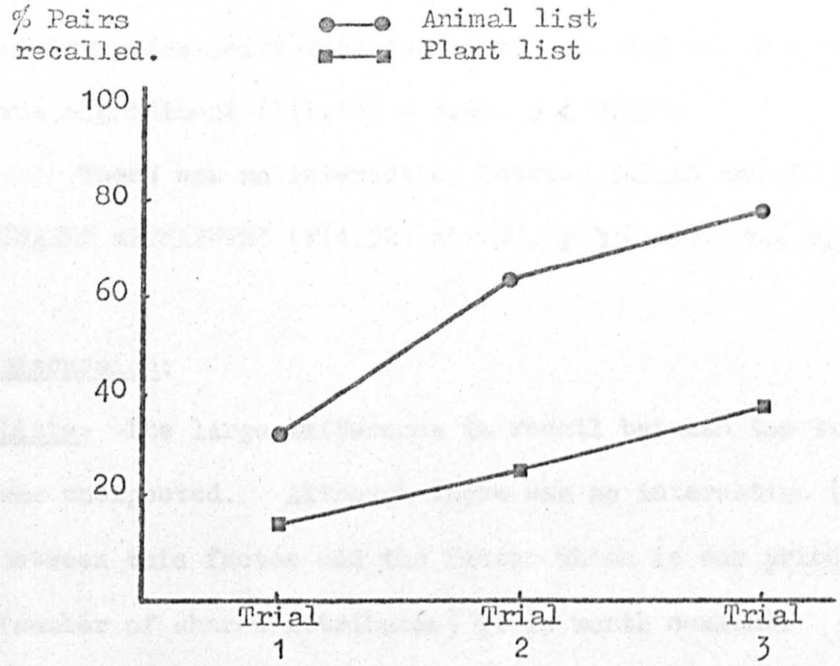
It can be seen from Figure 29 that most of the effect of number of shared attributes is caused by the superior recall of 1st order pairs compared with 2nd and 3rd order pairs. Comparison of 2nd and 3rd order pairs was not significant ([illegible] 3.66, p [illegible] 0.05).

		Mean	S.D.
Trial 1 :	Animal	6.4	3.9
	Plant	2.8	1.3
Trial 2 :	Animal	12.6	3.6
	Plant	5.4	2.7
Trial 3 :	Animal	15.4	3.4
	Plant	7.0	3.7

Mean numbers of pairs recalled (out of 18) and standard deviations for Fig.28.

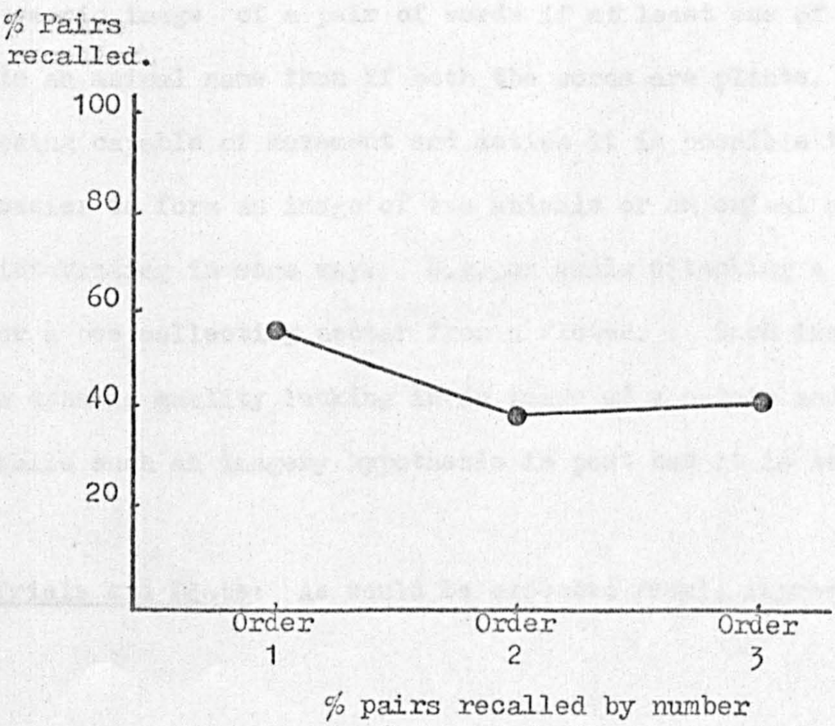
	Mean	S.D.
Order 1	10.3	3.9
Order 2	6.9	5.0
Order 3	7.6	5.2

Mean number of pairs recalled (out of 18) and standard deviations for Fig.29.



% pairs recalled by lists and trials.

Figure 28.



% pairs recalled by number of shared features.

Figure 29.

($F < 1$; see Winer, 1970 pp 65-70 and pp 207-211). Comparison of 1st order pairs with the average of 2nd and 3rd order pairs was significant ($F(1,16) = 5.96$, $p < 0.05$).

There was no interaction between TRIALS and NUMBER of SHARED ATTRIBUTES ($F(4,32) = 1.97$, $p > 0.05$). See Figure 30.

Discussion:

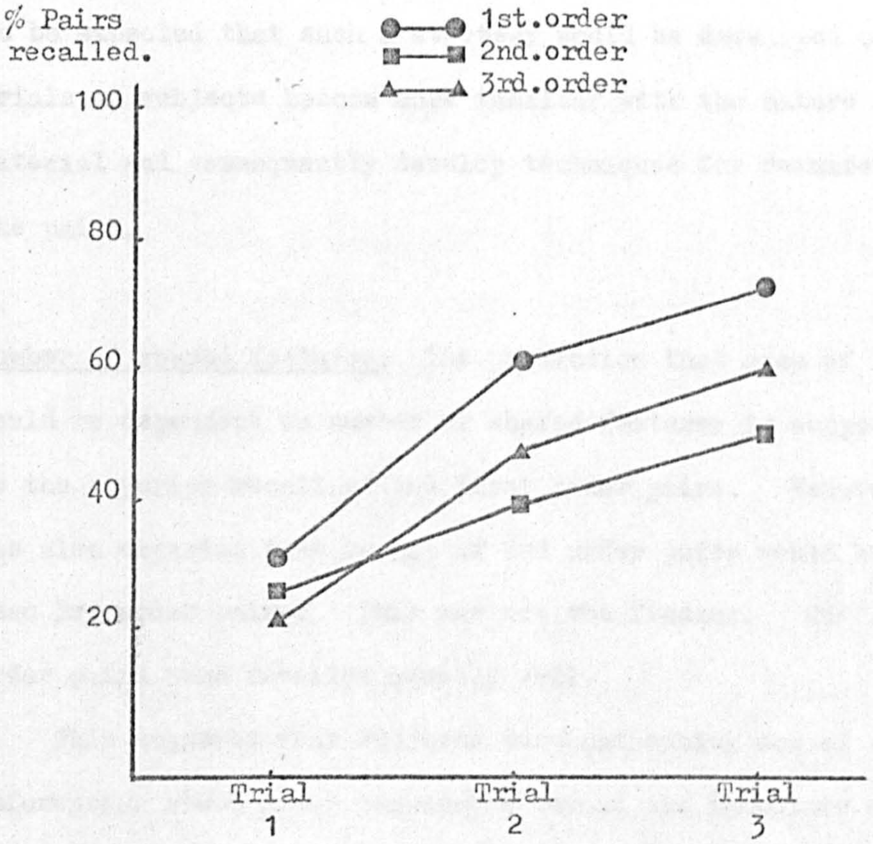
Lists: The large difference in recall between the two lists was unexpected. Although there was no interaction ($F < 1$) between this factor and the factor which is our prime concern (number of shared attributes) it is worth comment. Why should a list consisting largely of animal names be so much better recalled than a list consisting largely of plant names (overall 54% versus 25% correct recall)?

One plausible hypothesis is that it is easier to form a memonic image of a pair of words if at least one of the words is an animal name than if both the words are plants. Animals being capable of movement and action it is possible that it is easier to form an image of two animals or an animal and a plant interacting in some way. E.g. an eagle attacking a sparrow, or a bee collecting nectar from a flower. Such images have a dynamic quality lacking in an image of a potato and a tulip. While such an imagery hypothesis is post hoc it is testable.

Trials and Lists: As would be expected recall improved over

		Mean	S.D.
Trial 1 :	1st order	1.9	1.1
	2nd ''	1.5	1.2
	3rd ''	1.2	1.5
Trial 2 :	1st ''	3.9	1.6
	2nd ''	2.3	2.0
	3rd ''	2.8	1.9
Trial 3 :	1st ''	4.5	1.6
	2nd ''	3.1	2.2
	3rd ''	3.6	2.0

Mean numbers of pairs recalled (out of 6) and standard deviations for Fig.30.



% pairs recalled by Trials and Number of shared features.

Figure 30.

trials. More interesting is the interaction between TRIALS and LISTS. Recall of the animal list is 17% better than the plant list on the 1st trial and this difference more than doubles to 39% on the 2nd trial. It is suggested above that superior recall of the animal list may be due to greater ease in forming "unitary" images of pairs containing an animal name. It is to be expected that such a strategy would be developed over trials as subjects become more familiar with the nature of the material and consequently develop techniques for remembering the pairs.

<u>Number of shared features:</u> The prediction that ease of learning would be dependent on number of shared features is supported by the superior recall of the first order pairs. However, it was also expected that recall of 2nd order pairs would be better than 3rd order pairs. This was not the finding. 2nd and 3rd order pairs were recalled equally well.

This suggests that subjects were not making use of any information about group membership beyond the immediate superordinates. Although it is a fact that pairs such as EAGLE-TROUT share the feature of being animate, and OAK-CABBAGE share the feature of being plants, in the present task at least these objective facts had no subjective value as aids to recall, since such pairs were no better recalled than pairs such as WASP-CARROT and POTATO-WREN which do not share these features.

It seems that when presented with a pair of words a subject categorizes them in terms of their immediate superordinates (eagle → bird, carrot → vegetable) and then codes the pairs category membership as "same" or "different". At recall given retrieval of one member of the pair, knowing that it was paired with a member of the same category is clearly of more value since it restricts the range of possible alternatives far more the retrieval cue that it was paired with a member of a different category.

Why do people not make use of information about membership of higher level classifications? Certainly the subjects were capable of saying what characteristics TROUT and WASP or OAK and POTATO have in common, yet they do not use these features as aids to recall. The answer presumably lies in the nature of the memorization task used. It seems likely that a search is carried out to find a common attribute which can be used to code the pair. Information about membership of directly superordinate groups is quickly found and used as a mnemonic device. The work of Collins and Quillian (1969) suggests that finding higher order similarities will take more time than lower order similarities (e.g. both plants versus both trees). Other strategies may be quicker than searching for high order similarities. Given the limited time available for coding each pair time is likely to be an impòrtant factor. If comparison of immediate superordinates does not produce a match, this strategy may be abandoned in favour of some other strategy.

Conclusions:

The major conclusion of this experiment is that the hypothesis that the main factor in learning pairs of animal and plant names would be number of shared feature is insufficient. Subjects take advantage of shared features that are quickly found but failure to find such features rapidly leads to the adoption of alternative strategies.

An unexpected finding was that pairs containing an animal name were better recalled than pairs containing both plant names. One suggestion is that it is easier to form an image of both names if one is an animal since animals interact with objects in the world around them in a way that plants do not.

EXPERIMENT 8

The effects of priming information on recall from a category.

Introduction:

This experiment is an attempt to investigate what kinds of information people use in comprehending discourse. Since the early work of Bartlett (1932) it has been evident that people do not store what they hear in its original form. More recently Jarvella (1971) has shown that subjects only remember in its exact form the immediately preceding clause in running discourse, when extracting meaning from discourse. Howe (1970) has shown that people remember very few of the actual words in a passage but that they can recall the meaning well. Sachs (1967) found that subjects could detect changes in deep structure but not in surface structure. All these studies indicate that comprehension of sentences involves active recoding processes.

As well as these paraphrasing processes there is also evidence that comprehension involves processes of inference making. Bransford and Franks (1971) have shown that subjects are unable to discriminate between sentences they have seen and related sentences which they have not seen. Furthermore subjects falsely recognize consequences of sentences they have seen when the consequences themselves had not been presented. Kintsch

(1972) has found that subjects use information not actually presented in comprehending sentences. In particular they make extensive use of facts that are implied but not explicitly stated.

All this evidence suggests that comprehension is an active process involving an interaction between stored information and the incoming discourse. It is worth noting, however, that Collins and Quillian (1972) failed to find evidence of the use of such knowledge. They tested the hypothesis that subjects would be quicker to make a true/false decision about a concept if that concept had been implied in the comprehension of a previous sentence. For example, it was hypothesized that comprehending a sentence like "The gloves were in his coat" should facilitate deciding that "A coat has pockets" is true since the first sentence obliges subjects to make use of the information in the second sentence. Decision times were, however, not faster than when the preceding sentence was "The gloves were under his coat", which does not imply the second sentence.

The present experiment was an attempt to find another way of providing evidence about the use of stored information in comprehension. In particular the experiment investigated whether people use their knowledge of individual members of a category in comprehending discourse about the category. Evidence has been presented here showing the close relationship between a word and its superordinate. E.g. in Exp. 1. a category name cue facilitated verification of a proposition about a member of the category. In Exp. 7. being able to categorize two words together facilitated recall of the pair. In the light of this evidence it seems not unreasonable that people would use particular members of a category to aid comprehension and retention of facts about the category. E.g. given the sentence "Some birds feed in water" the subject may code this in terms of particular water birds that he knows. One reason for doing this may be that members of category are more "concrete" than the relatively abstract category name (e.g. table versus furniture). Paivio (1971) has shown the superiority of concrete over abstract words in a number of situations. If people do carry out such recoding operations this may manifest itself in increased recall of members of the category.

Method:

Subjects: 36 subjects were used. Subjects were lecturers and post-graduates of the Psychology Department of the University of Stirling. Relatively "sophisticated" subjects were used, since previous experience had shown that subjects from the 1st year undergraduate subject pool did not perform very well when asked to recall category members verbally.

Materials: Two passages were taken from the Encyclopaedia Brittanica. One passage concerned birds and one trees.

The passages were of a general nature and any reference to specific bird or tree names were excluded. The passages were tape recorded and each lasted approximately $3\frac{1}{2}$ minutes. (see Appendix)

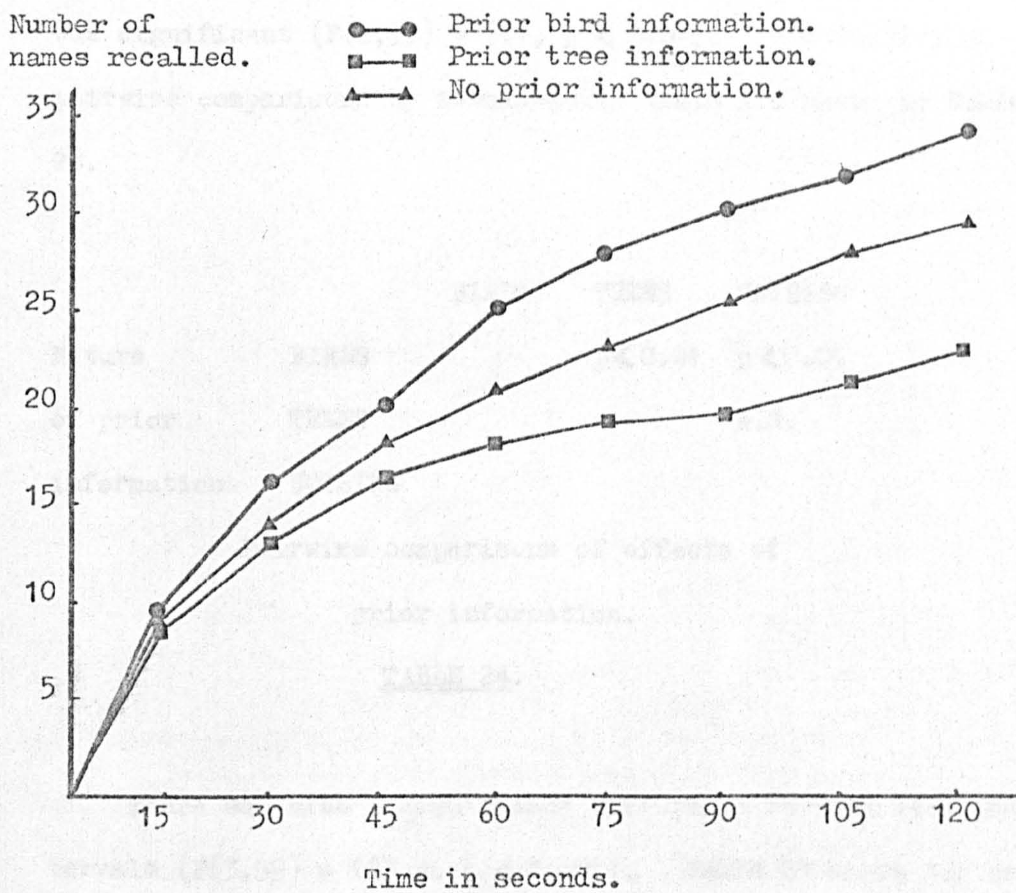
Procedure: The 36 subjects were randomly assigned to 3 groups. Group 1 heard the "bird" passage, group 2 heard the "tree" passage, group 3 heard no passage. Subjects were tested individually. Groups 1 and 2 were instructed to listen to the tape recorded passage and told that at the end of the passage they would be asked to carry out some unspecified task. When the passage ended both groups were asked to recall as many bird names as they were able. Subjects were allowed 2 minutes for recall and their performance was tape recorded. Group 3 subjects were simply asked to recall as many bird names as they could in two minutes.

Results and Discussion:

Figure 31 shows the cumulative total of bird names produced by each group plotted against time. Although this is the usual way of presenting such data Smith and Claxton (1972) have made the following comment: "Mean number of words produced does not appear to be a very sensitive statistic, because it is determined by two factors whose effects interact; the cumulative number of words produced by time t, N_t, is determined by (1) the asymptotic level that N_t gradually approaches, and

Seconds		15	30	45	60	75	90	105	120
Bird.	Mean	10.6	16.5	20.0	24.2	27.8	29.8	31.9	33.7
	S.D.	1.4	1.4	3.5	4.7	6.0	6.6	6.7	7.6
Tree.	Mean	8.9	13.3	15.8	17.8	19.1	20.5	22.1	23.5
	S.D.	1.9	2.5	3.6	4.8	4.7	4.7	5.0	4.8
No.	Mean	9.3	14.4	16.9	20.3	22.7	25.4	27.5	28.5
	S.D.	2.6	3.1	3.8	4.1	5.2	6.0	7.0	7.0

Means and standard deviations for Fig.31.



Cumulative record of number of bird names recalled over time.

Figure 31.

(2) the rate at which this asymptote is approached."

For the present, it is simpler to concentrate on the single factor of rate of production (N'_t). Figure 32 shows the rate of production in each 30 second interval for each group. An analysis of variance showed that the effect of the prior passages was significant ($F(2,33) = 7.1$, $p < 0.01$). The results of pairwise comparisons by Newman-Keuls tests are shown in Table 24.

		BIRDS	TREES	NOTHING
Nature	BIRDS		$p < 0.01$	$p < 0.01$
of prior	TREES			N.S.
information	NOTHING			

Pairwise comparisons of effects of prior information.

TABLE 24.

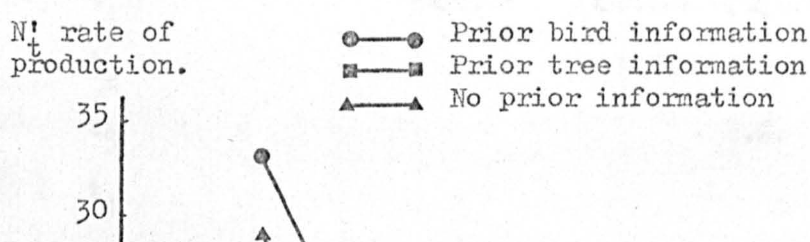
There was also a significant difference between time intervals ($F(3,99) = 163.3$, $p < 0.001$). Table 25 shows the results of pairwise comparisons of intervals by Newman-Keuls.

The interaction between time interval and prior information was not significant ($F < 1$).

An alternative way of plotting these results is that suggested by Smith and Claxton (see above). Instead of plotting N'_t against time interval (t) it is plotted against the total

Seconds	30	60	90	120
Bird. Mean	33.0	15.6	11.2	7.8
S.D.	2.9	8.6	4.6	6.1
Tree. Mean	26.6	9.0	5.6	6.0
S.D.	5.0	5.6	4.4	2.8
No. Mean	28.8	11.8	10.2	6.4
S.D.	6.2	5.0	5.8	5.0

Means and standard deviations for Fig.32.



Rate of production of bird names
in each 30 second interval.

Figure 32.

number of words produced by time t (N_t). The data is plotted in this way in Figure 33.

	30 secs.	60 secs.	90 secs.	120 secs.
t_1		$p < 0.01$	$p < 0.001$	$p < 0.01$
t_2			N.S.	N.S.
t_3				N.S.
t_4				

Pairwise comparisons of time intervals

<u>TABLE 25</u>.

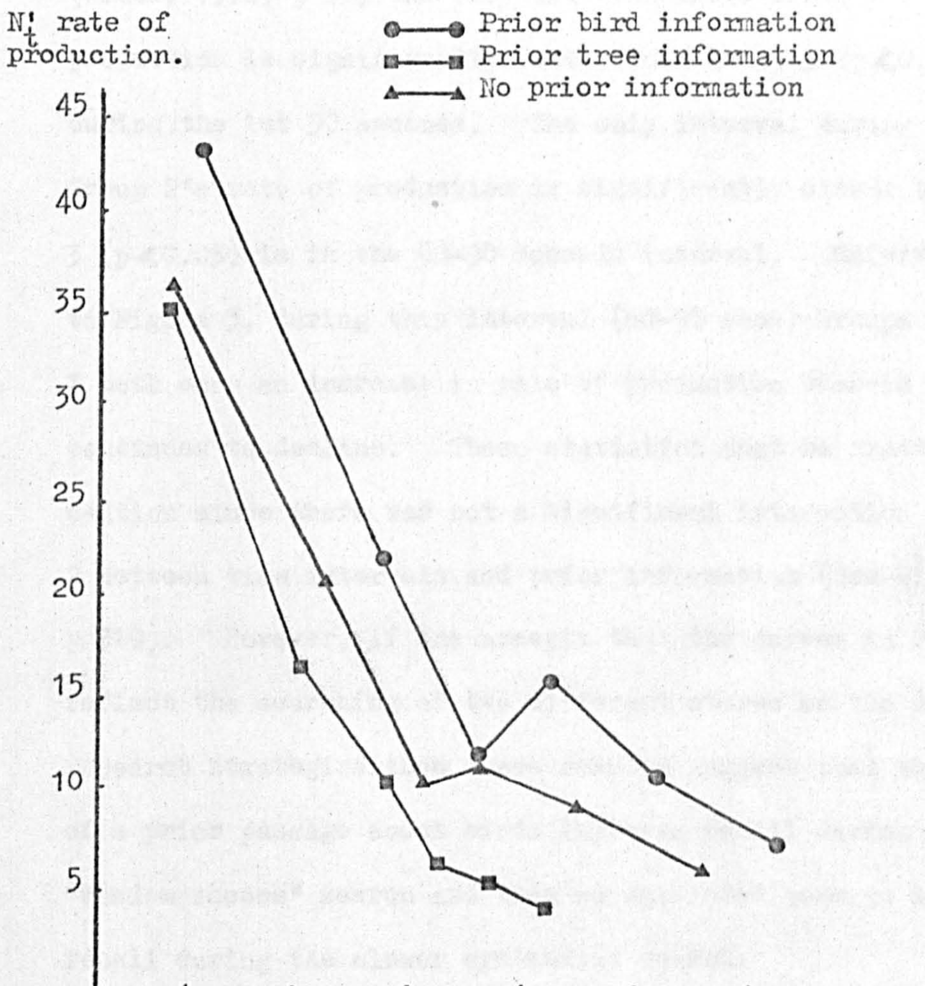
The curves in Figure 33 are similar to those obtained by Smith and Jones (1974). Smith and Jones interpret this shape as being incompatible with the simple random search model proposed by Bousfield (1953) and others. They suggest that there are two processes involved. The initial steep linear segment of the curves is assumed to reflect the search of a small capacity, random access store. When the rate of production becomes too low the subject switches to a slower systematic search of a larger store, which produces the shallow section of the curve.

Examination of Figure 33 reveals that the superiority of Group 1 over Group 3 is greatest in the initial part of the curve. Groups 2 and 3 differ only slightly in this section.

Seconds		15	30	45	60	90	120
Bird.	N'_t	43.6	22.0	12.0	16.4	10.0	7.2
	N_t	5.6	13.7	17.9	21.5	26.0	31.5
Tree.	N'_t	35.6	16.8	10.4	6.4	5.6	4.4
	N_t	4.5	11.0	14.4	16.5	18.7	21.7
No.	N'_t	36.8	20.8	9.6	13.6	11.0	7.0
	N_t	4.6	11.7	15.6	18.5	22.9	27.5

Means for Fig.33.

(These figures are based on group data, so standard deviations cannot be calculated).



N'_t plotted against N_t.

Figure 33.

However, in the shallow part of the curve rate of production for groups 1 and 3 are similar (certainly for the last two points), whereas Group 2 is now inferior to Group 3. Some statistical support is given to these claims by Figure 32. Using Dunnett's test for comparing all means with a control, (Winer, 1970, p 89) the only interval where Group 1's rate of production is significantly faster than Group 3 ($p < 0.05$) is during the 1st 30 seconds. The only interval during which Group 2's rate of production is significantly slower than Group 3 ($p < 0.05$) is in the 60-90 seconds interval. Referring back to Figure 3, during this interval (60-90 secs) Groups 1 and 3 both show an increase in rate of production whereas Group 2 continues to decline. These statistics must be treated with caution since there was not a significant interaction in Figure 2 between time intervals and prior information (See Winer, 1970, p 310). However, if one accepts that the curves in Figure 3 reflect the searching of two different stores or two different research strategies then these results suggest that the effect of a prior passage about birds improves recall during the first "random access" search and that an unrelated passage impairs recall during the slower systematic search.

In an attempt to gain a clearer understanding of the processes underlying these effects five protocols from each group were selected at random for closer examination. Using a technique suggested by Smith and Claxton (1972) each protocol was given a "relatedness" score. This score consists of rating each name on a 4 point scale for the degree of relatedness to the preceding word. E.g. rook - raven = 3, rook - eagle = 2, rook - blackbird = 1, rook - budgerigar = 0. Since such a score is to some extent subjective an independent rater was asked to score the protocols. Pearson product moment correlation between the author's scores and the independent rater's scores was +0.61 $p < 0.02$. Although the agreement was not 100% it was felt that the agreement was high enough to justify the use of the score. It was felt that this "1st order* measure did not exhaust the degree of relatedness in the protocols and so a "2nd order" measure was derived using the same technique but comparing each name with the name before the immediately preceding one. These scores are presented as mean relatedness per word produced. 1st and 2nd order relatedness were fairly highly correlated ($r_s = +0.6$ $p < 0.02$). A combined relatedness score was obtained by adding the 1st and 2nd order scores. Table 26 shows the correlation between 1st order, 2nd order and combined relatedness scores with the total number of names produced.

1st order	2nd order	Combined
+0.7**	+0.68**	+0.73**

Spearman Rank Correlations with Total number of names produced. ** = $p < 0.01$.

TABLE 26.

Bousfield, Mandler, Tulving, Bower and others have all shown that in list learning recall is closely related to the organization that the subject can impose on the list. The present results demonstrate that such organization is also important in producing members of a category.

It has been suggested here that recall can be divided into two stages, the first random and the second systematic. The comparisons in Table 25 provide some support for such a distinction since they show that rate or production in the 1st 30 seconds differs significantly from rate in the other intervals and that rates in these three intervals do not differ from each other. As well as the rate one would expect differences in the importance of organization in the two stages. For each subject in the sample a relatedness measure was computed for the 1st half of the names produced and for the 2nd half. A 1st order, 2nd order and combined relatedness score were computed and correlated with total number of names produced. The correlations are shown in Table 28.

	1st order	2nd order	Combined
1st ½	0.39	0.23	0.38
2nd ½	0.5*	0.74**	0.72**

Spearman Ranks correlations with Total Number produced. * = $p < 0.05$ ** = $p < 0.01$.

<u>TABLE 27</u>.

It can be seen that degree of relatedness in the later stages of recall is more closely related to total produced than relatedness in the first stage. This is particularly the case for 2nd order relatedness. The difference between the correlations for the 1st order and each half is not significant but the difference is significant for the 2nd order ($t = 2.37$, d.f. = 12, $p < 0.05$).

However, the main concern of the present experiment is the effect of the prior passages and the question here is, do the prior passages affect degree of organization? Figure 34 shows the mean combined relatedness score for each group. (1st and 2nd order relatedness both followsed similar patterns).

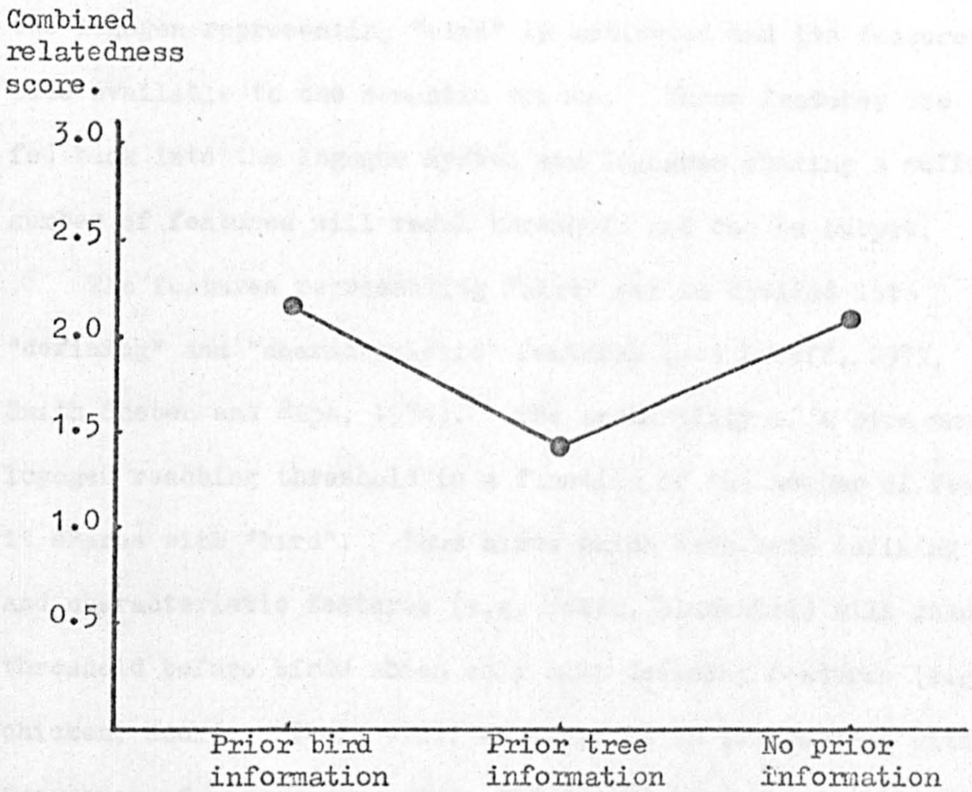
There is an overall significant difference between groups (Kruskall-Wallis $H=7.6$, $p < 0.05$). Groups 1 and 3 do not differ but both groups are significantly different from Group 2 (Mann Whitney $U=0.5$, $p < 0.008$; $u=3$, $p=0.028$ respectively).

It thus seems that the poor performance of Group 2 can be explained by a failure to organize their output. What causes this failure of organization is not clear. At present no suggestion is offered, although Smith and Claxton's (1972) proposal that relatedness is a measure of spare mental capacity may indicate a direction for future research.

Although we can account for the poor performance of Group 2 in terms of failure to organize output some other explanation is needed to explain the superior performance of Group 1 compared

	Mean	S.D.
Bird.	2.17	0.28
Tree.	1.48	0.27
No.	2.06	0.39

Means and standard deviations for Fig.34.



Mean combined relatedness score for each condition.

Figure 34.

to Group 3. To account for this finding it is necessary to outline a model of the processes involved in the generation task.

It is assumed that all bird names are stored in the logogen system, discussed elsewhere, in the form of a list of semantic features. When the instruction is given to recall bird names the logogen representing "bird" is activated and its features made available to the semantic system. These features are fed back into the logogen system and logogens sharing a sufficient number of features will reach threshold and can be output.

The features representing "bird" can be divided into a "defining" and "characteristic" features (see Lakoff, 1972, Smith Shoben and Rips, 1974). The probability of a bird name's logogen reaching threshold is a function of the number of features it shares with "bird". Thus birds which have both defining and characteristic features (e.g. robin, blackbird) will reach threshold before birds which only have defining features (e.g. chicken, duck). There will, however, be an interaction with frequency of occurrence, since there is a bias for high frequency names to reach threshold before low frequency names.

To summarize, it is suggested that order of production of members of an instance will be a function of two factors; what might be called how "typical" an instance is and how frequently it occurs in the language. Some support for these assumptions can be obtained from the Battig and Montague (1969) category norms. Smith et al. (1974) present a typicality rating of 12 bird names. These ratings have a rank order correlation of +0.89 ($p < 0.01$) with frequency of occurrence in the Battig and Montague norms, i.e. the more typical a bird is, the more likely it is to be produced as an instance of bird. Thorndike-Lorge frequency of occurrence and number of times a bird name appeared first in the Battig and Montague norms correlate +0.78 ($N = 10$, $p < 0.01$).

It is suggested here that this model can account for the two segments of the curves in Figure 33 as follows: in the normal "unprimed" situation the effect of feeding the features of "bird" into the logogen system is to cause the logogens representing bird names that are both highly typical (i.e. share the characteristic feature of "bird") and highly frequent to reach threshold and be available for output. (Of course as each name reaches threshold it supplies a list of features that can be used as new input to the logogen system). Once these names have been output conscious strategies are needed to cause names that are either of low typicality or low frequency to reach threshold. Such strategies may involve activating low saliency attributes of "bird". Attributes below a certain saliency may not be automatically transmitted to the semantic system at the initial activation of "bird". There may be a limited information capacity. The first part of the curve simply involves reading out names automatically made available by "bird" and subsequent output of highly typical birds, the second part involves active retrieval of features and this process is slower.

We can now return to the effect of the prior bird passage and to the question asked at the beginning of the experiment; what kind of information is used in comprehending discourse. Examination of the passage about birds (see Appendix) suggests that its effects may be to prime a number of kinds of birds that do not normally appear in the early part of recall (see Battig and Montague, 1969; Brown, 1972). Such sub-groups of birds (e.g. edible birds, water birds, flightless birds etc.) tend not to share the characteristic features of "bird". However, having been primed by the passage, when the subject is asked to recall bird names these names automatically reach threshold and can be read out during the initial strategy-free stage of recall. This can be contrasted with the group who heard no passage, who will only be able to recall such birds with the aid of strategies (i.e. in the second stage of recall).

Examination of the protocols using the ratings provided by Smith, Shoben and Rips, (1974) as a guide to typicality showed that overall there was no significant difference between groups in the percentage of non-typical birds recalled. Comparing only the first 10 names produced, however, showed that there was a significant difference between groups. Table 28 shows the percentage of non-typical bird names produced by each group. $X^2 = 6.2$, d.f. = 2 $p < 0.05$.

Prior bird passage	Prior tree passage	No prior passage
20	4	8

Percentage of non-typical birds produced in 1st 10 names.

TABLE 28.

Conclusions:

It was assumed in this experiment that recall of members of a given category is a function of two variables; how "typical" an instance is and how frequently it occurs in the language. It is further assumed that in normal recall there are two stages; an automatic retrieval of high typicality and high frequency instances, followed by a slower strategy dependent retrieval of low typicality or low frequency instances. The second stage is highly dependent on strategies that permit organization of the output. In this experiment it was found that these strategies were less effective if the subjects were unexpectedly asked to recall members of the category after listening to an unrelated passage.

It was found that a relevant passage improved recall by enabling the subject to gain faster access to low typicality but high frequency instances. This finding is consistent with a logogen model of word production and supports the view that comprehension of language involves activation of concepts in memory.

Final Discussion and Conclusions

I shall first briefly recapitulate the major findings of the experiments reported here and the implications of these findings for the version of the logogen model presented earlier. The model assumes that each word is represented by a recognition device that examines input from visual, acoustic and contextual sources for features belonging to its word.

In Experiments 1 to 3 a variety of associative priming effects were demonstrated. Priming effects were found using superordinates, subordinates and 'simple associates' as cues. (N.B. Rosch (1975) has recently produced evidence supporting the results concerning superordinates, using a task similar to Posner's (1973)). The results of these experiments are consistent with the logogen model. The finding of priming effects using a pronunciation task (Exp. 3) is consistent with the assumption of the logogen model that the function of contextual information is to facilitate the decoding of a visual representation of a word into a semantic representation rather than facilitating any search through memory. This confirms the finding of Meyer et al. (1974).

Experiment 4 shows that associative priming effects are complex and depend upon the frequency of occurrence of the words used and the time between the presentation of the cue and the associate. These findings are consistent with the general workings of the logogen model and further they enable us to specify in greater detail the way in which context is used in the logogen system. It seems likely that certain highly salient information (e.g. antonyms, membership of idioms etc.) is fed back very rapidly into the logogen system from the semantic system. The speed with which such information is made available (less than $\frac{1}{3}$ second) suggests that this process occurs automatically. Miller (1962) has argued that the speed with which decisions have to be made in word recognition is more compatible with a passive (i.e. automatic) system than an active system. Certainly as is argued by Laberge and Samuels (1974, see Introduction) the more the 'lower order' processes can be automated in language recognition the greater the mental capacity 'left free' for handling processes that cannot be automated. This relationship between the logogen system and the semantic system is discussed in greater detail later.

Experiment 4A suggests that common mechanisms are involved in associative priming effects and production of word associations. Again word frequency and kind of association were found to be important factors. This is consistent with the assumption made above that certain kinds of information about a word are made available automatically but that other kinds are only made available by the operations of the semantic system. The logogen model predicts that frequency will be an important variable but as the model is stated by Morton (1970) there is no reason to believe that one kind of information will be more important than any other kind. On the other hand Clark's (1970) rules for describing word association production are based on the assumption that a word's features will be ordered in some way (see also Katz and Fodor, 1963). Exps. 4 and 4A give some support to Clark's rules but demonstrate they are inadequate without taking frequency into consideration. It may be possible to redescribe Clark's rules in terms of a logogen system where ordering of features is equivalent to the order in which features are made available to the semantic system by the logogen system. Thus features will vary on some 'saliency' dimension. Saliency may be determined by how often a particular feature has been relevant to a word's comprehension in the past. This assumption is similar to Wilkins' (1971) notion of conjoint frequency.

It was argued that one of the primary functions of context in a system such as the logogen system would be to resolve ambiguity due to homonyms by causing the most likely meaning to reach threshold first. Experiments 5 and 6 do not provide unequivocal support for the logogen model or for any of the other models proposed. The perceptual suppression theory of Mackay, which is similar in many respects to the logogen model, can account for most of the results but only by making some post hoc assumptions. In general the position of theories of ambiguity resolution is unsatisfactory. Many of the reported findings are themselves ambiguous or have failed to replicate. (see Hogaboam and Perfetti, 1975). It seems possible that a logogen model incorporating some kind of suppression mechanism may account for most of the available data but at present there is no clear experimental evidence of how such a suppression mechanism might work. It is interesting to note that a recent theory of semantic memory not specifically concerned with ambiguity (Collins and Loftus, in press) also includes an assumption about suppression as follows: "If the total amount of activation is limited then the activation of one concept by another closely related concept may make a third, distant concept temporarily less accessible."

On a related note it should be pointed out that dividing meanings into primary and secondary may be an over-simplification. Homonyms vary from those with one clearly dominant meaning to those with both meanings approximately equiprobable. Future research should follow Mackay's example and include degree of bias as a factor in experiments.

Experiment 7 showed that subjects can use knowledge of immediate category membership in a learning task but do not use knowledge of higher order category membership. This again supports the assumption that information from the logogen enables retrieval of key information about a word very quickly (automatically?) but that other information may only be obtained as a result of operations in the semantic system. Whether infor-

Whether information about category membership is stored as a link between the instance and the category name (i.e. as a feature in its own right) or is obtained by comparing the feature lists of the instance and the category name will be discussed in more detail later.

Experiment 8 demonstrated how sensitive recall is to context. The most relevant conclusion from the point of view of the logogen model was that recall from a category can be divided into an automatic read-out stage and a strategy dependent stage. (cf. the assumption of retrieval of information about a word made above). It was also concluded that a passage about birds primed less salient features of 'bird' causing less typical members of the category to be accessed more easily.

Much of the discussion of these results has been in terms of attributes and features, yet in the Introduction it was argued that it was implausible to restrict a word's meaning to a finite set of features. Here I want to describe how a logogen model could be conceived that includes aspects of both the feature models and the association network models of semantic memory. The following account makes use of suggestions from a variety of sources, but notably Collins and Quillian (1972) and Kintsch (1970).

I want here to use the analogy of distinguishing a dictionary from an encyclopaedia. The logogen system can be regarded as the dictionary and is solely concerned with words and their definitions. This definition may include a visual imagery component. The semantic system is the encyclopaedia and is concerned with general knowledge of the world. It is neither verbal nor sensory but abstract in nature. The definition of word in the logogen system is in terms of features. Knowledge of the world in the semantic system is in the form of a network. The essential thing about a dictionary entry is that it is not exhaustive; it merely lists the key features of a word. These are both syntactic and semantic. What is the nature of these features? In Quillian's terminology they may be "token nodes" belonging to "type nodes" in the semantic system. In Kintsch's terminology they may be pointers that give the addresses of other words. Here it is preferred to regard them as pointers to abstract concepts in the semantic system. Some of the concepts may be more readily realized in verbal terms and others in some kind of sensory/imagery form. Features do not merely point to entries in semantic memory but indicate a type of relationship.

The operation of the logogen system is independent of strategies. It simply receives input and provides output. The output to the semantic system may take the form of activating a concept in semantic memory plus an indication of the relation of the word to the activated concept. The semantic system, on the other hand, is richly supplied with strategies that can operate on the activated concepts to extract further information about the word. These operations will themselves activate other concepts. Information about activated concepts will be fed back to the logogen system, whatever the source of the activation.

It is assumed that activation of the key concepts in semantic memory that define a word will occur automatically, i.e. without any need for attention. Attention is necessary for the selection and operation of strategies in the semantic system.

In reading or listening to speech in a normal everyday manner it is assumed that the concepts automatically activated by the logogen output are sufficient for comprehension. If the situation requires more information than is provided by these concepts then strategies may be used to obtain further information.

It has been suggested here that a word's logogen provides access to only the key information needed for its definition. Kintsch says that an entry will be encoded by whatever semantic markers are "relevant". Can we specify what features will be relevant for a given word? Although the idea that all words can be described by a limited set of features has been rejected it seems that it may be possible to identify certain dimensions within restricted semantic domains.

Rips, Shoben and Smith (1973) have shown that mammals and birds can be distinguished on the dimensions of size and predacity. Miller (1972) has attempted to define the dimensions underlying verbs of motion. In more general areas the results of Exps. 4 and 4A suggest that where a word has an opposite (or complement) this opposite will play a key role in its definition. This is especially the case for many common adjectives that can only be defined in terms of the dimension they define (e.g. GOOD-BAD, LARGE-SMALL) although it can also apply to verbs (e.g. GIVE-TAKE) and nouns (KING-QUEEN). Similarly the results from the parallel associates suggest that a word will be defined in terms of phrases in which it commonly occurs (e.g. BREAD-BUTTER, HORSE-CART, EGG-BACON). It should be noted that some of the antonym pairs may also fall in this class. Synonyms do not appear to be particularly key concepts in a word's definition. This conclusion is suggested by Exps. 4 and 4A and the generally low frequency of synonyms as associates in association norms. (Examination of the norms in Deese (1965) suggests that this may only be true for the more concrete nouns and adjectives). It may be that to obtain a synonym of a word requires operations by the semantic system on the features provided by the word's logogen.

Perhaps the most vexed question is the role of superordinates in the definition of a word. As noted earlier a number of models (e.g. Kintsch, 1970; Rumelhart et al., 1972) explicitly include links between a word and its superordinates. Indeed we have produced experimental evidence of a close relationship between a word and its superordinate (see in particular Exps. 1 and 7). However, these results are equally compatible with the feature comparison models proposed by Smith et al. (1974), Schaeffer and Wallace (1970), and others. As noted in the Introduction Smith et al's. version of the feature model, incorporating a distinction between defining and characteristic features can account for a number of findings that are difficult for models which assume an ISA link between a word and its superordinate. Out of the present experiments Exp. 8 is the only one which provides any evidence relevant to this argument. The findings of Exp. 8 can be more parsimoniously accounted for by a feature model.

In general then the evidence is more consistent with the view that categorization is the result of an operation carried out by the semantic system rather than a direct link between an instance and its category.

On the other hand it may perhaps be useful to emphasize the distinction (after Kintsch) between features that carry essentially linguistic information and those that carry imagery information. It is quite probable that theorists who argue for feature comparison are dealing largely with imagery and that theorists who argue for an ISA relationship are referring to knowledge that is verbal in origin. Most of the materials used in category experiments are familiar and more important highly concrete (i.e. high in imagery content). E.g. Rips et al. (1973) use common bird and animal names; Rosch (1975) uses nine categories explicitly chosen for their highly concrete nature. Paivio (1971) has shown the pervasiveness of imagery as a variable in a large variety of situations, which suggests that imagery, where available, is extensively used. It is probable that for highly concrete words imagery information is made available faster than semantic information. (For experimental support of this claim see Rosch, 1975). When the system is asked to verify an instance as a member of a category if the instance is concrete the first thing that will be made available will be an image. This can then be compared with some stereotyped image of the category. At the same time, although at a slower rate the semantic information will be made available. This may be needed to confirm or reject the imagery-based feature comparison. Such a check would be needed to avoid falsely affirming such sentences as "A whale is a fish" where imagery feature comparison is likely to report a match. Whether a decision will be made on the feature comparison alone will depend on the criterion that has been set. E.g. the linguistic hedges described by Lakoff (1972) may produce different criterion levels. "A whale is a true fish" may set a high criterion level, whereas "loosely speaking a whale is a fish" may set a low criterion. It is quite probable that imagery also plays a large role in the verification of physical properties. Seymour (1973) has given some indication of how non-verbal features may be handled by the logogen model.

For abstract words (i.e. those having no or few imagery features) category membership verification must be dependent on language based knowledge. It is difficult to see what role imagery could play in verifying a sentence such as "Christopher Wren was an architect". Since, however, most of the experiments on categorization have concentrated largely on concrete material the representation of abstract words is at present a matter of speculation.

Throughout the discussion of this research frequent use has been made of the concept of activation both in the activation of concepts in semantic memory and in the activation of feature detectors in the logogen system. It has been assumed that activation spreads from the logogen system to the semantic system, within the semantic system from concept to concept, and from the semantic system to the logogen system, but not however, from logogen to logogen. The last assumption is made explicitly by Morton (1970, p 247): "It should be emphasized that the logogen system can in no way be regarded as an associative net. There is no direct way of transferring information from one logogen to another. All associative phenomena that involve any semantic relationships are seen at the moment as proceeding via the Cognitive System."

Since spreading activation was first suggested by Quillian (1966) as a process in memory it has been extensively used as an explanatory device (e.g. Collins and Quillian, 1972; Meyer, Schvaneveldt and Ruddy, 1972; Marcel and Forrin, 1974). Recently Loftus has accounted for a number of findings by Rosch (1975) in terms of spreading activation. Collins and Loftus (in press) have recently extended the theory by adding additional processing assumptions (e.g. the assumption of a corresponding suppression effect). Loftus (1975, p 236) argues that these assumptions "enable the theory to account for widely differing empirical results". Rosch (1975) describes the theory as "interesting and powerful" and agrees with Loftus that it can "encompass virtually all the present data in semantic memory" (p. 243), although this in itself raises the problem of whether the theory is capable of being invalidated. The major problem is that, as yet, spreading activation is essentially a post hoc explanation. The ultimate goal must be to specify the theory with sufficient precision in advance so that it can be put to the test. It is not clear at the moment what assumptions will have to be incorporated into the theory. It seems probable that for some time to come we shall have to keep adding assumptions like those of Collins and Loftus to fit new data as they arise. As can be seen from Exp. 4 even apparently simple questions like "what is the time course of activation?" will require very complex answers. At present this author is basically in agreement with the comment of Rosch (1975, p 243): "Whilst post hoc prediction is a considerable virtue given the present state of psychological theory, such a relationship between a theory and research data does not invoke confidence that this is the only possible explanation of the data."

Finally what answer can we give to the question "What is the role of context in retrieving information from semantic memory?" At the level of the individual word context operates to bias the recognition system towards the most probable stimulus. The result of this biassing is to reduce to a minimum the amount of sensory analysis of the stimulus required. Much of the effect of context occurs automatically as a result of interaction between the word recognition system and stored knowledge about the words. Further contextual information may be obtained as a result of conscious strategies. Different types of words will provide different types of information to the context system and at different speeds depending upon whether the information is made available automatically or not. This will result in biassing of different word recognition units. A similar process is assumed to operate in the production system.

In this research we have largely concentrated upon semantic context. This type of context undoubtedly interacts with syntactic and non-linguistic context. It is possible that these sources of contextual information may be accounted for in terms of a model similar to the one described here for semantic context. This, however, is a problem for future research.

REFERENCES

Anderson, J.R. and Bower, G.H. (1974), Human Associative Memory, John Wiley and Sons.

Atkinson, R.C. and Schiffrin, R.M. (1968), Human memory: a proposed system and its control processes. In Spence, K.W. and Spence, J.T. (Eds.). The Psychology of Learning and Motivation, Vol. 2. New York, Academic Press.

Baddeley, A.D. (1968), How does acoustic similarity influence short-term memory? QJEP 20, 249-64.

Baddeley, A.D. (1972), Retrieval rules and semantic coding in short-term memory. Psychological Bulletin, 78, 379-85.

Baron, J. (1973), Phonemic stage not necessary for reading. QJEP 25, 241-246.

Bartlett, F.C. (1932), Remembering, Cambridge, Cambridge University Press.

Battig, W.F. and Montague, W.E. (1969), Category norms for verbal items in 56 categories: a replication and extension of the Conneticut Category Norms. JEP Monograph, 80, 1-46.

Bertelson, P. (1961), Sequential redundancy and speed in a serial two-choice responding task. QJEP, 12, 90-102.

Bloomfield, L. (1933), Language, New York: Holt, Rinehart and Winston.

Bousfield, W.A. (1953), The occurrence of clustering in recall of randomly arranged associates. Journal of General Psychology, 49, 269-73.

Bouveresse, J. (1974) Discussion with Parret, H. in Parret, H., Discussing Language, Mouton.

Bower, T.G. (1970), Reading by eye. In H. Levin and J.P. Williams (Eds.). Basic Studies in Reading, New York: Basic Books.

Bradshaw, J. (1975), Are we reading right? New Scientist, 65, 935, 321-323.

Bransford, J.D. and Franks, J.J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, 2, 331-50.

Bransford, J.D. and Johnson, M.K. (1973). Consideration of some problems of comprehension. In W.G. Chase (Ed.), Visual Information Processing, New York: Academic Press.

Brenker, J.A. (1973), Priming and availability of knowledge. Paper presented at Joint Conference of E.P.S. and Netherlands Psychonomic Foundation, Amsterdam.

Broadbent, D.E. (1967), Word frequency effect and response bias. Psych. Review, 74, 1.-15.

Brown, W.P. (1972), Studies in word listing: some norms and their reliability. Irish Journal of Psychology, 3, 117-59.

Bühler, K. (1908), The apprehension of thoughts (The comprehension of sentences). In O. Rapaport, (Ed.), Organization and Pathology of Thought. New York: Columbia University Press, 1969.

Campbell, R.N., Donaldson, M.C. and Young, B.M. (in press). Constraints on classificatory skills in young children. British Journal of Psychology.

Carey, P.W. Mehler, J. and Bever, T.G. (1970). When do we compute all the interpretations of an ambiguous sentence? in G. Flores d'Arcais and W. Levelt (Eds.). Advances in Psycholinguistics. Amsterdam: North-Holland.

Chomsky, N. (1957) Syntactic Structures, The Hague: Mouton.

Chomsky, N. (1959). A review of Verbal Behaviour by B.F. Skinner. Language, 35, 26-58.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, Mass.: M.I.T. Press.

Clark, H.H. (1970). Word Associations and Linguistic Theory in J. Lyons (Ed.). New Horizons in Linguistics. Pelican.

Clark, H.H. (1973). The language-as-fixed effect fallacy: A critique of language statistics in psychological research. JVLVB, 12, 335-359½

Collins, A.M. and Quillian, M.R. (1969), Retrieval time from semantic memory. JVLVB, 8, 240-7.

Collins, A.M. and Quillian, M.R. (1970a), Does category size affect categorization time? JVLVB, 9, 432-8.

Collins, A.M. and Quillian, M.R. (1970b). Facilitating retrieval from semantic memory: The effect of repeating part of an inference. In A.F.Sanders (Ed.), Attention and Performance III, Amsterdam: North Holland Publishing Co.

Collins, A.M. and Quillian, M.R. (1972b),How to make a language user. In Tulving, E. and Donaldson, W.A. (Eds.). Organization of Memory, New York, Academic Press.

Collins, A.M. and Loftus, E.F. (in press), A Spreading activation theory of semantic processing. Psych. Review.

Conrad, C. (1972). Cognitive economy in semantic memory. J.E.P., 92, 149-54.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. Memory and Cognition, 2, 130-138.

Corcoran, D.W.J. (1971). Pattern recognition. Penguin.

Craik, F.I.M. and Lockhart, R.S. (1972). Levels of processing: A framework for memory research. JVLVB, 11, 671-84.

Crystal, D. (1971), Linguistics. Pelican.

Dallett, K.M. (1964). Implicit mediation in PA learning. JVLVB, 3, 209-214.

Deese, J. (1965). The Structure of Associations in Language and Thought. Baltimore, John Hopkins Press.

Fillmore, C.J. (1968). The case for case. In E. Bach and R.T. Harms (Eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston.

Fodor, J.A., Garrett, M. and Bever, T.G. (1968). Some syntactic determinants of sentential complexity, II: Verb structure. Perception and Psychophysics, 3, 453-461.

Fordham, F. (1953). An introduction to Jung's psychology. Pelican.

Foss, D.J. (1970). Some effects of ambiguity upon sentence comprehension. JVLVB, 9, 699-706.

Foss D.J. and Jenkins, C.M. (1973). Some effects of context on the comprehension of ambiguous sentences. JVLVB, 12, 577-589.

Foss, D.J., Bever, T.G. and Silver, M. (1968). The comprehension and verification of ambiguous sentences. Perception and Psychophysics, 4, 304-306.

Garrett, M.F. (1970). Does ambiguity complicate the perception of sentences? In G.B. Flores d'Arcais and W.J.M. Levelt (Eds.), Amsterdam: North Holland. Advances in Psycholinguistics,

Gibson, E.J., Pick, A., Osser, H. and Hammond, M. (1962). The role of grapheme-phoneme correspondence in the perception of words. Amer. Journal Psychol. 75, 554-570.

Goodman, K.S. (1967). Reading a psycholinguistic guessing game. Journal of the Reading Specialist, May, 259-271.

Halle, M. and Stevens, K. (1962). Speech recognition: A model and a program for research. IRE Transactions on Information Theory, IT-8, 155-159.

Haviland, S.E. and Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. JVLVB, 13, 512-521.

Heider, E.R. (1973). On the internal structure of perceptual and semantic categories. In T.M. Moore (Ed.), Cognitive Development and acquisition of language. New York: Academic Press.

Herriot, P. (1974). Attributes of Memory. London: Methuen.

Hogaboam, T.W. and Perfetti, C.A. (1975) Lexical ambiguity and sentence comprehension. JVLVB, 14, 265-274.

Howe, M.J.A. (1970). Repeated presentation and recall of meaningul prose. Journal of Educational Psychology, 61, 214-19.

Hubel, D.H. and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. Journal of Physiology, 148, 574-591.

Hubel, D.H. and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 160, 106-154.

Jakobson, R., and Halle, M. (1956). Fundamentals of language. The Hague: Mouton.

Jarvella, R.J. (1971). Syntactic processing of connected speech. JVLVB, 10, 409-416.

Jenkins, J.J. (1963). Mediated Associations: Paradigms and situations. In C.N. Cofer and B.S. Musgrave. Verbal behaviour and learning. New York: McGraw H ill.

Katz, J. and Fodor, J. (1963). The structure of a semantic theory. Language 39, 170-210.

Katz, J. and Postal, P. (1964). An integrated theory of linguistic descriptions. Research Monographs, 26. Cambridge, Mass.: M.I.T. Press.

Katz. J. (1972). Semantic Theory, Harper and Row, New York.

Kausler, D.H. and Kollasch, S.F. (1970). Word Associations to homographs. JVLVB, 9, 444-9.

Kintsch, W. (1970). Models for free recall and recognition. In Norman, D.A. (Ed.), Models of Human Memory. New York, Academic Press.

Kintsch, W. (1972). Notes on the structure of semantic memory. In Tulving, E. and Donaldson, W. (Eds.), Organization and Memory. New York, Academic Press.

Kiss, G. (1972). LTM: A state-space approach. Brit. Journal of Psychology, 63, 3, 327-41.

Kolers, P.A. (1970). Three stages in reading. In H. Levin and J.P. Williams (Eds.), Basic Studies on reading. New York: Basic Books.

Kosslyn, S.M. and Nelson, K.E. (1972). Quoted in Anderson, J.R. and Bower, G.H, Human Associative Memory. John Wiley and Sons (1974).

Laberge, D.L. and Samuels, S.J. (1974). Towards a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-332.

Lackner, J.R. and Garrett, M. (1973). Resolving ambiguity: Effects of biassing context in the unattended ear. Cognition, 1-4, 359-373.

Lakoff, G. (1972). Hedges: A study in meaning criteria and the logic of fuzzy concepts. Papers from the eighth regional meeting, Chicago: University of Chicago, Linguistics Department.

Landauer, T.K. and Freedman, J.L. (1968). Information retrieval from long-term memory: Category size and recognition time. JVLVB, 7, 291-295.

Landauer, T.K. and Meyer, D.E. (1972). Category size and semantic memory retrieval. JVLVB, 11, 539-549.

Lashley, K.S. (1951). The problem of serial order in behaviour. In L.A, Jeffress (Ed.), Cerebral mechanisms in behaviour. New York; Wiley.

Leech, G. (1974). Semantics. London: Pelican.

Lettvin, J.Y., Maturana, H.R., McCulloch, W.S., and Pitts, W.H. (1959). What the frog's eye tells the frog's brain. Proc. Inst. Radio Eng., 47, 1940-1951.

Liberman, A.M., Cooper, F.S., Harris, K.S., Macneilage, P.F., and Studdert-Kennedy, M. (1967). Some observations on a model for speech perception. In W. Wathen-Dunn (Ed.). Models for the perception of speech and visual form. Cambridge: MIT Press.

Light, L.L. and Carter-Sobell, L. (1970). Effects of changed context on recognition memory. JVLVB, 9, 1-11.

Loftus, E.F. and Scheff, R.W. (1971). Categorization norms for 50 representative instances. JEP, 91, 355-364.

Loftus, E.F. (1975). Spreading activation within semantic categories: Comments on Rosch's "Cognitive representation of semantic categories." JEP (General), 104, No. 3., 234-240.

Mackay, D.G. (1966). To end ambiguous sentences. Perception and Psychophysics, 1, 426-436.

Mackay, D.G. (1970). Mental diplopia: Towards a model of speech perception at the semantic level. In G.B. Flores d'Arcais and W.G.M. Levelt (Eds.), Advances in psycholinguistics. Amsterdam: North Holland.

Marcel, A.J. and Steel, R.G. (1973). Semantic cueing in recognition and recall. QJEP, 25, 368-77.

Marcel, A.J. and Forrin, B. (1974). Naming latency and the repetition of stimulus categories. JEP, 103, No. 3., 450-460.

Meyer, D.E. (1970). On the representation and retrieval of stored semantic information. Cognitive Psychology, 1, 242-300.

Meyer, D.E. and Ellis, G.B. (1970). Parallel processes in word recognition. Paper presented at the meeting of the Psychonomic Society, San Antonio, Texas.

Meyer, D.E. and Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of dependence between retrieval operations. JEP, 90, 227-234.

Meyer, D.E., Schvaneveldt, R.W. and Ruddy, M.G. (1972). Activation of lexical memory. Paper presented at the meeting of the Psychonomic Society, St. Louis, Missouri.

Meyer, D.E., Schvaneveldt, R.W. and Ruddy, M.G. (1973). Loci of contextual effects on visual word recognition. In P. Rabbit and S. Dornic (Eds.). Attention and Performance, V, London: Academic Press.

Meyer, D.E., Schvaneveldt, R.W. and Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word recognition. Memory and Cognition, Vol. 2, No. 2, 309-321.

Miller, G.A. (1962). Decision units in the perception of speech. IRE Transactions on Information Theory, IT-8, 81-83.

Miller, G.A. (1970). The subjective lexicon. Presidential address, 78th Annual Convention, American Psychological Association.

Miller, G.A. (1972). English verbs of motion: A case study in semantics and lexical memory. In A.W. Melton and E. Martin (Eds.), Coding processes in Human Memory. Washington, D.C., Winston.

Miller, G.A., Heise, G.A. and Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. JEP, 41, 329-335.

Miller, G.A. and Isard, S. (1963). Some perceptual consequences of linguistic rules. JVLVB, 2, 217-228.

Morton, J. (1964). A preliminary functional model for language behaviour. International Audiology, 3, 216-25.

Morton, J. (1969). The interaction of information in word recognition. Psychological Review, 76, 165-178.

Morton, J. (1970). A functional model for memory. In D.A. Norman (Ed.), Models of Human Memory. New York, Academic Press.

Morton, J. and Broadbent, D.E. (1964). Passive versus active models or is your homunculus really necessary? In W. Watham-Dunn (Ed.), Models for the perception of speech and visual form. MIT press.

Mountcastle, V. (1961). Personal communication to J.C.R. Licklider. In W.A. Rosenblith (Ed.), Sensory communication. Cambridge, Mass.: M.I.T. Press.

Norman, D.A. (1968). Toward a theory of memory and attention. Psychological Review, 75, 522-36.

Norman, D.A. (1970). Memory and Attention, New York, Wiley.

Ogden, C.K. and Richards, I.A. (1923). The meaning of meaning. New York: Harcourt Brace Jovanovich Inc.

Oldfield, R.C. (1966). Things, words and the brain. QJEP, 18, 340-353.

Paivio, A. (1971). Imagery and verbal processes. New York, Holt, Rinehart and Winston.

Palermo, D.S. and Jenkins, J.J. (1964). Word Association norms Grade School Through College. Minneapolis: University of Minnesota Press.

Pavlov, I.P. (1927). Conditioned Reflexes, Trans. G.V. Anrep, Oxford University Press.

Perfetti, C.A. Lindsey, R., and Garson, B. (1971). Association and uncertainty: Norms of association to ambiguous words. Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA.

Posner, M.I. (1973). Coordination of internal codes. In W.G. Chase (Ed.), Visual Information Processing. New York: Academic Press.

Quillian, M.R. (1966). Semantic memory. Report AFCRL -66-189, Bolt, Beranek and Newman, Cambridge, Massachusets.

Quillian, M.R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. Behavioural Science, 12, 410-430.

Quillian, M.R. (1969). The teachable language comprehender. Communication of the Associationnfor Computing Machinery, 12, 459-476.

Rips, L.J., Shoben, E.J., and Smith, E.E. (1973). Semantic distance and the verification of semantic relations. JVLVB, 12, 1-20.

Romney, A.K. and D'Andrade, R.G. (1964). Cognitive aspects of English kin terms. In A.K. Romney and R.G. D'Andrade (Eds.), Transcultural studies in cognition. American Anthropologist, 66 (No. 3, part 2.).

Rosch, E. (1975a). Cognitive representations of semantic categories. JEP (General). 104, No. 3, 192-233.

Rosch, E. (1975b). Reply to Loftus, JEP (General), 104, No.3, 241-3.

Rubenstein, H., and Pollack, I. (1963). Word predictability and intelligibility. JVLVB, 2, 147-158.

Rubenstein, H., Garfield, L., and Millikan, J.A. (1970). Homographic entries in the internal lexicon. JVLVB, 9 487-494.

Rubenstein, H., Lewis, S.S., and Rubenstein, M.A. (1971). Evidence for phonemic recoding in visual word recognition. JVLVB, 10, 645-657.

Rumelhart, D.E., Lindsay, P.H. and Norman, D.A. (1972). A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press.

Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 2, 437-42.

Sanford, A.J. and Seymour, P.H.K. (1974). Semantic distance effects in naming superordinates. Memory and Cognition, Vol. 2, No. 4, 714-741.

Savin, H.B. (1973). Meaning and Concepts: A review of J.J. Katz's "Semantic Theory". Cognition, 2/2, 213-238.

Schaeffer, B. and Wallace, R. (1970). The comparison of word meanings. JEP, 86, 144-52.

Shulman, H.G. (1972). Semantic confusion errors in short term memory. JVLVB, 11, 221-7.

Selfridge, O.G, (1958). Pandemonium: A paradigm for learning. In Mechanisation and Thought Processes, H.M.S.O.

Seymour, P.H.K. (1973). A model for reading, naming and comparison. British Journal of Psychology, 64, 1, 35-49.

Smith, E.E., Shoben, E.J. and Rips, L.J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 81, 3, 214-241.

Smith, P.T. and Claxton, G. (1972). Lexical search and organization in memory. Paper read to Experimental Psychology Society, London.

Smith, P.T. and Jones, K.F. (1974). Phonemic organization and search through long-term memory. In P. Rabbitt and S. Dornic, (Eds.), Attention and Performance V. London: Academic Press.

Snodgrass, J.G. and Jarvella, R.T. (1972). Some linguistic determinants of word classification times. Psychonomic Science, 27, 220-2.

Stannerz, R.F. Forbach, G.B., and Headley, D.B. (1971). Decision and search processes in word-nonword classification. JEP, 90, 45050.

Sutherland, N.S. (1959). Stimulus analyzing mechanisms. In Symposium on the mechanization of thought processes. London: H.M.S.O.

Swanson, J.M. and Wickens, D.D. (1970). Preprocessing on the basis of frequency of occurrence. QJEP, 22, 378-383.

Taylor, J.A. (1956). The effect of set for associated words on duration threshold. Perceptual and Motor Skills, 6, 131-4.

Thorndike, E.L. and Lorge, I. (1944). The teacher's word book of 30,000 words. Columbia University.

Treisman, A.M. (1964), Verbal cues, Language and meaning in selective attention. Amer. J. Psychol., 77, 206-219.

Tulving, E., and Gold, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. JEP, 66, 319-327.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), Organization of Memory, New York, Academic Press.

Underwood, B.J. (1974). The role of the association in recognition memory. J.E.P. Monograph, Vol. 102, No. 5, 917-939.

Underwood, B.J. and Schulz, R.W. (1960). Meaningfulness and verbal learning. Philadelphia: Lippincott.

Wallace, A.F.C. and Atkins, J. (1960). The meaning of kinship terms. American Anthropologist, 62, 58-79.

Weinreich, M. (1966). Explorations in semantic theory. In T.A. Sebeok (Ed.), Current trends in linguistics. Vol. 3, The Hague: Mouton.

Wickens, D.D. (1972). Characteristics of word encoding. In A.W. Melton and E. Martin (Eds.), Coding processes in Human Memory. Washington, D.C., Winston.

Wilkens, A.T. (1971). Conjoint frequency, category size and categorization time. JVLVB, 10, 382-385.

Winer, B.J. (1971). Statistical principles in experimental design. New York: McGraw Hill.

APPENDIX TO EXP. 8.

PRIOR BIRD INFORMATION PASSAGE

Birds belong to the class Aves. This class contains over 8700 species. All members of this class have feathers. Since the earliest time birds have helped to satisfy man's material needs and to provide him with recreation. With increasing leisure and education in the 20th century, more people have become interested in their environment. As birds are one of the most attractive features of animate nature, a tremendous amount of writing and reading about birds is done each year and many people watch and study birds as a hobby.

Many birds have tasty flesh and palatable eggs which probably were eaten by most peoples throughout history. But it was only where birds were especially plentiful that they were important as food. Birds are economically important today in a number of parts of the world.

Because of their body structure and their feathery body coverning, birds are the best fliers among animals. Comparing a bird to an aeroplane, a bird's wing is both wing and propellor. The basal part of the wing supplies most of the supporting surface, the wing tip most of the propelling force. The record speed for a bird is 200 m.p.h. although this has been disputed. The record length of a migration is 7200 miles.

To serve their function a bird's feathers must lie smooth and neat. The grooming or preening of the plumage starts as the nestling's feathers are breaking out of the sheaths. The young bird spends a great deal of time combing the feathers with the bill and freeing them from bits of sheath. Similar behaviour continues throughout adult life. To aid its preening a bird often baths. Birds may use water or dust baths. Many birds have such poor night vision that they sit quietly all through the hours of darkness. The habitat in which a bird feeds during the day

/contd. over...

may serve as its resting place at night. A bird on a branch retains its grip without effort, for the bending of the leg automatically tightens certain tendons that strengthens the grip, rather there must be an effort to straighten the leg and loosen the grip. Sociability may be more pronounced at sleeping time than during waking hours.

Food getting has been a major factor in the course of evolution, in shaping birds structures and habits to fit the environment. The habitats used by birds are not so many or so varied as those occupied by other vertebrates: no birds burrow in the ground for their food as do some animals, nor do any birds live in the great depths of the oceans as do some fishes. But birds do feed in the air, in the water and on the land. Availability seems an important factor in determining diet. Within limits a bird eats what is available.

Birds may not need bodies of water but may drink droplets of rain or dew from leaves or grass, or may get what they need from moist foods, without any source of free water.

The behaviour of birds is caused by both instinct and learning. There is no doubt there is a broad range of bird activities, the paths of which are inherited. But just how and where the bird uses these activities may be greatly modified by the individual bird's experience.

PRIOR TREE INFORMATION PASSAGE

A tree is a woody perennial, usually seed-bearing plant 20ft. tall or more at maturity, in which the main stem dominates the lateral branches in growth, either through life to produce a conical or pyramidal outline, or only during early growth, after a few years forking one to several times to produce several ascending, almost equally important branches that collectively form a flat or rounded crown. There are countless graduations in size, form and growth habit among trees.

Trees made famous by historical events, religious beliefs, superstitions or by their sheet beauty, massive size, venerable age or bizarre appearance occur in many parts of the world. The tallest trees are over 360 ft. high. The oldest known trees are nearly 5000 years old!

Growth rate and longevity varies considerably among trees. Some trees grow 12ft. in a single summer. Some trees grow very slowly for the first 10-15 years and then faster after their root systems have become well established. Others grow rapidly in youth but drop to a moderate or slow rate after the first few years. Variations in soil, available water and winter and summer temperatures are among the factors affecting the growth of trees, so that a plant that grows rapidly in one site may progress slowly elsewhere. The total growth is often affected by the length of the dormant period during which no appreciable growth occurs. Growth by trees in tropical rain forests is nearly uniform throughout the year. In temperate zones, however, lower winter temperatures or rainless periods in arid regions result in distinct dormant periods.

When Europeans first reached North America forests covered most of the lands from the Atlantic to the inland prairies of Illinois and clothed vast areas in the Rocky Mountain region and along the Pacific

/contd. over..

coast. Today after several hundred years of exploitation by man, all that remains of this virgin continental cover are bands and dots of forest scattered over the land.

If one could have looked at Britain from the air before the Romans invaded the country, the forest cover would have appeared almost continuous from north to south, with only a few open areas on chalky downs and Scottish heaths. The forests of Britain have suffered many vicissitudes, however, from encroachment by man, from insect pests, fire and exploitation.

On the European continent the Scandinavian countries have a larger percentage of their surface still supporting forests than has any country to the south. Around the eastern end of the Mediterranean the forests have nearly disappeared having succumbed to the inroads of cutting, fire, insect pests and the browsing of domestic animals. A few mountain slopes and canyons contain relatively insignificant stands of forest.

The primary tropical rain forests of Africa are very complex, similar in structure to those in the wetter parts of South America. The trees grow in three welldefined layers or stories, the upper-most towering to 150ft. or slightly more and standing well above the tops of the continuous layer of the second story which terminates at about 75-100ft. A lower third story is made up of shrubs and small trees that tolerate heavy shading.