ASPECTS OF THE PSYCHOLOGY OF SECOND LANGUAGE VOCABULARY LIST LEARNING

GERARD F. GRIFFIN

THESIS SUBMITTED FOR THE DEGREE OF PH. D.

UNIVERSITY OF WARWICK

DEPARTMENT OF PSYCHOLOGY

SEPTEMBER 1992

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

My thanks are due first of all to Trevor Harley. He introduced me to psycholinguistics in the course of a very happy M.Sc. year at the University of Warwick and he has guided me through to the completion of this thesis with his usual mixture of erudition, enthusiasm, and wit. I would also like to thank him and Siobhan MacAndrew for all their kindness and hospitality over the last seven years. Being a part-time student can be a isolated task; they always made me welcome and gave me encouragement on the many occasions when I needed it. They are good friends.

I also wish to thank all my colleagues at Dorset Institute of Higher Education, Dorset Institute, Bournemouth Polytechnic, and Bournemouth University (the distinction is logical rather than ontological). The Directorate, my Heads of Department, departmental colleagues (academic, administrative, and secretarial), and the library staff have all been understanding, supportive, and helpful in many different ways. The same applies to the staff in the Department of Psychology and in the library at the University of Warwick.

Thanks are due to Wanda Arti, Tony Campolini, Nick Hobbs, and the children of their respective schools for making the running of the experiments possible.

I want to thank my children, Joe, Lucy, and Ben who supported me through some very difficult personal circumstances and who accepted my disappearance every evening to slave over a hot Macintosh with equanimity. They have been most understanding.

Finally, I want to thank Linda Butcher, my friend and partner, who has made me happy. I could not have done it without her.

LIST OF TABLES AND FIGURES

SUMMARY

The learning of second language vocabulary in lists of word-pairs is a widespread practice despite the disapproval of many in the second language learning domain. There is an acknowledged mismatch between psychological theories on the one hand and techniques of vocabulary learning on the other. Psychology does not address the relevant issues directly and second language learning practice is often atheoretical and unprincipled. This thesis reviews aspects of psychology which appear to be relevant to second language vocabulary learning and their applicability. A series of experiments is conducted with comprehensive school students learning French, aged 11-13.

The first part of the study deals with the presentation of vocabulary items to be learned. Presenting items in the order First Language - Second Language is the more versatile form of presentation if both generation and comprehension are required on the part of the learner. The transferability of list learning to testing in a sentential context depends on the ability of the learner and the task involved. Higher-ability list learners are inhibited in a generation task but not in a comprehension task; the opposite is true for lower-ability learners. Learning in a context improves the performance of higher-ability learners in generation but makes little difference to lower-ability learners. An explanation is suggested in terms of transfer-appropriate processing. The position of items in the list is not a reliable indicator of learnability. Primacy, recency, and serial effects may be obtained but none of them is consistent. The same conclusion applies to different ways of presenting word-pairs.

The second part of the study examines aspects of word learnability. Objective word frequency is not a reliable indicator of learnability in this context. Word category and the presence of an English word embedded in a French word are promising indicators of learnability.

CHAPTER 1

Introduction

The purpose of this thesis is to investigate some aspects of the psychology of learning vocabulary items in a second language (1). This is to be understood as "sitting down to learn vocabulary" as practised by students in schools, universities, and evening classes rather than in the sense of incidental learning of vocabulary as, for example, by so-called immersion methods, or by being brought up bilingually. Second language vocabulary learning, taken in this sense, is a rather odd phenomenon in that it is at the same time both "artificial" and "natural". It is artificial in the sense that it is a conscious and deliberate attempt to learn items of language whereas most normals appear to learn items of their native vocabulary in an unconscious and trouble free manner (McKeown & Curtis, 1987, passim). It is natural in the sense that it is something which many people undertake and indeed accomplish; when accomplished, the items learned function in much the same way as first language items for such purposes as understanding and making oneself understood.

The thesis is directly concerned with examining those areas of psychology which appear to underpin the theory and practice of second language vocabulary learning and it concentrates in particular on list learning by beginners in a foreign language, in this case English students learning French. For many years, the study of second language vocabulary learning was designated a "neglected area" (Broadbent, 1967; Ellis, 1985; Levenston, 1979; Meara, 1983). If this is no longer true in a general sense (see surveys by Meara, 1987; Nation, 1987, 1990), it still seems to be true for the study of list learning of second language vocabulary despite the widespread use of lists by language learners (see Carter & McCarthy, 1988; Nation, 1982).

Note (1). Although distinctions are made in the literature between foreign language learning and second language learning, the distinctions would serve no useful purpose here and the terms are used interchangeably.

The reasons why the study of vocabulary has been somewhat neglected are many. Second language vocabulary acquisition was often seen as a by-product of the student's attention to structure or to the process of communication (see Allen, 1983; McCarthy, 1984, for a discussion). According to Stern (1983), this notion goes back at least as far as Gouin (1880) and it is still found as late as Rivers (1981). There is an implied analogy between second language vocabulary learning and a child's acquisition of his/her first language, the vocabulary for which is assumed to have been achieved effortlessly and "naturally". Krashen's (1978, 1981, 1985) distinction between acquisition, "a subconscious process identical in all important ways to the process children utilize in acquiring their first language" (Krashen, 1985, p. 1), and learning, "a conscious process that results in 'knowing about' language" (ibid.), with its preference for methods of acquisition as opposed to learning, formalised this analogy. Both the analogy and Krashen's distinction have been seriously questioned in recent years (see, for example, McLaughlin, 1987). However, even if the analogy is accepted, there is no evidence to suggest that indirect learning cannot be complemented by direct learning (Nation, 1982) and this is particularly true when time is limited.

Again, it is often felt that acquiring vocabulary in a second language (L2) is a complex activity which cannot simply be reduced to learning an L2 equivalent for a first language (L1) item. J. C. Richards (1976) argued that knowing a word involves knowing not only its denotation but its frequency of use, the constraints on its use, its syntactic behaviour, its forms and derivatives, and its network of associations. Balhouq (1976) stressed the difficulties associated with polysemy, homonyms, non-isomorphism, connotations, taboo words, *faux amis*, idioms and clichés, and over-extension of collocations. However, the claim of Hughes (1968) that : "[No] English word is really the equivalent of any foreign word" (p. 85), is surely too strong, as he himself implicitly acknowledges, and many writers have made the point that the use of an L1 equivalent can be an important first step in learning a word thoroughly (see, for example, Cornu, 1979; Green, 1970; Nagy & Herman, 1987; Smith, 1969; Sternberg, 1987).

Even when vocabulary learning is accepted in principle, list learning is often seen as an unsatisfactory learning method. This matter will be discussed in detail in subsequent chapters, but Hughes (1968), for example, is typical in claiming that list learning of translations is soon forgotten and not readily available for use.

It is clear that there should be a reciprocal relationship between psychological theory and the theory of second language vocabulary learning. A complete psychological theory must have something to say about the processes involved in second language vocabulary learning and should be able to inform practice; conversely, theories associated with second language vocabulary learning must be compatible with general psychological theory, and empirical data from second language learning will constrain that theory to some extent. As things stand, it is clear that this relationship is not as productive as it might be. First, this is because the state of psychological theory is more satisfactory than the state of the theory of second language vocabulary learning. Second, because the use made of psychological theory in the study of second language vocabulary learning appears to be inadequate. These points will be discussed in turn.

Meara (1987) contrasted the development of psycholinguistic research, taking place in a substantial theoretical framework, with that of applied linguistics which seems not to have established its theoretical underpinning in the same manner. Many writers have made the point that research into second language vocabulary learning is unconvincing. The area is described as not being genuine research at all but merely the reporting of new teaching methods without adequate controls (Carroll, 1963); as falling short of "scientific standards of a purist psychologist" (Broadbent, 1967); as having been "largely atheoretical and unsystematic" (Meara, 1980, p. 221); as having received "relatively little systematic attention" (Stern, 1983, p. 131); as needing to be directed to a "systematic and principled basis" (Corder, 1973, p. 109); as being "quite old and rather patchy" (Meara, 1983, p. ii); as being of "poor quality" (Meara, 1987, p. 3).

Where the use made of psychological theory is concerned, there are two aspects to be considered. On the one hand, the complaint is made that a good deal of psychological research is of limited value to second language teaching practitioners because foreign language acquisition is often not the point of experimentation. This means that any findings need to be interpreted; that the experimentation underestimates the complexity of the process because relevant variables are ignored; that results are badly explained or misleading in that averaging obscures important differences (Carroll, 1963; Nation, 1982).

On the other hand, it appears that the process for using psychological theory is somewhat *ad hoc* and often inaccurate, with theory "contributing little more than scientific patter and an impressive-sounding new jargon" (Stern, 1983, p. 24). Thus, for example, it is questionable whether it is useful to state that: "Research also shows...that there is a tight relationship between 'cognitive depth' and retention" (Nattinger, 1988, p. 65), unless the status of that research is indicated (as, for example, in Kolers and Roediger, 1984). Research does not support the notion of cognitive depth in relation to retention because, as will be discussed later, the notion of cognitive depth cannot satisfactorily be defined; at best the notion of cognitive depth has heuristic value. It is also questionable whether references to "rote learning" (e.g., Carter & McCarthy, 1988, p. 12) serve much purpose unless it is made clear how it is established that subjects engage in this activity as opposed to some more elaborated form of processing. As a final example, Richards (1970, pp. 90-91) among others quotes Michéa (1964): "An available word is a word which though not necessarily frequent, is always ready for use, and comes to mind when it is needed." It is difficult to envisage what model of language processing such a concept could fit into.

There appear to be two main tasks to be accomplished. The first is to review those areas of psychological research which are relevant to the present domain. Psychological theory is not static but dynamic, not simple but complex, therefore for use to be made of that body of theory a process of updating is always going to be necessary. It is doubtful whether it will be adequate merely "to scan the field of psychology and psycholinguistics so as to be cognizant of theories, concepts, studies, and research findings that appear relevant" (Stern, 1983, p. 333). The second task is to evaluate the extent of their usefulness and applicability by a set of controlled experiments. A side-effect of the experiments will be any contribution their results may make to the more general psychological debates with which the process started.

Nation and Coady (1988) asked the question whether "research experiments still have a value for teachers, even when contradictory claims result" (p. 109). The answer would seem to be that experiments which result in contradictory claims may have a value for practitioners provided an explanation is given for those conflicting results. In other words, it must be acknowledged that vocabulary learning is a complex activity; that slight changes in materials, subjects, tasks, procedures may lead to different results (Jenkins, 1979); that concepts from psychology cannot be applied "neat" if only because sets of factors interact differently in different circumstances. It is probably more important for practitioners to understand this than it is for them to be led to believe, for example, that a particular factor will lead to improved learning whatever the circumstances (McDonough, 1986). Subjects used in these experiments are all taken from a relatively homogeneous population. Over-generalisation is always a problem in psychological research and the variability of results even within the sample used is a timely reminder of the difficulty of applying theories in any straightforward manner.

It is clear that a number of disciplines, and subsets of those disciplines, impinge on the study of second language vocabulary learning and it is important at the outset to indicate the limits of the research undertaken here. In general terms, it is not concerned directly with Applied Linguistics, with the literature on second language teaching, or with the psychology of bilingualism (at least to the extent that bilingualism is taken to refer to more or less fluent bilinguals; see Romaine, 1989).

Applied Linguistics in this context is taken to be "a mediating discipline between theoretical developments in the language sciences and the practice of language teaching" (Stern, 1983, p. 35). It is concerned with issues such as the nature of language, the description of language, the concept of difficulty, and contrastive analysis (Corder, 1973). Although some of these issues will be touched on in the thesis, the line of enquiry followed is basically psychological and the wider interests of Applied Linguistics will not be of concern here.

The literature on second language teaching concerns itself with a range of issues such as the difference between learning and acquisition (Krashen, 1978, 1981, 1985), interlanguage (Selinker, 1972), contrastive analysis (Lado, 1957), linguistic universals, order of acquisition, and teaching methods. Useful surveys include Dulay, Burt, and Krashen (1982), Klein (1986), Larsen-Freeman and Long (1991), McLaughlin (1987). The intention here is not to study teaching techniques as such but, in the phrase of Hamers and Blanc (1989), "to analyse the psycholinguistic and psychological processes upon which sound teaching methodology should be based" (p. 215).

There is an extensive literature on bilingualism covering such issues as the social aspects of bilingualism; neuropsychological aspects; code-switching; the bilingual lexicon; the bilingual child (see Romaine, 1989, for a survey).

The over-riding reason for delimiting the thesis in this way is simply one of available space; each of the areas mentioned is a separate undertaking in its own right. Even so, indirect reference will be made to these areas of study and this is particularly true of some aspects of the literature on bilingualism. Meara (1983) argued that studies of bilingualism have much to offer in this field both because bilingualism is the goal of the undertaking and because many of the experiments on bilingualism are conducted with subjects who are learners rather than fluent bilinguals and the results of these experiments are of particular interest in the present context.

Assumptions

Although the thesis is concerned with issues in a specific domain, the domain does not exist in isolation and at this stage some of the more important assumptions that lie behind the enquiry will be outlined since it will be a major concern throughout that local explanations should be compatible with more general theories about language processes.

The first assumption is that any final explanation in psycholinguistics, as in psychology generally, must be a neurological explanation. In the present state of research such a goal is unattainable. However, acknowledging this as the goal is a useful reminder that boxes with arrows can only ever be temporary explanations; even as temporary explanations, those which seem incapable of any neurological implementation should be avoided. An important corollary of this assumption relates directly to the domain of second language acquisition. The system as a whole, and the language system in particular, is not "interested" in language as such, let alone in second, third, or fourth languages. The system is designed to process input and output; it is designed for a two-way process of communication for purposes of survival and well-being. Discussions about second language acquisition, therefore, must sit easily within wider explanations about memory and language processes. An example of where this point leads is related to the debate about differences between compound and co-ordinate bilingualism (see Stern, 1983, for a discussion). Ervin and Osgood (1954) envisaged an individual who can use two languages as fitting into one of two categories. Co-ordinate bilinguals handle their two languages as separate entities and they are considered to be "true bilinguals". Compound bilinguals on the other hand have their languages linked and typically they may well interpret one language through the medium of the other. This initial distinction has led to a great deal of debate in subsequent years and whole teaching methodologies have been based on it, particularly since the original work was taken to endorse the co-ordinate approach as superior. Further research led to conflicting results (Macnamara, 1967) and eventually the suggestion was made that the distinction be abandoned as no longer tenable or serving a useful purpose (McLaughlin, 1987). The principle of seeking a neurological and ecologically valid explanation could have cut through this debate by putting the burden of proof on those who would argue for co-ordinate bilingualism and by assuming in the absence of such proof that all bilinguals are compound bilinguals. There is no obvious reason for thinking that the system would benefit from having separate language systems, even if this were possible. The language of thought, often dubbed "mentalese", is natural-language independent (Fodor, 1975). It is accessed by various means, natural language input being one of them. A bilingual's two languages come together at this semantic level at least. But access to the semantic system is achieved through the activation of memory. If activation spreads, not only is a target "word" and its meaning activated by the input, but related "words" and their meaning are activated also. This effect is clearly demonstrated in the phenomenon of semantic priming (Meyer & Schvaneveldt, 1971) and associative priming (Ratcliff & McKoon, 1981). It is difficult to envisage how or why spreading activation could be restricted to a particular language when the conceptual system is abstract and language independent, and indeed Schwanenflugel and Rey (1986) showed that cross-language priming can take place. This argument, of course, relies on yet another assumption, that of spreading activation. But spreading activation does have general explanatory power; the idea of co-ordinate bilingualism does not appear to have any general theory of memory supporting it. In conclusion, those explanations will be preferred which are compatible with wider explanations of human language processes rather than explanations which appear to relate only to a subset of those processes and to be incompatible with them.

A second assumption is that the language process is more complex than a serial process model can account for. One possibility is that the process is parallel, interactive, and compensatory (McClelland, 1987; Seidenberg, 1989). Without going into detail at this stage, the process is envisaged as being carried out by multiple and simultaneous sources of information; in the case of reading, for example, these sources include feature information, orthographic knowledge, lexical knowledge, syntactic knowledge, semantic knowledge. The system is interactive because the output from lower level processes constrains the possibilities available to higher level processes. Conversely, higher level processes influence the accessibility of data in lower level processes. Finally, the system is compensatory because shortage of information from one source can be compensated for by contributions from other sources. Thus in the case of poorer readers or of degraded stimuli, there may well be reliance on higher level sources of information if the lower level processes do not provide sufficient information for comprehension to be achieved. The word "simultaneous" needs to be qualified in this context. It is intended to mean that sources of information are simultaneously active, and receiving input, not that they output information at the same rate. The distinction is important. Automatic processes deliver information quickly; controlled processes take more time (Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). This simple time difference can provide good explanations for otherwise confusing experimental data, as, for example, why adults appear not to suffer from incongruous context effects in the way that younger readers do (Stanovich, 1981; West & Stanovich, 1978). The assumption is justified therefore because of its explanatory power.

A third assumption is that the language system is typified by few mechanisms and many strategies. Parallel systems, whether in the brain or in computing, are well-known for being both powerful and difficult to control and predict and it is clear from the research literature that experimental results are heavily task-dependent (Kolers & Roediger, 1984; Postman & Schwartz, 1964; Smith, 1982). The system appears to adopt those strategies which are most appropriate to the task in hand. Thus Coltheart and Funnell (1987) reported the case of HG whose strategy for reading words aloud was different (and more successful) when reading from a list of words only, than it was when reading from a list of words and nonwords. This example could be multiplied many times from normals as well as from brain-damaged subjects. The corollary must be that parsimonious explanations should normally be sought and new mechanisms should only be invoked when explanations based on automatic process or strategies changed in response to task demands, are ruled out. (It should be noted that "strategy" in this context is used in a neutral sense; it does not necessarily imply conscious control on the part of the agent.) This assumption is related to the first one made and precludes *ad hoc* explanations which, although possibly convincing in themselves, are not compatible with wider explanations of mental processes.

Methodology

As mentioned previously, a number of writers have pointed out that a good deal of psychological research is of limited value to second language teaching practitioners. In addition to Nation's (1982) arguments on interpretation, complexity, and averaging, there are concerns about materials, subjects, and tasks. Higa (1965) pointed out the care which needs to be taken with materials. He suggested that there is a need to take into account the previous knowledge of learners, the pronounceability of words, the familiarity of words in

L1, the part of speech, and item imageability. Where subjects are concerned, account must be taken of learners' ability and the fact that subjects may not be employing techniques or methods which the experimenter wishes and assumes them to be using (Gershman, 1970; Lado, Baldwin, & Lobo, 1967). Where tasks are concerned, the relationship of the learning condition to the test condition must be taken into account (Bialystock, 1985; Durgunoglu & Roediger, 1987; Fry, 1960; Kopstein & Roshal, 1955; Stoddard, 1929).

With the above points in mind, several constraints have been adopted in the choice of experimental procedures. First, the basic task is the list learning of English-French or French-English word-pairs. Second, materials have been controlled as far as possible for frequency and syntactic category. Where more than one list has been used in an experiment, lists have been controlled for frequency (checked by a *t*-test). Unless experimental requirements dictated otherwise, the same pattern of items from syntactic categories (four concrete nouns, four abstract nouns, five verbs, seven "other words") has been maintained throughout; exceptions are noted in individual experiments. In addition, in Experiments 1-4, where lists were not sub-divided, an item or F_2 analysis has been carried out to ascertain whether effects are able to be generalised beyond the words sampled (Clark, 1973). This was not possible in Experiments 5-10; when lists were sub-divided, the number of items in each category was insufficient for F_2 analyses to be carried out.

All subjects used were in their first year of secondary education, aged between 11 and 13, and in their first year of learning French. This choice of subjects in one sense represents a restriction, but it was a restriction deliberately imposed on the study. It was felt that the use throughout of a relatively homogeneous groups of subjects would make a sound basis for comparing various methods of presentation and various aspects of word difficulty; as far as subjects were concerned, like would be compared with like. In the event, from Experiment 1 onwards considerable variations were found even within this sample. This unexpected outcome cast doubts on the usefulness of generalisations based on more heterogeneous groups of subjects, at least at this stage in the development of the area of study.

Three schools from the same geographical area (Bournemouth) were used. School A was a mixed-sex state comprehensive school; School B was an all-girls state comprehensive school; School C was a mixed-sex state comprehensive school. Different cohorts from the three schools were used in different experiments and these are identified by a numerical indication of the experiment in which they took part (e.g., A1, B3....).

In the first five experiments, subjects from two schools were available. The original intention was to pool results from the two schools because the similarity of age of subjects, their belonging to the same stage of foreign language learning, and the similarity of the type of school involved pointed to the groups of subjects being of comparable ability. However a preliminary analysis of results in Experiment 1 revealed a significant difference in performance between subjects from the two schools concerned. It was decided therefore that for each experiment where two schools were involved, a preliminary analysis of the data would be carried out. If there was a significant difference in performance between the two schools, separate analyses would be carried out. Where results from two schools are treated separately, subjects are identified as being of "higher-ability" or of "lower-ability". The terminology refers merely to subjects' performance in the particular task under consideration; higher-ability subjects are those who performed more successfully; lower-ability subjects are those who performed less successfully. The terminology is not intended therefore to have any explanatory value and is used only for ease of identification.

All experiments were conducted in as "natural" a manner as possible which in this case meant that learning and testing took place in the course of normal lessons. The advantage of this constraint is that in avoiding highly artificial experimental environments there is less chance of artificial effects being obtained. The disadvantage is that certain experimental techniques are ruled out and results are limited to a relatively high level of description. An appropriate next stage would be to move to a lower level of description and in particular to reaction timing.

Instructions to teachers and subjects were kept to a minimum with the intention of maintaining consistency across groups and conditions. Examples can be found in the Materials Appendix.

Pattern of experiments

In broad terms, the first six experiments are concerned with presentation issues. Experiment 1 seeks to find out whether, for beginners, it is more effective to present listed vocabulary items in the order English-French, or *vice versa.* The issue addressed in Experiments 2 and 3 is whether list learning successfully transfers to the demands on vocabulary recall in more "normal" situations. Experiment 4 deals with the question of whether provision of a context at learning is more satisfactory than presentation of items in a simple list. Experiments 5 and 6 look at alternative forms of word-pair presentation.

In addition to matters of presentation, another important aspect of vocabulary list learning is the effect of characteristics of the items themselves on the learning process; this is the concern of Experiments 7-10. The characteristics examined relate directly to psychological research on memory and reading. Experiments 7 and 8 look at word frequency in relation to list position and serial order. Experiment 9 examines the effect of word category on ease of learning. Experiment 10 considers the effect on the learning process of English words embedded in French items.

It is hoped that the results of these experiments will add to the understanding of the process of second language vocabulary learning and will indicate which areas of psychological research are likely to prove most productive in future research.

CHAPTER 2

The effect of the order of presentation of L1 and L2 items on word-pair list learning

A substantial amount of work was done, particularly in the 1960s, on list learning (e.g., Bower, Clark, Lesgold, & Winenz, 1969; Murdock, 1962; Segal & Mandler, 1967; Tulving, 1968). After many years of neglect, there is a growing interest in second language vocabulary acquisition (Meara, 1983, 1987; Nation, 1982, 1987, 1990). However, for various reasons which will be discussed in Chapter 3, an area which has largely escaped attention has been the kind of list learning carried out by beginners in foreign language learning. Learners will often be presented with lists of vocabulary items arranged as word-pairs, either in the order First Language - Second Language, L1-L2, or in the order Second Language - First Language, L2-L1. (Since the first language referred to throughout this thesis is English and the second language is French the terms English and French will be interchangeable with L1 and L2 respectively). Additionally, students will normally make their own vocabulary lists arranged in a similar manner.

The theoretical issues raised in this chapter have an immediate practical application because it is not clear whether, for beginners, it is more effective to present listed vocabulary items in the order English-French, or *vice versa*. A principled approach to this matter would seem to be somewhat overdue and there is a need for answers to five questions: given a word-pair A-B, whether the association between the two components of the word-pair is bi-directional; if it is bi-directional, whether the forward-association, A-B, is stronger than the backward-association, B-A, in the sense that A is more likely to lead to recall of B than *vice versa*; given that one component is familiar and the other unfamiliar, whether it is more effective to learn the familiar-unfamiliar association (L1-L2) or the unfamiliar-familiar association (L2-L1); whether generation or comprehension is the easier task; and whether direction of learning has an effect on remembering over time.

The directionality of the word-pair bond

The first question to be addressed is that of the directionality of the word-pair bond. When subjects learn a word-pair, A-B, is the bond between A and B uni-directional or bi-directional? Historically, Behaviorism accounted for the association between A and B in stimulus-response terms (see Voss, 1979, for a review). This meant, by definition, that the bond was taken to be uni-directional. However, Asch (1968) and Asch and Ebenholtz (1962) showed that either A or B can serve as a recall cue for an A-B word-pair. They argued that when a word-pair is learned, this involves the formation of both a forward-association and a backward-association of equal strength; this they designated associative symmetry. They saw the two associations as aspects of one entity, the relationship between the two items, and therefore as logically rather than ontologically distinct. Johnston (1967) showed this not to be true. The associations must be independent because the loss of the association in one direction does not necessarily entail the destruction of the association in the other direction. Lockhart (1969) showed that with a word-pair made up of an adjective and a noun, the noun is the more effective cue whether it is in the stimulus or response position at learning. Wolford (1971), developing the point, argued that the relationship is not one of associative symmetry, but of associative asymmetry. There are two separate and distinct associations, A-B and B-A; the learned forward-association is mainly responsible for recall and the learned backward-association for recognition. Finally, Voss (1972) argued that the word-pair is a complex entity involving not one bond but at least five processes which relate the two elements of the word-pair. If the word-pair is A-B, then given that A results in response r^{a}, and that B results in response r^{b}, and given that response r^{a} and response r^{b} have attached a whole set of aspects or dimensions:

> "It would seem that in order to have an adequate understanding of what happens when an association is acquired, it is necessary to have knowledge of at least five processes. The first is the question of how A is encoded, ?1. The second is how B is encoded, ?2. The third is what components of r^{a} become associated to what component of r^{b}, ?3. The fourth issue is how the component of r^{b} that is related to

a component of r^a becomes related to a component of r^b that initiates a response, ?4, that is, are these two aspects of r^b the same or are they different and, if different, how are they related. Finally, the last question is how the particular r^i initiates certain responses." (Voss, 1972, pp. 171-2)

It seems safe to assume that the bond is bi-directional, although the nature of that relationship is a subject of debate. There is a certain amount of independence between the forward- and the backward-association since one association can be lost without loss of the other. They are not simply aspects of the same entity. The choice of what features of A are associated with what features of B in learning an A-B word-pair is subject to individual differences. In other words, word-pair learning is not a question of simple association but is open to learner strategy. Indeed, mnemonic methods such as the loci method (Neisser, 1976), the keyword method (Atkinson, 1975; Merry, 1980), and the so-called Direct Method (Curran, 1976), all rely on strategies of association adopted by the learner.

Forward-association and backward-association

The second question concerns the relative relative strength of the forward- and backward-association. Although conceptually this is a different question from the third question concerning the effectiveness of learning when one component of the word-pair is familiar and the other unfamiliar, as is the case in learning L1-L2 word-pairs, in practice similar arguments apply in both cases since effectiveness of learning is established by performance at testing. Noble (1952) and Noble and McNeely (1957) reviewed evidence to show that under certain circumstances response factors rather than stimulus factors can be a better predictor of recall; in other words, under certain conditions the backward-association may be stronger. When stimulus and response meaningfulness (Noble, 1952) are manipulated, it is the meaningfulness of the response item which dictates recall whatever the meaningfulness of the stimulus item. A similar conclusion was reached by Cieutat, Stockwell, and Noble (1958), Glanzer (1962), Postman (1962), and Underwood and Schulz (1960). However, not all agree. Cason (1933) saw equal effects from changing the meaningfulness of the stimulus as well as the response (see Rodgers, 1969) and Gannon and Noble (1961) showed an effect derived from stimulus manipulation without an effect from response manipulation. Crothers and Suppes (1967) extended the scope of the discussion by calling attention to the range of factors which can affect the performance of stimulus or response. Using always L2-L1 learning, and following on from the work of Gannon and Noble (1961) and Underwood and Schulz (1960), they argued that the relevant factors for the L2 stimulus (in this case Russian) are pronounceability and association values, whereas syntactic class is the relevant factor for the L1 (English) item in the response position. Given the L2-L1 order of learning, they also showed that the two stimulus properties mentioned are more important indicators of performance than the effect of syntactic class in the response. They argued, however, that it is the degree of unfamiliarity of the L2 item which is important rather than its being in the stimulus position; thus a (difficult/unfamiliar L1)-(easy/familiar L2) word-pair would be easier to learn than an (easy/familiar L1)-(difficult/unfamiliar L2) word-pair.

Paivio (1969) made what was to become for him a characteristic distinction between association values and imageability values. He argued that where nouns are concerned, it is the imageability of the word in the stimulus position which is important; he saw this conclusion as being in conflict with the "common empirical generalization" that it is response factors which are crucial in paired-associate learning (Paivio, 1969, p. 245). He pointed to the finding of, for example, Goss and Nodine (1965) and Underwood and Schulz (1960) that meaningfulness (defined by the number of associations generated) in the response position is the important factor in learning paired nonsense-words and words. However, when familiar words rather than nonsense-words are used his claim was that imageability is a reliable indicator of performance whereas the effect of meaningfulness can range from slightly positive to slightly negative. Paivio (1971) returned to the issue suggesting that imageability and meaningfulness might be indicators of different sorts of mediation; where visual mediation is concerned, it is the stimulus position which is important; where verbal mediation is concerned, it is the response position which is important. Paivio (1971) carried out a detailed review of findings to date showing that a range of experimental effects can be obtained by manipulation of variables such as the use of children or adults as subjects, the use of words or nonwords, the use of concrete or abstract words, the relative imageability and association values of items in the stimulus or response position, the frequency of words used, and the syntactic class of words used. On the matter of the associative symmetry, Paivio argued that visual associations are symmetrical whereas verbal associations are directional with the forward-association being stronger than the backward-association.

Wolford (1971) saw the question of the strength of the forward- and backward-association as related to that of task demands. Where recall is concerned it is the forward-association which is going to appear to be stronger whereas the opposite is the case where the task is recognition. Bower (1972), in arguing for the use of mediational imagery, showed that image-evoking values varied on the stimulus and response sides of word-pairs are positively correlated with learning in both cases, but much more so on the stimulus than on the response side.

There is, therefore, no reliable answer to the question of the relative strength of the forward- and backward-association in isolation from a consideration of the task and materials involved. Either the item in the stimulus position at learning or the item in the response position at learning can be the more effective cue depending on the particular experimental conditions obtaining (Underwood, 1982). Horowitz and Gordon (1972) specifically addressed associative symmetry in the context of second-language learning. They argued that when a word-pair, A-B, is learned, then in addition to the forward-association, A-B, a latent backward-association, B-A, is formed. Symmetry between the two associations can be achieved by making the response to the backward-association, A, more available through overt practice. They concluded from this that learning L2 items in the order L2-L1 will result in efficient learning of the forward-association because the L1 item is relatively more available than the L2 item. Overt practice of the L2 item will then mean that the backward-association will make that association overt. Although the results of their first experiment is equivocal, their second experiment does show faster learning to a criterion than a control group. However, the relevance of their finding to the present discussion is questionable since the performance they report is dependent on additional learning, in the form of further practice of the B-A response, which would seem to present motivational problems with respect to young learners.

Another aspect of the discussion was introduced by Jones (1976). He proposed that a remembered item is to be thought of as a fragment of a perceived situation, or an all-or-none cluster of components of the original stimuli. If a cue overlaps with any aspect of the fragment, then the whole fragment becomes available. Therefore, one component of the L1 cue could cue the L2 item or *vice versa*. In this sense, memory traces are symmetrical. However, he argued that reflexivity is not a corollary of the theory if multiple cues are involved. In other words, if A can cue C, and B can cue C, this does not mean that B will necessarily activate A. In this sense, the relationship may be seen as symmetrical or asymmetrical depending on the characteristics of the target item. Spyropoulos and Ceraso (1977) discussed the notion of limited access to memorised items in terms of categorised and uncategorised attributes. They found that in the case of a unitary to-be-remembered item, for example a red triangle, then the categorised aspect of the item is the most effective cue for recall of all aspects of the item. The categorised aspect is taken to be the defining aspect of the item. Thus if subjects were instructed that x was a triangle (which was red) as compared with a red object (which was triangular), then its being a triangle would be its categorised aspect. Following the reasoning of Asch and Ebenholtz (1962), and Horowitz and Gordon (1972), it would be expected that the weaker (non-categorised) attribute would be the better cue because in this situation the response (the categorised attribute) would be more available. The explanation offered by Spyropoulos and Ceraso (1977) was that the unitary item forms a cohesive unit in memory which is difficult to access except through its salient aspect. However, where the to-be-remembered item is non-unitary or poorly integrated, such as a word-nonword pair, then the less salient aspect, presumably the nonword, will be the better cue because the word part of the item will be more available. The relationship between attended (categorised) and unattended (uncategorised) attributes of memorised units was taken up by Jones and Martin (1980). Over three experiments, inconsistent results were obtained and the conclusion was reached that the idea of cue categorisation as a key to recall is not robust since its effect can be varied by relatively slight changes in experimental design.

Generation and Comprehension

The fourth question concerns task demands. In this thesis, generation is defined throughout as the production of an L2 item in response to an L1 cue, irrespective of the direction of learning. Comprehension is defined as the production of an L1 item in response to an L2 cue, again irrespective of the direction of learning. It could be argued that the comprehension task is not as difficult as the generation task. Stoddard (1929) estimated that recall for comprehension was twice as easy as recall for generation; Horowitz and Gordon (1972) similarly found comprehension to be the easier task, as did Mägiste (1979). The target in the case of comprehension is a word which, presumably, is already well integrated into the network of memory and which is therefore accessible in a variety of ways; the target in the case of generation is a word which has little "status" in memory and means of access to it are for that reason restricted. In comprehension, then, subjects are working towards the well known; if the provision of the French cue results in even a minimal activation of memory, then subjects are in a position to choose between candidate responses, and those responses will be real English words; the task then becomes very close to being a recognition task which is known to be easier than recall (see Eysenck, 1984, for a survey). In the case of generation, subjects are working from the known to the less well known or even the unknown. The provision of the English cue will result in the activation of memory, but it is quite feasible that no French candidates for a response will be automatically generated. If subjects set out to generate candidates, the set of potential candidate words is small where a beginner is concerned and there is a good possibility that the candidates generated will not be words at all but pseudo-words, mere strings, or attempts to "Gallicise" English words. The matter is of some practical importance to the main issue of whether presentation in the form English-French or French-English is to be preferred. If comprehension is indeed easier than generation, there would be a case for learning in the direction English-French. The relative strength of the forward-association would make for the more successful completion of the more difficult task; the relative weakness of the backward-association would be offset by the relative ease of the comprehension task.

Trace persistence over time

The final question is concerned with whether there is any connection between direction of learning (whether subjects learn L1-L2 or L2-L1) and persistence of learning over time measured by the ability of subjects to recall items over a relatively extended period. As will be discussed more fully in Chapter 5, in list learning it is possible for vertical connections to be set up as well as the desired horizontal word-pair associations. When learning is in the direction L2-L1, the expected target words are English items. It would therefore be relatively easy for subjects to build on the vertical connections between the English items which would already be primed by spreading activation. For French-English learners expecting to be tested in the same direction as learning, therefore, a reasonable strategy would be to make use of the assistance which the list provides and to spend correspondingly less time on the word-pair itself. Where learning takes place from English to French and the expected target words are French, the same strategy would not be appropriate. Subjects are faced with forming associations between words which are new to them, none of which would have much in the way of memory connections in their own right (although, as discussed in Chapter 8, there is always the possibility that during the learning process individual French items become associated with orthographically similar English words and that associations are formed between these English words). It could be, then, that learning English-French is more difficult than learning French-English. However, there is evidence to suggest that more difficult learning can lead to better retention over time. Battig (1979), Heim, Watts, Bower, and Hawton (1966), Jacoby and Craik (1979), and Jenkins (1979) all showed that initial learning difficulty can lead to better long-term retention, although Bahrick and Phelps (1987) found no disadvantage attached to difficulty of learning as opposed to any advantage in it (see also Lado, Baldwin, & Lobo, 1967). The idea of depth of processing will be discussed in detail in Chapter 4 but, in brief, it is argued that any item has many facets, features, or dimensions along which it can be processed and the more varied the processing that takes place, the stronger the bond between the word-pair items, and the more cues there are for retrieval. It is argued that where initial learning is difficult, subjects may involve themselves in more varied processing and will in consequence form a stronger word-pair link which is less dependent on the list context. Another factor in these experiments is that the test list is in a different order from the learning list. This could be seen to favour the less list-dependent form of learning of the English-French learners whose strategy prepared them better for the randomised test list.

All of these issues are addressed in Experiment 1 which besides having practical implications for vocabulary learning would influence the design of subsequent experiments.

EXPERIMENT 1

The purpose of Experiment 1 was to examine the five questions discussed above. First, whether L1-L2 word-pair associations are uni-directional or bi-directional (Asch & Ebenholtz, 1962; Johnston, 1967; Lockhart, 1969; Wolford, 1971). Second, if word-pair associations are bi-directional, whether the forward association is equal to the backward-association (Asch & Ebenholtz, 1962), stronger than the backward-association (Bower, 1972; Paivio, 1969, 1971), or weaker than the backward-association (Cieutat, Stockwell, & Noble, 1958; Goss & Nodine, 1965; Noble, 1952; Noble & McNeely, 1957; Underwood & Schulz, 1960). Use of the forward-association is defined as being tested in the same direction as learning. Thus for the forward-association, English-French learners will be tested English-French, being required to produce a French response to an English cue; French-English learners will be tested French-English being required to give an English response to a French cue. Use of the backward-association means being tested in the opposite direction from learning; this would be French-English for English-French learners and English-French for French-English learners. Third, given that L1 items are familiar to subjects and L2 items unfamiliar, whether learning L1-L2 is more effective than learning L2-L1 (Jones, 1976; Spyropoulos & Ceraso, 1977). Fourth, whether generation or comprehension is the easier task. Generation in this context means giving an L2 item in response to an L1 cue, irrespective of the direction of learning. Comprehension in this context means giving an L1 response to an L2 cue irrespective of the direction of learning. Fifth, whether direction of learning has an effect on remembering over time (Battig, 1979; Heim, Watts, Bower, & Hawton, 1966; Jacoby & Craik, 1979; Jenkins, 1979).

Method

Design

This experiment was a 2 x 2 x 4 design, with two between-subjects factors and one within-subjects factor. The between-subjects factors were: use of forward- or backward-association at testing; direction of learning (English-French or French-English). The within-subjects factor was time (the four test days). The first day was the same day as learning; it was followed by tests on the third, seventh, and twenty eighth day.

Materials

There are some generalisations which can usefully be made about the materials used throughout the following experiments. In all cases, there were 20 word-pair items to be learned. The length of lists was decided on the basis of an informal pilot study. As far as could be ascertained after consultation with teachers concerned, none of the French items had been encountered by subjects in their studies prior to the experiments. Materials used for all experiments are detailed in the Materials Appendix and identified by the number of the experiment.

In Experiment 1, the same list was used for Group 1 and Group 2 and a different list for Groups 3 and 4; this was because subjects in Groups 3 and 4 came from the same pool of subjects who had previously been members of Groups 1 and 2. There was no significant difference between the overall frequency of the English items in the two lists. The French components, by definition, were all equally unfamiliar to the subjects.

Word-pair ordering, English-French or French-English, was manipulated appropriately. Thus, for example, the first item for Group 1 was **a wardrobe - une armoire**; for Group 2, **une armoire - a wardrobe.** The pattern of words was similar for both lists; in each case there were four concrete nouns and four abstract nouns, five verbs, seven others.

Cue words on test papers were identical to the original but the order in which they appeared was randomised across groups; this was designed to reduce possible effects of list dependency.

Subjects

In Experiment 1, 47 subjects were used from School A1, and 63 subjects from School B1. As stated in Chapter 1, all subjects were aged between 11 and 13 and were in their first year of learning French. Both schools were state comprehensive schools; however, School A1 was a mixed-sex school, the pool comprising 24 girls and 23 boys, and School B1 was an all-girls school.

Procedure

In general terms, the procedure followed was consistent across experiments. Subjects were advised that the experiment was designed to aid language learning, in the long run, and they were encouraged to do the test as well as possible and without communication with their companions. Being right or wrong was not as important as doing the tests as well as possible. The term "list" was not used and no instruction was given on either learning technique or mode of testing. Examples of instructions are contained at the start of the Materials Appendix.

In all the experiments standard learning and testing times were used. Learning was restricted to eight minutes and testing to two minutes. These times were established on the basis of informal pilot studies. Eight minutes was the appropriate learning time for

Table 2.1. Experiment 1.

Arrangement of Groups.

Group	Direction of learning	Use of forward- or backward-association at testing (Note 1)	Test condition (Note 2)
1	English-French	Forward	Generation
2	French-English	Forward	Comprehension
3	English-French	Backward	Comprehension
4	French-English	Backward	Generation

Note 1. Use of forward-association means being tested in the same direction as learning. Use of backward-association means being tested in the opposite direction from learning.

Note 2. Generation in this context means giving an L2 item in response to an L1 cue irrespective of the direction of learning. Comprehension in this context means giving an L1 response to an L2 cue irrespective of the direction of learning.

adequate learning to take place while engaging the attention of subjects. Two minutes was an adequate amount of time for subjects to respond. It was assumed that the activity of handing in learning lists and receiving test lists was sufficiently long and distracting to preclude short-term memory effects such as rehearsal.

Experiment 1 took place over the Spring term and Summer term of the school year. Groups 1 and 2 were tested in the Spring term. Groups 3 and 4 were tested in the Summer term. Subjects had had, therefore, six months of formal French teaching when they began the experiment. Tests subsequent to the day of learning were not announced to subjects in advance; this was intended to avoid encouraging relearning.

In Experiment 1, four experimental groups, numbered 1-4, were formed in each school. The arrangement of the groups is shown in Table 2.1. Group 1 learned English-French and was tested English-French. Group 2 learned French-English and was tested French-English. Group 3 learned English-French and was tested French-English. Group 4 learned French-English and was tested English-French. Membership of the four groups related to the between-subjects factors and the five questions posed in the following way.

Groups 1 and 2 used the forward-association at testing and Groups 3 and 4 used the backward-association at testing. Therefore a comparison between the performance of Groups 1 and 2 on the one hand and Groups 3 and 4 on the other hand would provide answers to the first two questions about the existence of a backward-association and its strength relative to the forward-association.

Groups 1 and 3 learned English-French; Groups 2 and 4 learned French-English. Therefore a comparison between the performance of Groups 1 and 3 on the one hand and Groups 2 and 4 on the other hand would provide an answer to the question of the relative effectiveness of the two directions of learning.

Groups 1 and 4 were tested for generation; Groups 2 and 3 were tested for comprehension. Therefore a comparison between Groups 1 and 4 on the one hand and Groups 2 and 3 on the other hand would provide an answer to the relative ease of the generation and comprehension task.

The final question would be answered by a possible interaction between direction of learning and ability to recall over time.

In School A1, Group 1 was comprised of 23 pupils (12 boys, 11 girls); Group 2, 23 pupils (11 boys, 12 girls); Group 3, 25 pupils (12 boys, 13 girls); Group 4, 22 pupils (11 boys, 11 girls). In School B1, Group 1 was comprised of 30 pupils; Group 2, 33 pupils; Group 3, 16 pupils; Group 4, 17 pupils.

Results and Discussion

As discussed above, a preliminary analysis of variance was carried out to ascertain whether there was a significant difference between the performances of subjects from the two schools. The ANOVA revealed a significant difference in performance between subjects in School A1 (percentage mean of correct responses, 46.70%) and subjects in School B1 (percentage mean of correct responses, 29.04%), F (1, 181) = 45.51, $p < 0.01$. There was also a significant interaction between School x Day of testing, F (3, 543) = 5.52, $p <$ 0.01. For these reasons separate analyses were conducted on the data for the two schools.

Results for School A1: Higher-ability subjects

An analysis of variance was performed. The ANOVA summary table for this and all subsequent experiments is to be found in the Results Appendix and identified by the number of the experiment. Between-subjects factors were: use of forward-association or backward-association at testing; direction of learning. The within-subjects factor was the four test times. Mean scores are contained in Table 2.2.

The first question concerned the directionality of the word-pair association. If the association were uni-directional presumably the performance of Group 3 and Group 4 would be nil, since subjects in these groups were tested in the opposite direction from learning. It is clear that the association is bi-directional. Learning in one direction did not preclude performance in the opposite direction, Group 3 averaging 37.41% items recalled and Group 4 averaging 30.00% items recalled.

The second question concerned the relative strength of the forward-association and the backward-association. It is clear that the forward-association is stronger than the backward-association. This can be determined by comparing the performance of Groups 1 and 2 with the performance of Groups 3 and 4. The mean percentage score for Groups 1

Table 2.2. Experiment 1. School A1.

Mean percentage scores for items recalled: Forward- or backward-association, direction of learning, and test condition.

Association		**Direction of learning**		**Test**	
Forward		**English-French**		**Generation**	
Group 1	52.98	Group 1	52.98	Group 1	52.98
Group 2	66.41	Group 3	37.41	Group 4	30.00
Mean	59.69	Mean	45.19	Mean	41.49
Significance of difference (Tukey test)	Not significant	Significance of difference (Tukey test)	$p < 0.05$	Significance of difference (Tukey test)	$p < 0.01$
Backward		**French-English**		**Comprehension**	
Group 3	37.41	Group 2	66.41	Group 2	66.41
Group 4	30.00	Group 4	30.00	Group 3	37.41
Mean	33.70	Mean	48.20	Mean	51.91
Significance of difference (Tukey test)	Not significant	Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	$p < 0.01$
Difference between forward- and backward-association significant at: $p < 0.01$.		Difference between directions of learning not significant: $p > 0.4$.		Difference between generation and comprehension significant at: $p < 0.01$.	

and 2, using the forward-association, was 59.69%; the mean percentage correct score for Groups 3 and 4, using the backward-association, was 33.70%. This difference was significant in relation to the subjects used, F_1 (1, 89) = 44.26, $p < 0.01$. The analysis also suggests that the difference between the forward-association and backward-association is able to be generalised beyond the words sampled, F_2 (1, 76) = 46.61, $p < 0.01$ (see Clark, 1973).

The third question concerned the effectiveness of learning English-French or learning French-English. Direction of learning did not have a significant effect, F (1, 89) = 0.59, $p > 0.44$. English-French learners (Groups 1 and 3) averaged 45.19% correct responses and French-English learners (Groups 2 and 4) averaged 48.20%. Taking the question in isolation, therefore, there is no reason for deciding to present materials in one direction rather than the other. There is nothing inherently more difficult about learning in the direction English-French as compared with learning French-English. It appears that for School A1, at least, subjects were well motivated and learned in both directions equally well.

The fourth question concerned the relative ease of generation and comprehension as tasks. This is effectively measured by the interaction between forward- and backward association and direction of learning. The results show (Table 2.2) that the mean percentage score for correct responses for generation (Groups 1 and 4) was 41.49% and for comprehension (Groups 2 and 3) was 51.91%; this difference was significant, F_1 (1, 89) = 7.10, $p < 0.01$; F_2 (1, 76) = 7.34, $p < 0.01$. Overall, therefore, comprehension is an easier task than generation.

The possibility had been considered that although the French-English bond might appear to be easier to establish, the English-French bond might be stronger over time due to initial difficulty of learning and due to its lack of list dependence. In general terms, day of testing had a significant effect, F_1 (3, 267) = 23.60, $p < 0.01$; F_2 (3, 228) = 28.84, $p < 0.01$. Performance on Day 1 was better than performance on all three subsequent days, $p < 0.01$ (pairwise comparison, Tukey test). Percentage means are contained in Table 2.3 (and see Figure 2.1). Forgetting took place between Day 1 and Day 2 but after that the

Table 2.3. Experiment 1. School A1.

Mean percentage scores for items recalled: Direction of learning, day of testing.

Learning direction	**Day 1**	**Day 2**	**Day 3**	**Day 4**	**Overall**
English-French	51.92	43.69	45.41	39.76	45.19
French-English	54.27	46.86	45.80	45.87	48.20
Overall	53.10	45.28	45.61	42.81	46.70

memory trace remained relatively stable. However, there was no significant interaction between the language-order of learning and day of testing, F (3, 267) = 1.68, $p > 0.16$. The English-French bond and the French-English bond decayed at much the same rate.

So far, then, it is clear that the forward-association is stronger than the backward-association; there is no significant difference due to the direction of learning; comprehension is an easier task than generation. In relating these findings to the domain in question, it is necessary now to consider which direction of learning would be the more versatile in the sense of being more suitable for both generation *and* comprehension. The results indicate that the best way of achieving equal fluency in the two processes would be to learn vocabulary items in both directions. For purposes of comprehension learning in the direction French-English is the more effective, since it is the forward-association of French-English which will be used at recall. For generation, learning in the direction

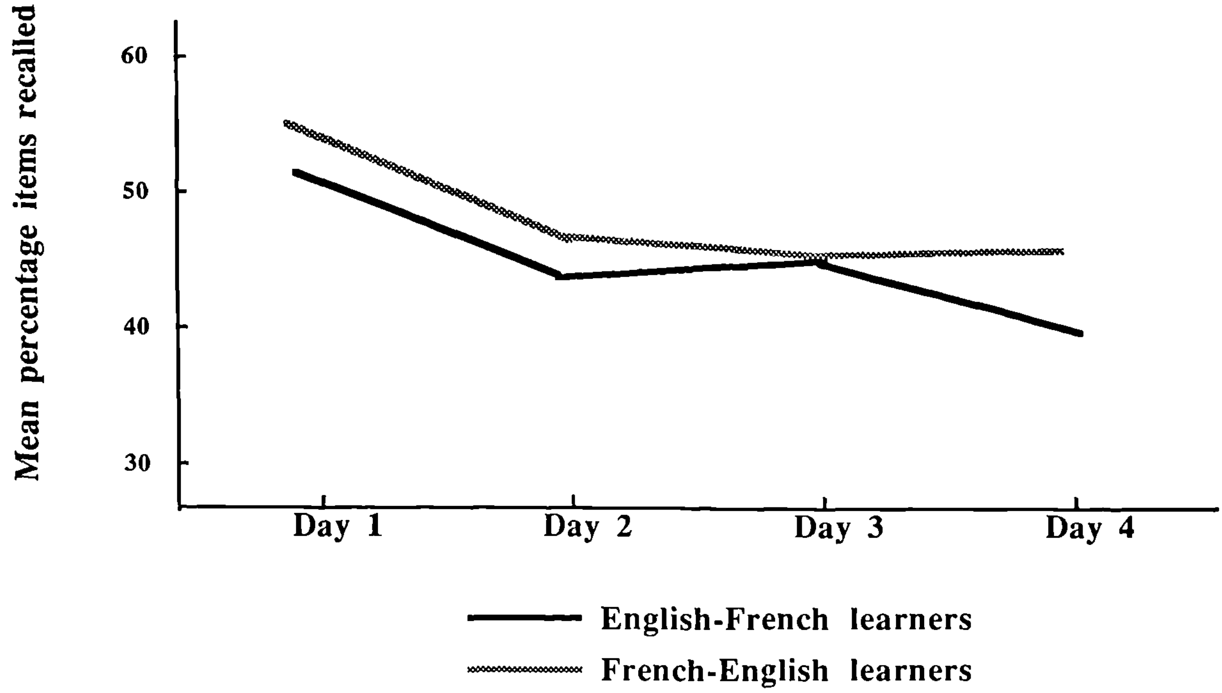

Figure 2.1. Experiment 1. School A1. Mean percentage scores for items recalled: Direction of learning, day of testing.

English-French is the more effective, since it is the forward-association English-French which will be used at recall. However, this solution, as mentioned above, seems to present severe practical and motivational problems since vocabulary learning even in one direction is usually considered a tedious task. If one learning condition is to serve both purposes, then some sort of trade-off is going to be necessary. The data is inconclusive in this respect although it does seem to favour slightly the English-French learning condition. There is no significant difference between the combined performance of Group 1 and Group 3 (English-French) compared with the combined performance of Group 2 and Group 4 (French-English), F (1, 89) = 0.59, $p > 0.44$. However, a comparison between the two English-French groups (Group 1 and Group 3) and the two French-English groups (Group 2 and Group 4) over the generation and comprehension tasks shows that the rate of

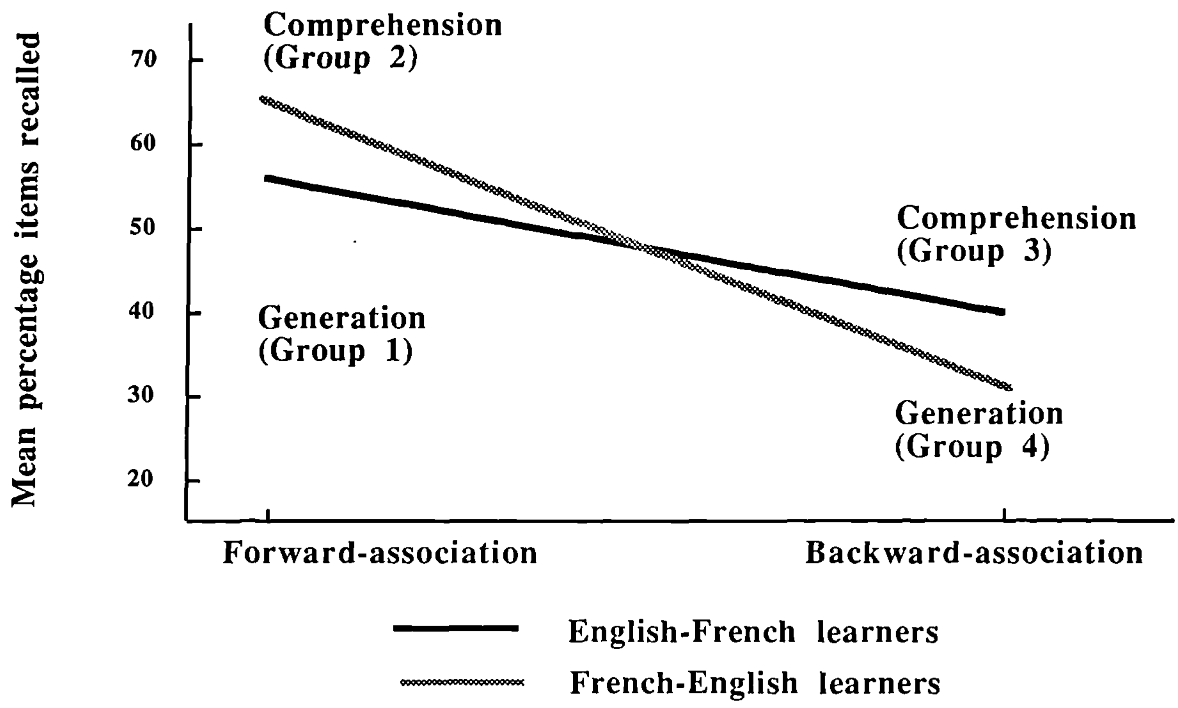

Figure 2.2. Experiment 1. School A1. Mean percentage scores for items recalled: Use of forward- or backward-association, direction of learning, test condition.

fall-off of performance is steeper for French-English learners than it is for English-French learners. As Table 2.2 indicates (and see Figure 2.2), the decrement for English-French learners (Groups 1 and 3) is significant only at $p < 0.05$, whereas the decrement for French-English learners (Groups 2 and 4) is significant at $p < 0.01$ (pairwise comparison, Tukey test). In other words, the English-French learning condition is less disadvantaged by the demands made on it by the reverse condition than is the French-English learning condition.

Another reason for using the English-French learning condition, which is an expansion of the point just made, arises from the relative ease of comprehension compared with generation, and the relative strength of the forward-association compared with the backward-association. English-French learners use the stronger forward-association for the more difficult generation task; French-English learners must use the weaker backward-association for the more difficult task. It would seem reasonable, therefore, to conclude that English-French is on balance the more versatile direction for learning when both generation and comprehension are required.

Results for School B1: Lower-ability subjects

An analysis of variance was performed (see Results Appendix for ANOVA summary tables). As with School A1, this experiment was a 2 x 2 x 4 design, with two between-subjects factors and one within-subjects factor. The between-subjects factors were: use of forward-association or backward-association at testing; direction of learning. The within-subjects factor was the four test times. The first day was the same day as learning; it was followed by tests on the third, seventh, and twenty eighth day. The mean scores are contained in Table 2.4.

In answer to the first question, it is clear that a bi-directional bond is established because learning in one direction did not preclude performance in the opposite direction. Although Group 3 and Group 4 were tested in the opposite direction from learning, Group 3 averaged 26.39% items recalled and Group 4 averaged 10.14% items recalled.

Given this bi-directional association, the forward-association is stronger than the backward-association. The mean percentage score of correct responses for the forward-association (Group 1 and Group 2) was 39.81%; the mean percentage score for the backward-association (Group 3 and Group 4) was 18.26%. The difference was significant over subjects, $F_1\ (1, 92) = 38.53, p < 0.01$, and over items, $F_2\ (1, 76) = 28.62, p < 0.01$.

As with subjects from School A1, language-order at learning did not have a significant effect on performance, $F\ (1, 92) = 0.08, p > 0.76$. English-French learners averaged 28.52% correct responses and French-English learners averaged 29.55%.

Table 2.4. Experiment 1. School B1.

Mean percentage scores for items recalled: Direction of learning, association, and task.

Association		Direction of learning		Task	
Forward		**English-French**		**Generation**	
Group 1	30.66	Group 1	30.66	Group 1	30.66
Group 2	48.97	Group 3	26.39	Group 4	10.14
Mean	39.81	Mean	28.52	Mean	20.40
Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	Not significant	Significance of difference (Tukey test)	$p < 0.01$
Backward		**French-English**		**Comprehension**	
Group 3	26.39	Group 2	48.97	Group 2	48.97
Group 4	10.14	Group 4	10.14	Group 3	26.39
Mean	18.26	Mean	29.55	Mean	37.68
Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	$p < 0.01$
Difference between forward- and backward-association significant at: $p < 0.01$.		Difference between directions of learning not significant: $p > 0.76$.		Difference between generation and comprehension significant at: $p < 0.01$.	

The fourth question concerned the relative difficulty of generation and comprehension which is measured by the interaction between forward- and backward-association and direction of learning. The results show that comprehension is an easier task than generation, F_1 (1, 92) = 24.76, $p < 0.01$, F_2 (1, 76) = 18.75, $p < 0.01$. Groups 2 and 3, tested for comprehension, had a mean percentage score of 37.68%; Groups 1 and 4, tested for generation, had a mean percentage score of 20.40%.

In general terms, day of testing was a main effect, F_1 (3, 276) = 58.91, $p < 0.01$, F_2 (3, 228) = 119.46, $p < 0.01$. The point of interest here is that although the only significant difference at $p < 0.01$ is between results from Day 1 and results from the other three days, for the lower-ability learners in this experiment there was also a tail-off on Day 4, significant at $p < 0.05$ (pairwise comparison, Tukey test). This would seem to be a further indication of the relative weakness of the word-pair link established by School B1

Table 2.5. Experiment 1. School B1.

Mean percentage scores for items recalled: Direction of learning, day of testing.

Direction of learning	Day 1	Day 2	Day 3	Day 4	Overall
English-French	40.73	25.20	26.98	21.18	28.52
French-English	39.47	27.52	26.92	24.31	29.55
Overall	40.10	26.36	26.95	22.74	29.04

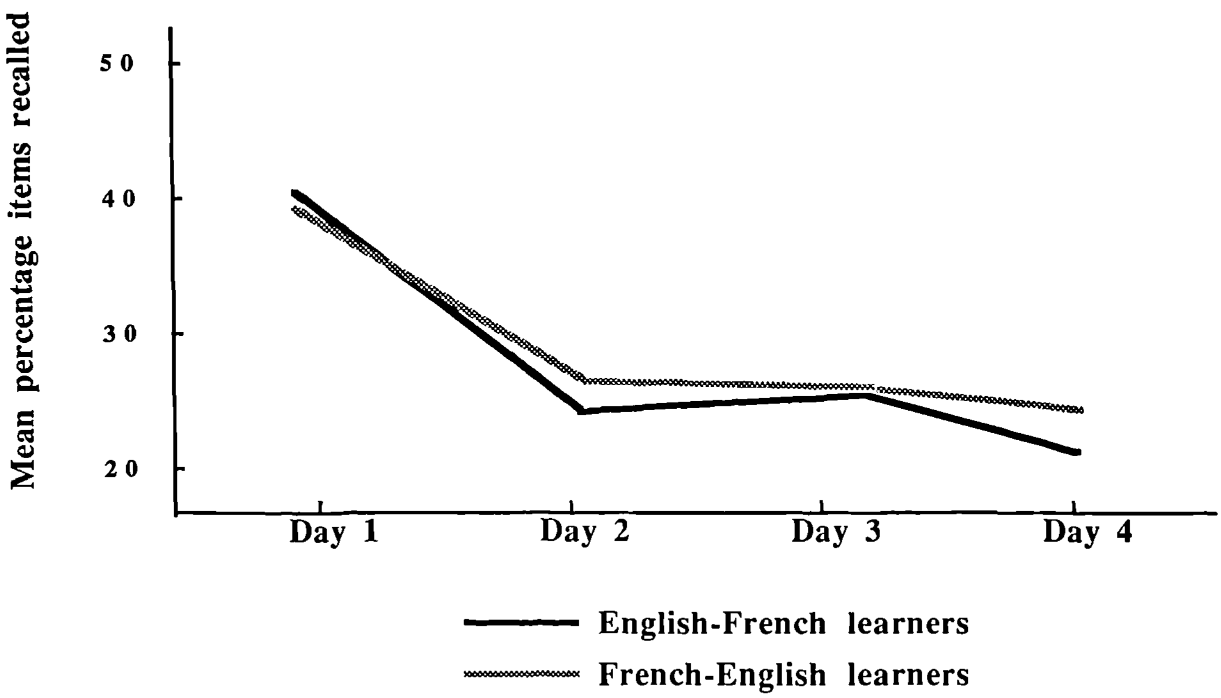

Figure 2.3. Experiment 1. School B1. Mean percentage scores for items recalled: Direction of learning, day of testing.

subjects. However, there was no significant interaction between day of testing and language-order at learning, F (3, 276) = 1.06, $p > 0.36$. Percentage means are contained in Table 2.5 (and see Figure 2.3). Language-order at learning had no significant effect on the strength of the bond over time measured by subjects' ability to recall items over the four test days.

Moving on now to the question of the more versatile direction of learning for both generation and comprehension, the position is more clear-cut than was the case for School A1. The more versatile direction for learning is English-French. Here, pairwise comparison (Tukey test) shows that the decrement for French-English learners tested in the reverse direction is 38.83%, (the difference between the performance of Group 2, 48.97% and Group 4, 10.14%). This difference, as Table 2.4. shows, (and see Figure 2.4), is

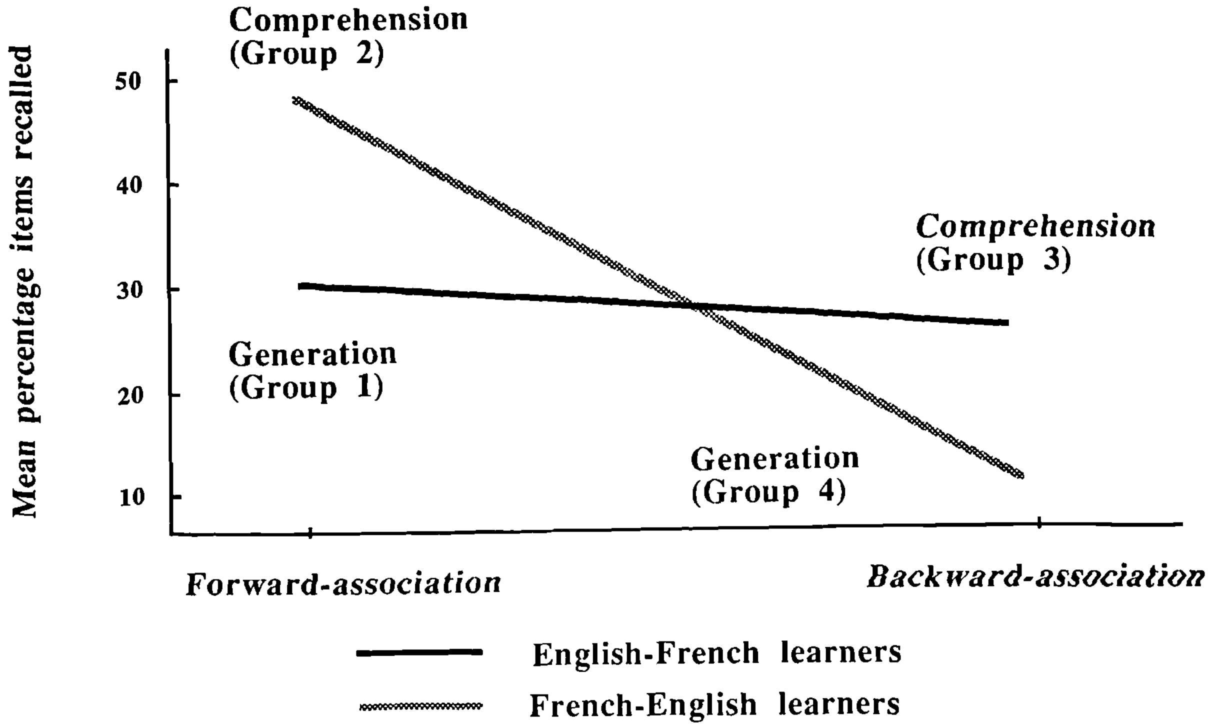

Figure 2.4. Experiment 1. School B1. Mean percentage scores for items recalled: Use of forward- or backward-association, direction of learning, test condition.

significant at $p < 0.01$. For English-French learners, on the other hand, the difference between Group 1, 30.66%, and Group 3, 26.39%, was not significant.

It seems to be the case that, where lower-ability learners are concerned, the relative ease of comprehension over generation comes much more into play. Thus although the French-English learners in Group 2 (48.97%) were significantly more successful than the English-French learners in Group 1 (30.66%), $p < 0.01$, the French-English learners in Group 4 (10.14%) were significantly less successful than the English-French learners in Group 3 (26.39%), $p < 0.01$. Assuming that subjects in Group 2 and Group 4 learned equally well initially, (they both learned French-English), then the difference in performance between them must be accounted for by the combination of the use of the backward-association and the more difficult task. When faced with the more difficult generation task, English-French learners, using the forward-association, are better equipped to cope than are French-English learners. For English-French learners, the relative ease of the comprehension task offsets any disadvantage from the relative weakness of using the backward-association in the comprehension test condition. For these lower-ability learners, therefore, learning in the direction English-French would be equally appropriate for generation and comprehension. Learning French-English would mean that learners would have great difficulty with generation of the foreign language items.

Conclusion

In terms of the discussion about word association, it is clear from these results that, as Wolford (1971) argued, the relationship of associations between two items in a word-pair is not symmetrical. The forward A-B association, whether English-French or French-English, is stronger than the backward-association and the item originally in the stimulus position is a more effective cue for the word-pair than the item originally in the response position. On the other hand, it is also clear that a substantial backward-association is formed between items in a word-pair and this is irrespective of whether the known item (the L1 word) or the unknown item (the L2 word) is in the stimulus position. As far as the English-French or French-English bond is concerned, there is finally no difference between the strength of the word-pair bond for English-French learners and for French-English learners.

Comprehension, that is responding L2-L1 irrespective of the order of learning, is an easier task than generation, responding L1-L2. It is not clear from this experiment whether this is due to a general facilitating effect of working from the less known (the L2 item) to the more known (the L1 item) or whether it is a specifically linguistic phenomenon.

In view of the inconsistency of results obtained in other domains, as discussed, it is clear that the issue cannot be resolved without reference to the kind of task involved.

Turning now to the implications of these findings for the present domain, it seems that for subjects in both schools, and irrespective of the overall differences in performance between the two schools, learning in the direction English-French is on balance the better all-purpose direction. The English-French association is more effective than the French-English association for the more difficult generation task. The weaker backward-association, therefore, is needed only for the less difficult comprehension task. Where second language learning practice is concerned, the matter receives little if any attention although it would seem to be fundamental to the whole learning process, particularly where lower-ability learners are concerned, and the motivational implications need further investigation.

CHAPTER 3

The transferability of list learning to testing in a simple context

Review of the second language learning literature on context effects

There is, in the second language learning literature, a strong, though not unanimous, view that learning word-pairs in a list is an ineffective way of learning second language vocabulary items and that learning in a context is to be preferred. There are two sorts of argument against the use of list learning. If it is assumed that list learning in effect means learning L1-L2 (or L2-L1) word-pairs, then there is an argument touched on in Chapter 1 to the effect that there is more to learning the meaning of a word than learning its denotation (Balhouq, 1976; Meara, 1980; Nation, 1987; J. C. Richards, 1976; Wallace, 1982). It would be difficult to argue with this point of view though it should be noted that it does not, of necessity, preclude the use of list learning as a first stage in the learning process. The second sort of argument is based on the notion that list learning is difficult and ineffective. This was the view of Hill (1965) who felt that because there is no one-to-one translation across language, time spent on word-pair learning is wasted time; of Hughes (1968) who claimed that list learning is soon forgotten; of Judd (1978) who claimed that "most people agree that vocabulary should be taught in context" otherwise it is not retained (p. 135); and of Turner (1983) who claimed of list learning that normally "only those students with unusually good memories can master the isolated words" (p. 81).

On the other hand, many writers make a case for the simultaneous presentation of L1 and L2 equivalent items and, in that sense, for word-pair learning. Seibert (1930) found word-pair learning to be superior to various forms of what she calls context learning. Kopstein and Roshal (1955) found an advantage for the simultaneous presentation of both parts of a word-pair. Mishima (1966) supported the use of translation and Lado, Baldwin, and Lobo (1967) identified simultaneous presentation of L1 and L2 as crucial to the learning process. Smith (1969) while not fully approving list learning accepted its efficiency for learners who do not have unlimited learning time. Green (1970) argued that:

"Where other presentation techniques seem likely to waste time or cause confusion, the most sensible solution is to give a vernacular explanation or approximate equivalent" (p. 219). George (1972) argued that because the system is adapted to the search for meaning, the orthographic form of the target word is often ignored; presentation of words out of context avoids this problem. Oskarsson (1975) found bilingual presentation to be more effective than monolingual means. Nilsen (1976), quoted by Judd (1978) as being against word-pair learning, in fact argued *for* what he calls paradigmatic learning before syntagmatic meaning. Cornu (1979) while accepting that translation can be misleading saw it as bringing security to the learner. Strick (1980) argued that learners use L1 initially in the learning process. Nation (1982) said that experimental evidence shows that simultaneous presentation is effective for a first encounter. Nation (1987) saw a place for list learning, particularly for high frequency words. Ostyn and Godin (1985) made the perhaps obvious but telling point that logically it is necessary for learners to understand before they can generate language and that banning L1 from the process would be ineffective even if it could be enforced. Bialystock (1985) argued for concentration on orthographic word form. Nation (1987), despite his reservations about the limitations of word-pair learning, pointed out that list learning saves time, can cover all categories of word (unlike pictures), and is easy to test. He concluded that not to use L1 in this way entails losing one way of encoding new vocabulary and that word-pair learning is better than learning by synonym or definition. Drum and Konopak (1987) suggested that learners supply definitions for words in isolation in any case and that the meaning of those words is filled out over time. Pressley, Levin, and McDaniel (1987) also argued that learners make lists anyway even when this is not allowed for in the "method" being used. There is a very large body of literature supporting the effectiveness of the Keyword Method which is based on a strategy of elaborating word-pairs (Atkinson, 1975; Bellezza, 1983; Paivio, 1983; Paivio & Desrochers, 1981; Pressley, Levin, & Delaney, 1982) although the Keyword Method does seem to be more effective with learning for comprehension than with learning for generation (Meara, 1980; Pressley, Levin, Hall, Miller, & Berry, 1980). Nagy and Herman (1987) argued for the value of any meaningful encounter with a word, even if the information gained from that one encounter is relatively small. Vaid (1988) found the use of translation to be effective. Hamers and Blanc (1989), Paivio (1971), and Paivio and Lambert (1981) saw dual coding by language as an aid to remembering.

In summary, then, despite the rejection by many of word-pair list learning, it is difficult to find convincing evidence to support this rejection; on the other hand, there does appear to be experimental evidence in favour of this form of learning under certain conditions. An alternative to list learning is learning in context and the arguments for and against this procedure are considered in Chapter 4. For the present, if it could be shown that list learning does not transfer well to testing in a simple context, this would seem to be a fundamental reason for devising some alternative way of presenting vocabulary items, always assuming that a better alternative exists and that learners could be persuaded to use it.

Review of the psychological literature on context effects

What is in question here is the ability of a word learned in a list of word-pairs to cue an appropriate response in a dissimilar test condition. In the psychological domain, the question is addressed in terms of the amount of overlap which is required between the learning condition and the testing condition if effective priming and cueing are to take place. The psychological discussion of this issue is extensive. It will be argued that although a strong version of encoding specificity cannot be sustained, there is an element of context-dependent information encoded at learning which can affect the probability of recall.

Encoding specificity and single word priming

An item x is said to cue or prime a target y when the presence of x facilitates the recall of y. "Priming" and "cueing" often seem to be used synonymously in the literature but here for the sake of clarity "priming" is taken to mean that kind of facilitation which is studied through reaction-time testing, and "cueing" is taken to mean that kind of facilitation which is more normally discussed in terms of memory tasks. Despite the distinction, it could be argued that priming and cueing are simply different ways of measuring the same phenomenon; that is, the reduction of the threshold of activation of an item in memory which makes its retrieval more likely. Priming and cued recall, as concepts, sit easily within a model of memory based on spreading activation. If the storage of information, including verbal information, is envisaged as being in a highly related and connectionist network (e.g., Anderson, 1983; Anderson, 1990; Collins & Loftus, 1975), then activation of a particular lexical representation will spread to related representations. Thus if the representations for the two halves of an L1-L2 word-pair have been related to each other at learning, one half of the word-pair can increase the activation of the other half of the word-pair, thus making its recall more likely. From Experiment 1 it is clear that the presence of one half of a word-pair can lead to successful retrieval of the other half of that word-pair and it will be assumed that priming and cueing have to some extent taken place. The question being asked here is whether a single-word component of a learned word-pair can function successfully as a recall cue when it is embedded, at testing, in a simple context. In outline, the issue is concerned with whether closeness of match between the learning situation and the test situation is necessary for recall; what emerges is that there is a range of factors which can affect retrieval among which task demands are of particular importance.

The argument against the effectiveness of a cue when there is a substantial difference between the learning and testing condition was originally formulated in terms of cueing and encoding specificity. It was particularly associated with Tulving (e.g., Thomson & Tulving, 1970; Tulving, 1974; Tulving & Osler, 1968; Tulving & Psotka, 1971; Tulving & Thomson, 1973). It has more recently reappeared as a debate about repetition priming effects where encoding specificity (in slightly different form) is defended in the work of Feustel, Shiffrin, and Salasoo (1983), Jacoby (1983a, 1983b), and Kirsner, Dunn, and Standen (1987), and is opposed to logogen-based explanations for repetition priming effects, which derive mainly from Morton (1969, 1970), for reasons to be discussed below.

In its original form, the encoding specificity principle (Tulving, 1974, insisted that it was a principle rather than a hypothesis) stated that for A to be effective as a cue for B, A must be encoded as a cue for B when B is learned. Tulving and Osler (1968) showed that weak associates effectively aid recall if presented with the target at both learning and recall but have no effect if presented only at recall. They also showed that there is no advantage to be gained by presenting one cue at learning and one at recall, or by presenting two cues rather than one. The argument of Tulving and Osler was that the word used as a cue does not enable recall due to any property it might have as a word, but only due to its having been processed as a cue in that particular learning episode.

The strong version of the encoding specificity principle cannot be sustained for empirical and theoretical reasons. Bahrick (1969, 1970) showed that cues of varying associative levels used at the time of testing but not encoded at the time of learning do cue target items differentially. It follows that the possibility of a particular word being an effective cue is not just a function of its being encoded as a cue at learning. Other factors must be involved, and notably how contiguous the cue is to the target in associative terms. Reder, Anderson, and Bjork (1974) showed that encoding specificity does not apply when target words of low frequency are used. In this case, cues which are highly associated with the low-frequency target are effective whether or not they are presented at learning. A more fundamental objection to the principle is that it is finally untestable in the manner undertaken by Tulving (passim). It has been pointed out by Battig (1968), Crothers and Suppes (1967), Jenkins (1979), Kolers (1979), and Madigan (1969) among others that the experimenter has no way of knowing that the item as given is the same as the item as remembered. Subjects adopt strategies appropriate to the perceived task and these strategies may be more or less successful. Remembering an episode, for example remembering a list of words in a given order, clearly requires that cues and targets be closely related to each other at encoding so that that specific episode can be recalled. A semantic memory task, on the other hand, is somewhat less rigorous in its demands because it is essentially abstract and concerned only with the meaning of items. Thus one can easily recall the word for "a fear of enclosed spaces" without being able to remember the exact circumstances (the episode) in which one first encountered the word.

Bahrick (1979) argued that if the task is specific and episodic, as the tasks set by Tulving tended to be, an encoding involving distinctiveness is likely to be effective. If the task is less specific, then an approach which is more flexible and which gives a variety of opportunities for retrieval will be more appropriate. Whereas encoding specificity can be shown to be an important factor in direct episodic retrieval, the principle simply does not apply to indirect retrieval where overlap between encoding and retrieval conditions rather than encoding distinctiveness is more important (Bahrick, 1979). Encoding distinctiveness can contribute to the retrieval process but is not necessary for it.

The more recent, and weaker, version of encoding specificity is to be found in the context of the debate on repetition priming (see Monsell, 1985, for a review). Repetition priming means that when a word is encountered twice in an experiment, then response to the word on the second occasion is faster and more accurate than is the response to words encountered for the first time. The debate is about what constitutes a repetition, in this context, and why that repetition is effective. Models of the language process which envisage the lexical unit as being an abstract representation of the word have no particular problems with the idea of variability between the learning and testing condition. However, models of the process which envisage the lexical unit as an episodic record appear to be committed to some extent at least to the idea of encoding specificity. If the lexical unit is an episodic record (e.g., Feustel, Shiffrin, & Salasoo, 1983; Jacoby, 1983a, 1983b; Kirsner & Dunn, 1985; Kirsner, Dunn, & Standen, 1987), there is no abstract representation of a word but records or instances of each specific encounter with that word. Repetition priming within this model occurs, therefore, only when the new episode is sufficiently similar to the original learning episode that matching is rapidly accomplished and reaction time is reduced.

The crucial difference between this model and logogen-based models, for example, is that the meaning of the word having been accessed, the codes which have enabled access are not discarded in favour of an abstract representation; on the contrary, it is just these codes which are the representation. According to Kirsner, Dunn, and Standen (1987) the incoming description is matched with existing records; if, as they put it, "contact" is not made immediately (by some process not specified) further parsing and re-description are carried out until recognition is achieved. If even then a match is not found, codes can be discarded until the best structural match is found. Codes are organised hierarchically, with morphological codes high in the hierarchy and surface characteristics lower in the hierarchy; it is their contention that surface characteristics will be the first to be discarded. The crucial factor in priming is, then, similarity. Performance will be enhanced by repetition to the extent that the repetition description matches the recently established record. Performance is highly sensitive to all the stimulus properties in both the prime and the repetition.

Kirsner, Dunn, and Standen (1987), in support of the instance model, claimed that for word identification it is similarity between learning and testing which is critical. Without similarity there is no priming. They summarised evidence "based on all studies of that type known to the authors" (Kirsner, Dunn, & Standen, 1987, p.152) which is taken to show that changes between upper-case and lower-case presentation, and changes between a male voice and a female voice, have a negligible inhibitory effect on transfer, that is, the ability of an item to prime a target. At the other extreme, morphologically unrelated translations show no priming effect at all. Changes in modality are somewhere in between. They do not mention cueing directly but claim that the record-based model has "general implications for perception and memory".

It is clear, however, that this claim is too strong. First, although it may be the case that morphologically unrelated translations show no priming effect in the sense of reducing lexical access time, word-pair learning does take place and a morphologically unrelated component of a word-pair can cue response to the other element of the word-pair (see Experiment 1). There is no reason to suppose that other changes between learning and testing conditions will eliminate cued recall. Second, even where priming is concerned, the sources quoted for the data on which they base their conclusions are not as clear-cut as they suggest. Cristoffanini, Kirsner, and Milech (1986) showed that non-cognates (words in two languages which though similar in meaning are not similar in orthographic form) do not prime under certain conditions and this is taken to be a function of the morphology of the items concerned. However, Kirsner, Dunn, and Standen (1987) also referred to Kirsner, Smith, Lockhart, King, and Jain (1984) where cross-language priming effects are obtained, as is the case in Meyer and Ruddy (1974). Another source quoted is that of Scarborough, Gerard, and Cortese (1984) but here too the possibility of cross-language priming under certain circumstances is accepted. It seems somewhat premature to talk of a zero effect of cross-language priming in general terms under these circumstances and, indeed, the difficulty of generalising about priming effects because of task dependencies and individual strategies is a point made in the rest of the Kirsner, Dunn, and Standen (1987) study.

The case for the possibility of effective cueing where learning and testing conditions are different was well reviewed by Monsell (1985). The argument is conducted on two levels. First, it can be shown that single word priming is not necessarily dependent on significant overlap between the learning and test condition. Second, it can be shown that a target word can be primed by a sentence context even when the sentence has not been encountered at learning. In an interactive model of the language process lexical representations receive information from a variety of sources including detectors at feature level, letter level, word level, syntactic level, word-sense level, and scenario level (see Elman & McClelland, 1986; Marslen-Wilson, 1989; McClelland, 1985, 1986, 1987; McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Seidenberg, 1989; Seidenberg & McClelland, 1989). Activations of logogens and connections between logogens are continuously graded rather than all-or-nothing and the degree of activation represents the degree of confidence the system has that a particular unit is the appropriate representation for what is being perceived. Each processing unit computes the difference between the excitation it is receiving and the inhibition it is receiving; its level of activation represents the result of this computation. The activation process is non-linear; at high and low levels of excitatory and inhibitory input, activation levels off. This means that factors which have an important effect on processing in some circumstances have little observable effect on other occasions. Finally, activation builds up and decays over time. Activation is gradual and incremental; decay is also gradual.

A main feature of logogen-based models is that the logogen is abstract in relation to any particular episode, rather than a record of an episode, but its activation is open to contextual information at each level. At an early level, for example, there is a word-superiority effect which demonstrates the influence of a lexical context on letter recognition. At a later level, words are recognised better when they occur in sentences. Words are disambiguated on the basis of their sentential context. Words are parsed on the basis of their immediate context. The system is highly interactive; through a process of spreading activation it makes potentially relevant information available in a parallel and automatic manner; it combines top-down and bottom-up processing in the most effective way possible relevant to the task in hand. It is a system which is "interested" above all in the message rather than in the medium. In other words, it is looking for the information, from whatever source, which will enable it to decode the message.

If lexical representations are indeed abstract in relation to particular episodes, then priming effects should transfer across tasks and across learning and testing conditions. If the effect is due to the recovery of a specific episode, then the effect should be highly context-dependent and again would support encoding specificity. In fact, the evidence points to clear transference of priming effects across task and context within the same modality. Thus priming effects have been shown for naming tasks when the prime has been: pronunciation of the opposite word to the prime (Clarke & Morton, 1983; Jacoby, 1983b); rhyme judgements and definition matching (Jacoby & Dallas, 1981); prior lexical decision (Monsell, 1985). Priming effects for identifying spoken words with a degraded stimulus have been shown by Kempley and Morton (1982). Visual lexical decision tasks have been primed by: naming (Scarborough, Cortese, & Scarborough, 1977; Scarborough, Gerard, & Cortese, 1979); reading the word in a sentence (Monsell, 1985); and sentence completion (Monsell, 1985). There is good evidence, then, for cross-task facilitation though, as Monsell pointed out, this does not mean that there is not a task-specific component to the priming effect but this will be in addition to the task-independent component. It seems to be the case that lexically "shallow" tasks such as lexical decision and naming do transfer across tasks; lexically "deep" tasks involve an additional task-dependent element. There is therefore a task-independent and long lasting component established at the point of lexical identification, in addition to a task-dependent component to priming which is effective at a later stage in processing. At learning, some attributes are included in the lexical representation which are not of the essence, but which are not "filtered out" in the process of abstraction. If these attributes are encountered on another occasion they contribute to the activation effect though they are not crucial to it. Thus the lexical unit is abstract in terms of any given context though it may incorporate these "accidental" attributes from a range of contexts. The claim is that this account clearly differs fundamentally from the Jacoby notion of separate records for each encounter but deals with the data just as effectively and has the added advantage of being wholly compatible with wider issues of interactive processing and memory performance.

Sentence priming

The evidence discussed to this point has been concerned with single word priming effects. Its relevance is that it shows that priming and cueing effects are not entirely record-based. A further question, however, is how these effects relate specifically to effects which derive from the provision of a sentence context at testing.

Given what has just been said about priming, it is self-evidently true that a word needed to complete a sentence will be primed by semantically or orthographically associated words which may happen to be present in that same sentence (although it is not clear, as has been seen, to what extent and in what circumstances this effect normally transfers across non-cognates in different languages). However, what is being examined here is rather different; this is the facilitation of word recall in sentences which do not provide an associative context. Lack of associative context is probably true of the majority of sentences, and in any case Foss (1982) showed that even when there is association between words in a sentence, the facilitatory effect on a target word is of longer duration than normal associative effects so that something more than simple association must be taking place.

That context should have a role in the understanding of individual words in sentences makes intuitive sense since it is context alone which allows for the determination of referents of pronouns and deictic terms, the form class of words such as *sleep* and *shovel* which can function as either nouns or verbs, and the appropriate reading of polysemous words such as *table* and *letter* (Gough, Alford, and Holley-Wilcox, 1981). In a much more general sense, however, the presence of a context speeds word recognition. Garnham (1985) gave the example of a sentence such as:

On the train the man read his **paper.**

Here the target word has no obvious associations with any other word in the sentence, yet the context will speed its recognition compared with the recognition of the same word without a context. Becker and Killion (1977), Fischler and Bloom (1979), and Schuberth and Eimas (1977) gave further examples of this kind of experimental result. The strength of the effect, however, is to some extent dependent on congruity and this could have important implications for tasks such as generation and comprehension. Morton and Long (1976) showed in a phoneme monitoring experiment that a target phoneme in a probable word in a sentence is more quickly accessed than the same target in an improbable word in that sentence. Goodman (1976), Morton (1964), Schuberth and Eimas (1977), Smith (1971), Stanovich and West (1979), and West and Stanovich (1978) all showed, in broad terms, that a congruous context will enhance the speed of processing of words within that context. On the other hand, sentence contexts can be inhibitory under certain circumstances. They can produce misrecognition (Goodman, 1976; Kolers, 1970; Weber, 1970). Incongruous contexts can inhibit word recognition (Stanovich & West, 1979; West & Stanovich, 1978). Efforts have been made to specify further the effects of congruity. Fischler and Bloom (1979) distinguished between probable sentence completions and congruous sentence completions. They found that highly probable congruous sentence completions were enhanced by context; congruous but unlikely completions were not enhanced; and incongruous completions were inhibited. In other words, the complete set of congruous words was not enhanced, only those that were probable. Kleiman (1980) found a broader range of effects than Fischler and Bloom. Kleiman used the categories of predictable completions, targets associated with predictable completions but not themselves congruous, and unrelated completions. He found that predictable and congruous completions were enhanced to the greatest extent. However, facilitatory effects were also found for words related to predictable and congruous completions and for acceptable completions by unlikely words. A different kind of discrimination is to be found in the work of Stanovich and West (Stanovich & West, 1979; West & Stanovich, 1978). Whereas the effect of congruity does not change in relation to age and reading ability, the effect of incongruity is not found in adult skilled readers; this point will be returned to presently.

Explanations of contextual effects are made in the context of wider-ranging models of the language process (see Connine, 1990, for a discussion) and it has to be said that much of the experimental evidence is directed to specific models and, probably as a consequence of this, confusing if not contradictory. At one extreme is the heavily top-down explanation of Goodman (1976) and Smith (1971) which seems to claim that the context is all important and the role of the individual word is simply to confirm expectations. This could be called the "facilitation by prediction" model. It is clear that there is some relationship between predictability and readability (see Rubenstein and Aborn, 1958) but it is not clear how far the model can go by way of explanation. Shannon (1951), for example, argued that the average uncertainty of an English word in context is 2^8; that is, sentential context reduces the number of possibilities to 256 (see Gough, Alford, & Holley-Wilcox, 1981). Clearly this is an important restriction on the set of possible "next words", but Shannon had nothing to say about the next and crucial step which is the system's way of choosing between these possibilities. Gough, Alford, and Holley-Wilcox (1981) favoured the idea of location shift, which means that when an incongruous target word is encountered, a different location within the network is activated and the shifting of attention to that new location by the system takes time. The level of confidence they have in this idea can be judged by their admitting that they have no evidence for it and that its status is therefore that of a metaphor. Forster's (1976, 1979) search model and the verification model (Becker, 1976, 1979; Becker & Killion, 1977) both envisaged a two-stage process of multiple activation followed by search and verification.

In Marslen-Wilson's (1989) Cohort Model, which combines spreading activation, verification, and connectionism, all lexical access must combine perceptual or form-based information and context-based information because the speed of spoken word identification is such that it rules out identification based on acoustic-phonetic information alone (Marslen-Wilson, 1984, 1987). The system's selection of a particular lexical item is based on multiple access to lexical forms constrained by assessment of the contextual appropriateness of those forms. The process of selection is parallel rather than serial to accommodate the speed of lexical selection. It is, therefore, not so much that the unit for the individual word is primed by context, but the *system* is primed in the sense that once the perceptual information is available, context can amplify the cues contained in the signal. The amount of perceptual evidence for a particular candidate would be less for a congruous word than for an incongruous word and the recognition criteria would be affected by context.

There is a different but related explanation of the effect of context based on the notion of two kinds of activation taking place at different speeds (Levy, 1981; Stanovich & West, 1979; West & Stanovich, 1978) and thus making information available to the system differentially. Stanovich (1981) put the explanation into the context of a compensatory-interactive model of reading. An interactive model envisages multiple and continuous information sources such as feature, orthographic, lexical, syntactic, and semantic information. Interactive constraint takes place and any level can compensate for shortage of information from other levels. Children and poorer readers tend to rely on context information rather than perceptual information when the former is available. As Spoehr and Schuberth (1981) pointed out, there is evidence to suggest that the context effect is more marked in younger and poorer readers than it is in older and more skilled readers (e.g., Biemiller, 1977-78; Doehring, 1976; Perfetti, Bell, Hogoboam, & Goldman, 1977; Perfetti & Roth, 1981; Samuels, Begy, & Chen, 1975-76; Schvaneveldt, Ackerman, & Semlear, 1977; West & Stanovich, 1978). Samuels, Begy, and Chen (1975-76) and

Schvaneveldt, Ackerman, and Semlear (1977) showed that the magnitude of a single word context effect is inversely related to reading level in ten year olds. West and Stanovich (1978) argued that in younger and poorer readers automatic word recognition processes are slow, relative to their speed in adult skilled readers. Therefore the influence of the slower contextual effects is more marked in this group than it is in the adult group where the automatic processes are that much more advanced. Stanovich and West (1979) took the argument one stage further by suggesting that what happens in both children and poorer adult readers is that higher level compensatory mechanisms come into play when lower level mechanisms are, for whatever reason, inefficient. This would explain why poorer readers would rely more on contextual information (see also Crowder, 1982). It should be noted that this does not mean that skilled readers do not use context along with all the other sources of information. It does mean that the effects of the use of context are more marked in weaker readers since their other sources of information are relatively impoverished. Thus Frederiksen (1978) argued that whereas low-skilled readers are only able to use context when it is constraining, more skilled readers are able to use whatever information there is even in a relatively unconstrained context. It could be also that the sentence context may occupy the attention of learners at the expense of the retrieval task.

Context effects are not only dependent on age and skill, as discussed, but also on the quality of the stimulus. Many studies show that when the stimulus is degraded, then the semantic context effect is more marked (e.g., Becker & Killion, 1977; Forster, 1976; Massaro, Jones, Lipscomb, & Scholz, 1978; Meyer, Schvaneveldt & Ruddy, 1975; Sanford, Garrod, & Boyle, 1977; Stanovich & West, 1979). Two types of compensation can be identified. One is obligatory and is provided by spreading activation and in this connection it is interesting to note that non-consciousness of the prime can be advantageous where these automatic processes are concerned (see Fischler & Goodman, 1978; Humphreys, Evett, Quinlan, & Besner, 1987; Marcel, 1976). The other is optional and under conscious attention. Mitchell and Green (1978) argued that in normal situations obligatory compensation is more likely but this is true only of fluent readers. Thus although adult readers, as we have seen, do not experience inhibitory effects due to context, this can be induced either by using a degraded stimulus or by increasing the interval between the initial response and the onset of the new stimulus.

The kind of strategy adopted by the system is in part, then, dictated by the quality of the stimulus. Another factor appears to be the general "clarity" of the environment. Dell and Newman (1980) and Stanovich and West (1979) claimed that a confusable environment and a degraded stimulus both slow down processing in what seems like a reallocation of resources. Drewnowski and Healy (1977) gave support to this notion pointing out that function words tend not to be detected in appropriate settings but this is not the case in inappropriate settings or when mixed-case presentation is made. What seems to happen is a kind of "exception reporting" which triggers more extensive and consciously controlled processing whenever something unusual occurs.

The conclusion of Levy (1981) was that there is no single form of the context effect. If the situation is tightly constrained, then the sentence context automatically spreads activation and facilitation. Sometimes the context seems to facilitate processing where the input is degraded and this suggests a change from automatic bottom-up processing to a top-down strategy to assist stimulus perception. It is possible that the two forms of facilitation work in parallel (Broadbent & Broadbent, 1980; Kleiman, 1980) and something of this can be seen in the system's ability to detect error (Cole & Jakimik, 1978, 1980) and make error restorations (Marslen-Wilson & Welsh, 1978).

General context effects

A footnote to the discussion on context is the finding that internal and external context at learning have a significant effect on the process of recall. Godden and Baddeley (1975) carried out a well known experiment with divers which demonstrated that the external context, in this case learning on land or underwater, can have a strong effect on recall (though not on recognition). State-dependent retrieval is the equivalent effect in terms of internal context; various studies have shown that physiological and psychological states of subjects at learning enable the recall process when reproduced at testing (see Eich, 1980; Goodwin, Powell, Bremer, Hoine, & Stern, 1969; Kumar, Stolerman, & Steinberg, 1970). It could be the case that being tested in a list and being tested in a context are sufficiently different for performance to be impaired.

Conclusion

In summing up this discussion, it is worth taking account of some important points made in memory study over the years: the distinction between episodic memory and semantic memory; the importance of task demands; the difference between direct and indirect retrieval. On the matter of episodic and semantic memory, it could be said that these concepts are good examples of unfortunate reification. That is, instead of being convenient ways of referring to two kinds of data, the terms have come to refer to two kinds of entity. Nevertheless, there is a useful distinction to be made between two kinds of data; there is memory for episodes and memory for meaning. Memory for episodes requires attention to detail across a range of features; memory for meaning requires the ability to abstract from a particular episode and to relate what is abstracted to the network of stored meanings. In any particular instance, and with reference to a particular individual, it would appear to be impossible to designate "a memory" as belonging to either one category or the other.

It is clear also that the kind of processing undertaken by a subject may well be different according to anticipated task demands (see, for example, Hall, 1971; Kolers & Roediger, 1984; Perfetti, 1979). Broadly speaking, if subjects are expecting a task which will require recognition and discrimination among similar items, then they will need to pay attention to the details of the episodic encounter; if subjects are expecting a task which will require recall of a word's meaning, then the specifics of the encounter will not be to the point.

Finally, there is a useful distinction to be made between what Eysenck called "direct" and "indirect" retrieval (Eysenck, 1984, pp. 157 ff.) and what Jacoby and Craik (1979) called "spontaneous" and "directed" memory (see also LaBerge & Samuels, 1974; Neely, 1977; Schneider & Fisk, 1982; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977; Solso, 1974). Direct or spontaneous recall occurs when the cue, either specific or environmental, elicits an immediate response because the information sought is familiar and regularly used, and possibly because there is considerable overlap between cue information and target information. Indirect recall occurs when no immediate response is elicited; in this situation a process of reconstruction, possibly involving the generating and checking of likely candidates for an answer, given the nature of the contextual cue, can lead to recall. Retrieval, therefore, like encoding, can vary along the traditional dimensions of depth, elaboration, and distinctiveness in the sense that the reconstructive process can be conducted to a greater or lesser degree. The point is that episodic direct retrieval would be heavily dependent on overlap between the cue and the target whereas semantic indirect retrieval may have very little dependence on such an overlap. While a strong form of encoding specificity cannot be sustained, the difference between record-based accounts of cued recall and logogen-based accounts seems to be more one of emphasis than of substance. The record-based model allows for abstraction into codes, for best match, for re-description, and for discarding of codes; it is difficult to see what encoding specificity finally means when the original encoding context can be reshaped in what appears to be an *ad hoc* fashion. To this extent, the model seems to be merely a reformulation of logogen-based explanations in different terms, but without the appropriate theoretical underpinning.

What emerges from this discussion is that the relationship between the learning condition and the testing condition is not straightforward and is highly responsive to perceived task demands. Experiments 2 and 3 were designed to test the ability of words learned in word-pair lists to cue appropriate responses in a simple sentence context. The task is taken to be sufficiently different from the learning condition to test the possibility of transferability of list learning.

EXPERIMENT 2

Experiment 2 was designed to test whether list learning transferred well to testing in a simple sentence context when generation, response of an L2 item to an L1 cue, was the task. Subjects learned in a list but were required to generate French items in response to an English cue when both the cue word and the space for the target word were presented not in isolation but in English and French sentences respectively.

Although the cue word remains the same at learning and testing, the conditions are clearly different. Even if subjects attempt to treat the test as a list test, the presence of the English and French sentences will function as distractors to some extent. If subjects base their learning strategy on the expectation of a list test, then they may well undertake task-dependent or episodic learning which will not transfer well to testing in context (Feustel, Shiffrin, & Salasoo, 1983; Hall, 1971; Jacoby, 1983a, 1983b; Kirsner & Dunn, 1985; Kirsner, Dunn, & Standen, 1987; Kolers & Roediger, 1984; Postman & Schwartz, 1964). It is difficult to envisage any advantage accruing from the presence of the context, where generation is the task, unless learners have engaged in sufficiently elaborated processing that the English and French contexts help to prime the French target through a process of spreading activation (Bahrick, 1969, 1970; Reder, Anderson, & Bjork, 1974; interactive models passim). This appears to be unlikely in the time available for learning what are entirely new vocabulary items. Again, context only facilitates word-access when the context is congruous (Becker & Killion, 1977; Fischler & Bloom, 1979; Goodman, 1976; Morton, 1964; Morton & Long, 1976; Schuberth & Eimas, 1977; Smith, 1971; Stanovich & West, 1979; West & Stanovich, 1978). It is likely that subjects will be paying attention to the English sentence context in which case the French target may well be perceived as incongruous. Under these circumstances, inhibitory effects are possible (Fischler & Bloom, 1979; Goodman, 1976; Kolers, 1970; Stanovich & West, 1979; Weber, 1970; West & Stanovich, 1978). The expectation was, therefore, that subjects would receive no advantage from being tested in a sentence context and that they would not recall items as successfully as subjects tested in a list when list learning had taken place.

Method

Design

Having concluded in the previous chapter that English-French is the more adaptable direction of learning for second language vocabulary items, all learning in the present experiment and subsequent experiments was English-French. All subjects in Experiment 2 learned in an English-French list.

The experiment had a 2 x 2 factorial design. The between-subjects factor was the condition of testing, with two levels; one group was tested in a list, as in Experiment 1, and the other group was tested in a simple sentence context. The within-subjects factor was the time of testing, again with two levels; the first test took place on the same day as learning and the second test took place five days later. Although the passage of time seemed to have a consistent effect over conditions in Experiment 1, it was considered worthwhile to test over two days in the present experiments since the introduction of a context at testing was a significant departure from the previous paradigm.

Materials

Items were prepared in lists of 20 word-pairs. The same list was used for Group 1 and Group 2. The pattern of words was similar to that used in Experiment 1; there were four concrete and four abstract nouns, five verbs, seven others. Short-medium length words were chosen but these were not controlled for frequency since both groups learned the same list. In the list test condition, cue words on test papers were identical to the original, but the order in which they appeared was randomised; this was designed to reduce possible effects of list dependency. In the context test condition, for each item two sentences were given; an English sentence containing the cue word was presented above its French equivalent. The English cue word was highlighted and positioned directly above the slot to be filled. The cue words were identical to the original list words, and the (randomised) order of the list test condition was followed. (Full lists are contained in the Materials Appendix.)

The wardrobe is in the bedroom
.... est dans la chambre

He was **seated** under a tree.
Il était .... sous un arbre.

Young people are forbidden **to bet.**
Il est interdit aux jeunes de ....

Subjects

Subjects were taken from the same two schools as in the previous experiment; 46 subjects were used from School A2 (24 boys and 22 girls), and 38 from School B2. The pool of subjects was completely different from that used in the previous experiment.

Procedure

The procedure was as for Experiment 1, except for there being two days of testing rather than four. Five days after the first test a second test was carried out. This was not

Table 3.1. Experiment 2.

Arrangement of Groups.

Group	**Direction of learning**	**Direction of testing**	**Type of test**
1	English-French	English-French	List
2	English-French	English-French	Context

announced to subjects in advance in order to avoid relearning by subjects.

Two experimental groups were formed in each school. Group 1 learned in a list and was tested in a list. Group 2 learned in a list and was tested in a sentence context. In School A2, Group 1 comprised 21 pupils (11 boys, 10 girls), Group 2, 25 pupils (13 boys, 12 girls). In School B2, Group 1 comprised 19 pupils, Group 2, 19 pupils. The arrangement of the groups is as in Table 3.1.

The experiment took place in the Spring term of the school year. Subjects had had, therefore, six months of formal French teaching when they began the experiment.

Results and discussion

A preliminary analysis of variance revealed a significant difference in performance between School A2 and School B2, School B2 (45.98% correct responses) being more successful than School A2 (34.91% correct responses), F (1, 80) = 9.68, $p < 0.01$. There was also a significant interaction between School x Group, F (1, 80) = 15.64, $p < 0.01$. For these reasons, separate analyses were conducted on the data for the two schools. For the sake of consistency, results for the higher-ability school will be discussed first.

Results for School B2: Higher-ability subjects

An analysis of variance was performed. The between-subjects factor was testing, with two levels. Subjects in Group 1 were tested in a list and subjects in Group 2 were tested in a sentence context. The within-subjects factor was time; subjects were tested on the day of learning and again five days later. Mean percentage scores are contained in Table 3.2.

The main question being asked in this experiment was whether list learning transferred well to more "normal" testing circumstances. The use of a list or context at testing had a significant effect on performance, F_1 (1,36) = 47.10, $p < 0.01$, F_2 (1, 19) = 108.57, $p < 0.01$. The mean percentage score for subjects tested in a list, that is in a

Table 3.2. Experiment 2. School B2.

Mean percentage scores for items recalled.

Group	Day 1	Day 2	Mean
1. List test	74.21	53.94	64.07
2. Context test	36.57	19.21	27.89
Overall	55.39	36.57	45.98

condition similar to the learning condition, was 64.07% and for subjects tested in the sentence context was 27.89%. To that extent therefore, list testing did not transfer well.

It had been anticipated that the provision of a context at testing would not contribute positively to performance; in effect, the difference between the two groups suggests that the provision significantly impaired performance. Assuming that subjects in both groups learned equally well at outset, the performance in context was surprisingly poor and, as discussed previously, various explanations are possible for this performance. The most likely explanation would seem to be that subjects adopted a strategy for episodic learning which was highly successful for list testing but not well adapted to the unexpected test condition (Feustel, Shiffrin, & Salasoo, 1983; Hall, 1971; Jacoby, 1983a, 1983b; Kirsner & Dunn, 1985; Kirsner, Dunn, & Standen, 1987; Kolers & Roediger, 1984; Jensen, 1962; Lesgold & Bower, 1970; Postman & Schwartz, 1964). According to the encoding specificity principle, this would be because there was not sufficient overlap between the learning condition and the test condition. According to the logogen-based interactive explanation it could be because there was a high degree of task-specific information in the learning process which was not available, or not perceived, in the test condition. Another possibility would be that the performance of Group 2 is an example of state-dependent retrieval (see Eich, 1980; Goodwin, Powell, Bremer, Hoine, & Stern, 1969; Kumar, Stolerman, & Steinberg, 1970). Being tested in a list and being tested in a context are sufficiently different for performance to be impaired. A less likely explanation, given the overall level of competence of School B2, is that the sentence context led learners in Group 2 to a misidentification of the task. In other words, faced with sentence contexts rather than the expected lists, subjects undertook a recall task rather than a cued recall task thus failing to make full use of the word-pairs which they had learned. These explanations are not mutually exclusive and the conclusion must be that, for whatever reason or combination of reasons, list learning did not transfer well to the more "natural" test condition.

The day of testing also affected performance. The percentage mean of correct responses for Day 1 was 55.39, and for Day 2 was 36.57, F_1 (1, 36) = 93.28, $p < 0.01$, F_2 (1, 19) = 120.26 $p < 0.01$. As in Experiment 1, therefore, there was a significant decrease in performance after Day 1.

There was no significant interaction between the day of testing and the type of test undertaken, F (1, 36) = 0.55, $p > 0.46$. This suggests that the difference in performance between the two groups was due mainly to the test condition and that both groups learned equally well at the outset.

Results for School A2: Lower-ability subjects

An analysis of variance was again performed with the same between-subject and within-subject factors as for School B2. The between-subjects factor was testing, with two levels. Subjects in Group 1 were tested in a list and subjects in Group 2 were tested in a sentence context. The within-subjects factor was time; subjects were tested on the day of learning and again five days later. Mean percentage scores are contained in Table 3.3.

Day of testing had a significant effect on performance, F_1 (1,44) = 72.96, $p <$ 0.01, F_2 (1, 19) = 42.07, $p < 0.01$. The percentage mean of correct responses for Day 1

Table 3.3. Experiment 2. School A2.

Mean percentage scores for items recalled.

Group	Day 1	Day 2	Mean
1. List test	49.76	28.09	38.92
2. Context test	39.80	22.00	30.90
Overall	44.78	25.04	34.91

was 44.78%; for Day 2 it was 25.04%. The effect of time appears to have been comparable for both groups since there was no significant interaction between the day of testing and the type of test undertaken, $F\ (1, 44) = 0.70, p > 0.30$.

The most interesting feature of these results, however, is that the change from list learning to testing in context did not have any significant effect on performance, $F\ (1, 44) = 2.81, p > 0.1$. Group 1 had a mean percentage score of 38.92% correct items, and Group 2 had a mean percentage of 30.90% correct items. Despite the tendency in favour of Group 1, this difference was not significant.

Discussion

There appears to be something of a paradox in these results. Subjects from School B2, who learned more successfully initially, were less able to transfer that learning to being tested in a sentence context than subjects from School A2 who learned less successfully overall.

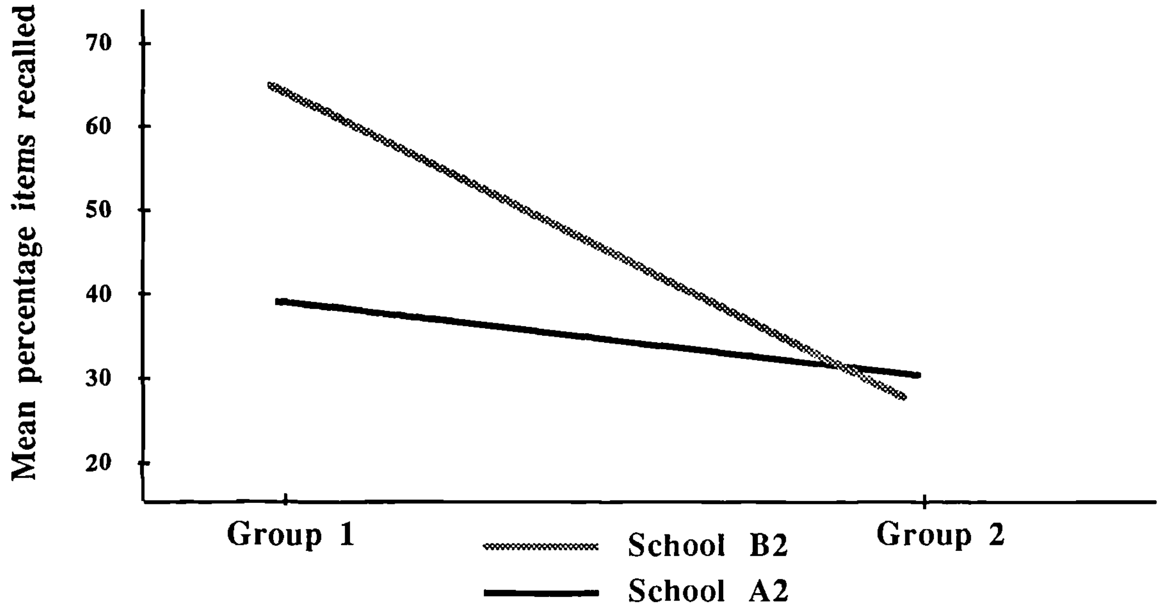

Figure 3.1. Experiment 2. Interaction: School x Group (defined by test condition). Group 1 was tested in a list and Group 2 in a sentence context.

An attempt was made to explore the source of this result by re-analysing the data with School as a factor. The analysis shows a significant interaction between School and Group, F (1, 80) = 15.64, $p < 0.01$. A pairwise comparison (Tukey test) shows that there was no significant difference between the performance of Group 2 in each case, $p > 0.05$; Group 2 in School B2 had a mean percentage score for correct items of 27.89%; Group 2 in School A2 had a mean percentage score of 30.90%. It would appear therefore that list learning transferred equally badly to testing in a sentence context for subjects in both schools. On the other hand, the difference in performance between Group 1 in School B2 (64.07%) and Group 1 in School A2 (38.92%) was significant, $p < 0.01$.

It could be that where School B2 is concerned, assuming that Group 1 and Group 2 learned equally well at the outset, the reason for the disparity between the performance of the groups is the difference between the learning condition and the test condition as has already been discussed. Group 1 were able to use the strategy well adapted to list learning and list testing; Group 2 were not able to use that strategy.

Where subjects in School A2 are concerned, Group 1 and Group 2 did not adopt a task-specific strategy with the result that Group 1 performed no better than Group 2 at testing despite the advantage of similarity of learning and test condition available to Group 1. If no task-specific strategy had been adopted at learning, the similarity of test condition to learning condition for those tested in a list would not be enabling.

As Figure 3.1. illustrates, it is not the case that the subjects with overall lower-ability were able to adapt more successfully to testing in a sentence context but that subjects with higher ability performed particularly well in the list testing condition.

EXPERIMENT 3

Experiment 3 was the reverse of Experiment 2 and was designed to test whether list learning transferred well to testing in a simple sentence context when comprehension, an L1 response to an L2 cue, was required. As in the previous experiment, there is a clear difference between the learning condition and the test condition but whereas where generation was concerned it was difficult to envisage a positive effect from the provision of a context at testing, where comprehension is concerned there seems to be the possibility of a trade-off between advantages offered by the context and the disadvantages associated with the encoding specificity effect. The main advantage is that the English context, through the priming effect of spreading activation, could direct a generation and selection process with greater accuracy than could be achieved in list testing (Bahrick, 1969, 1970; Reder, Anderson, & Bjork, 1974; interactive models passim). Again, since the target is congruous with the context, facilitation is possible (Becker & Killion, 1977; Fischler & Bloom, 1979; Goodman, 1976; Morton, 1964; Morton & Long, 1976; Schuberth & Eimas, 1977; Smith, 1971; Stanovich & West, 1979; West & Stanovich, 1978). Given that the target is a familiar word, and the sentence-context is meaningful, there is scope for what Neely (1977) called cognitive associations coming into play in the event of difficulty of retrieval (see Becker & Killion, 1977; Dell & Newman, 1980; Forster, 1976; Levy, 1981; Massaro, Jones, Lipscomb, & Scholz, 1978; Meyer, Schvaneveldt & Ruddy, 1975; Mitchell & Green, 1978; Sanford, Garrod, & Boyle, 1977; Stanovich, 1981; Stanovich & West, 1979; West & Stanovich, 1978). However, the influence of the difference between the learning and testing condition cannot be ignored and, in any case, any potential advantage is dependent on learners' ability to make use of the context and to exploit the opportunities offered. It might be expected, then, that if there were a difference between schools, higher-ability subjects would benefit more noticeably from the provision of a context at testing than lower-ability subjects.

Method

Design

The experiment had a 2 x 2 factorial design. The between-subjects factor was the condition of testing, with two levels; one group was tested in a list, as in Experiment 1, and the other group was tested in a simple sentence context. The within-subjects factor was the time of testing, again with two levels; the first test took place on the same day as learning and the second test took place five days later. All subjects learned in an English-French list and were tested in the direction French-English.

Materials

Materials were prepared as in the previous experiment with the exception that at testing the order of languages was reversed.

Subjects

Subjects were taken from the same two schools as in the previous experiment. No subjects participated in both experiments. There were 45 subjects (23 boys and 22 girls) from School A3, the mixed-sex school, and 41 subjects from School B3, the all-girls school.

Table 3.4. Experiment 3.

Arrangement of Groups.

Group	Direction of learning	Direction of testing	Type of test
1	English-French	French-English	List
2	English-French	French-English	Context

The experiment took place in the Summer term of the school year. Subjects had therefore had nine months of formal French teaching when the experiment took place.

Procedure

Two experimental groups were formed in each school. In School A3, Group 1 comprised 20 pupils (10 boys, 10 girls), Group 2, 25 pupils (13 boys, 12 girls). In School B3, Group 1 comprised 19 pupils, Group 2, 22 pupils. Group 1 learned in a list and was tested in a list. Group 2 learned in a list and was tested in a sentence context. The arrangement of the groups is shown in Table 3.4. In all other respects, the procedure was as in Experiment 2.

Results and discussion

A preliminary ANOVA with Schools as a factor revealed a significant difference in performance between School A3 and School B3, School B3 (54.39% correct responses) being more successful than School A3 (33.65% correct responses), F (1,82) = 43.89, p <

0.01. There were also significant interactions between School x Day, F (1,82) = 10.03, $p < 0.01$, and School x Group x Day, F (1,82) = 8.25, $p < 0.01$. Separate analyses were therefore conducted for the two schools.

Results for School B3: Higher-ability subjects

An analysis of variance was performed. The between-subjects factor was testing, with two levels. Subjects in Group 1 were tested in a list and subjects in Group 2 were tested in a simple sentence context. The within-subjects factor was time; subjects were tested on the day of learning and again five days later. *The mean scores are contained in Table 3. 5.*

As in Experiment 2, day of testing had a significant effect on performance. The percentage mean of correct responses for Day 1 was 62.82%, and for Day 2 was 45.97%. This difference was significant, F_1 (1,39) = 76.28, $p < 0.01$, F_2 (1, 19) = 24.07, $p < 0.01$.

However, the use of a list or context at testing, which was the main point of interest in this experiment, did not have a significant effect, F (1, 39) = 1.82, $p > 0.1$. Group 1, tested in a list, averaged 57.89% correct responses, and Group 2, tested in a context, averaged 50.9%. It appears that the relative ease of comprehension as a test, established in Experiment 1, is not enhanced by the provision of a context at testing. This result is surprising. If a single word cue is an effective cue in a comprehension task, then despite the change of direction between learning and testing, there is reason to expect that a congruous sentence will be an even more effective cue whether because of automatic spreading activation or because of cognitive association (Neely, 1977). However, there is possibly a useful analogy here with the performance of stronger learners in Experiment 2. If learners in Group 1 and Group 2 in School B3 adopted a task-dependent list learning strategy, then this strategy would be well adapted to being tested in a list. To that extent, subjects in Group 2 would be disadvantaged. However, subjects in Group 2 would have this disadvantage offset by the information available in the sentence context and this could account for their performance being comparable with that of Group 1 subjects who were

Table 3.5. Experiment 3. School B3.

Mean percentage scores for items recalled.

Group	Day 1	Day 2	Mean
1. List test	68.15	47.63	57.89
2. Context test	57.50	44.31	50.90
Overall	62.82	45.97	54.39

tested in a list. If this is so, the lack of a significant advantage for context testing would offer some support for the notion of encoding specificity, and would indicate the influence of task-dependent learning on subject performance.

In terms of second language list learning, for these stronger learners at least, list learning does transfer well to the more "natural" demands of a sentential context where comprehension is concerned. There is something of a trade-off between the disadvantage of differences between learning and testing conditions and the advantage offered by the sentence context.

Results for School A3: Lower-ability subjects

An analysis of variance was performed. The between-subjects factor was testing, with two levels. Subjects in Group 1 were tested in a list and subjects in Group 2 were tested in a simple sentence context. The within-subjects factor was time; subjects were tested on the day of learning and again five days later. Mean scores are contained in Table 3. 6.

For subjects in School A3 being tested in a list or in a context had a significant effect on performance. Subjects tested in a list, Group 1, averaged 37.50% correct responses whereas subjects tested in the unexpected context, Group 2, averaged 29.80%. This difference was significant, F_1 (1,43) = 4.35, $p < 0.05$, F_2 (1,19) = 5.25, $p < 0.05$. Again, this result is, to some extent, puzzling. If, as was argued in Experiment 2, lower-ability learners do not adopt any particular strategy in response to an expected list test, then there is no reason here either to expect subjects tested in a context to be more disadvantaged than subjects tested in a list. One possible explanation is that lower-ability learners are more influenced than higher-ability learners by context in the more general sense; that is, they are heavily influenced by the simple familiarity of doing certain things in a certain way. It is not so much a question of strategy as one of habit. Although in the generation task in Experiment 2 the change to testing in context for Group 2 was also unfamiliar, on that occasion the forward-association was used in the "normal" way. Here for Group 2 the comprehension task is unfamiliar in two ways; not only is testing in a context, but the weaker backward-association is being used. It could be that the combination of these two elements is responsible for the poor performance. It is a point made in the second language learning literature that the effective use of context is not necessarily intuitive and often needs to be taught; "unusual" learning and testing conditions may be inhibitory (see, for example, Bialystock, 1985; Nation & Coady, 1988; Sternberg, 1987; Turner, 1983). This might be the case here.

Performance was significantly affected by the day of testing. The average number of items recalled for Day 1 was 37.87% and for Day 2 it was 29.42%. This difference was significant, F_1 (1,43) = 21.48, $p < 0.01$, F_2 (1, 19) = 8.43, $p < 0.01$. There was also an interaction between manner of testing and day of testing, F_1 (1,43) = 4.69, $p < 0.05$, although the result must be treated with caution because the interaction by item is not significant, F_2 (1, 19) = 0.35, $p > 0.10$. A simple effects analysis shows that the interaction relates to the performance of Group 2 on Day 2, F (1,43) = 218.59, $p < 0.05$, which might suggest that for lower-ability list learners the adverse effects of a change of test condition increase over time.

Table 3.6. Experiment 3. School A3.

Mean percentage scores for items recalled.

Group	Day 1	Day 2	Mean
1. List test	39.75	35.25	37.50
2. Context test	36.00	23.60	29.80
Overall	37.87	29.42	33.65

General discussion and conclusion

The data from Experiment 2 and 3 are inconclusive but this in itself constitutes a contribution to the debate about transferability in both the second language learning and the psychological domain. Some support is given to the notion of encoding specificity but a strong version of encoding specificity (Thomson & Tulving, 1970; Tulving, 1974; Tulving & Osler, 1968; Tulving & Thomson, 1973) cannot be sustained because its effect is not manifest on all occasions. Under certain circumstances, lack of similarity between learning condition and testing condition does not inhibit recall; this can be seen in School A2 in Experiment 2 and in School B3 in Experiment 3. Encoding specificity appears to be more related to different learning styles than to some quality inherent in a mental record of an episode as Jacoby and Kirsner would have it (Jacoby, 1983a, 1983b; Kirsner & Dunn, 1985; Kirsner, Dunn, & Standen, 1987). In this sense it is probably better described in terms of task-related behaviour (see Monsell, 1985).

Higher-ability learners appear to be more responsive to learning and testing conditions than lower-ability learners and to be more ready to develop what they consider to be appropriate strategies. Where generation is concerned, the case of School B2 in Experiment 2 shows that presentation is particularly important for higher-ability learners since it is used by them as an indicator or the type of strategy to be adopted. In the comprehension task, the disadvantage of the change from the expected task is offset by their exploitation of the context provided; this could be as a result of either automatic processes (Bahrick, 1969, 1970; Becker & Killion, 1977; Fischler & Bloom, 1979; Goodman, 1976; Morton, 1964; Morton & Long, 1976; Reder, Anderson, & Bjork, 1974; Schuberth & Eimas, 1977; Smith, 1971; Stanovich & West, 1979; West & Stanovich, 1978) or conscious processes (Becker & Killion, 1977; Dell & Newman, 1980; Eysenck, 1984; Forster, 1976; Jacoby & Craik, 1979; Levy, 1981; Massaro, Jones, Lipscomb, & Scholz, 1978; Meyer, Schvaneveldt & Ruddy, 1975; Mitchell & Green, 1978; Neely, 1977; Sanford, Garrod, & Boyle, 1977; Stanovich, 1981; Stanovich & West, 1979; West & Stanovich, 1978). Nevertheless, it might be expected that the full advantage of the use of context would only be seen if subject strategy were not "misled" by the learning condition. This matter will be considered in the following chapter when a comparison will be made between learning in a list and learning in a context when the testing itself, as here, is in a context.

Lower-ability learners appear to have been impaired not so much by task-dependent behaviour as by a more general dependence on what might be called "familiarity" - the performing of familiar tasks in familiar ways (see Eich, 1980; Goodwin, Powell, Bremer, Hoine, & Stern, 1969; Kumar, Stolerman, & Steinberg, 1970). There is no evidence to suggest that a particular list learning strategy is undertaken by subjects in School A2 in Experiment 2 because subjects tested in a list did no better than subjects tested in a context. Although subjects tested in a list in School A3 in Experiment 3 performed more successfully than subjects tested in a context, it could be that the difference was not due to a successful strategy adopted by subjects tested in a list, but due to the poor performance of subjects tested in a context who had to cope with using the backward-association and the unexpected task.

The conclusion would seem to be that learning involves a somewhat unpredictable combination of task-independent and task-dependent elements and that the combination reacts to subject abilities and subject strategies.

In terms of second language learning, the important conclusion is that higher-ability learners are more likely to be disadvantaged by list learning where generation in a sentence is concerned and lower-ability learners are more likely to be disadvantaged by list learning where comprehension in a sentence is concerned. However, list learning is not to be discounted for this reason. First, because it is relatively successful for the reverse cases (comprehension for higher-ability learners; generation for lower-ability learners). Second, because it cannot be assumed that a better alternative is available (or that it would be used if available); this topic will be addressed in the next chapter.

CHAPTER 4

Learning in a list versus learning in a context

If there is, in the second language learning literature, a strong sense that learning word-pairs in a list is not an effective way of learning second language vocabulary items, then there is an equally strong sense that learning in a context is an effective procedure. Although the main purpose of this thesis is to examine aspects of list learning, it is worthwhile finding out whether the preference for learning in a context has any empirical basis if only for the reason that if learning in a context is not more effective than learning in a list, list learning could be accorded more attention than is at present the case.

A review of the second language learning literature

Briones (1937) argued that a context provides more word associations than a list. Carroll (1963), without necessarily agreeing with the practice, claimed that foreign language teachers favour the use of context. Hughes (1968) saw context learning without any L1 information as "natural". In similar vein, Wind and Davidson (1969) said that learning in context is "normal", leads to better processing, and provides syntactic cues as well as semantic redundancy. Judd (1978) claimed that "most people agree that vocabulary should be taught in context" (p. 135) otherwise vocabulary is not retained and the full meaning of the word is not learned. Taylor (1983) made a case for learning words in collocations (phrases such as "bed and breakfast", for example) since words naturally associated are more easily learned. Wallace (1982) insisted that vocabulary should be taught in what he calls a natural environment. Allen (1983) felt that a sentence context gives the opportunity for learners to infer the meaning of unknown words; it is not clear how this could apply to generation rather than comprehension.

However widespread the practice of teaching vocabulary in context might be, it is difficult to find hard evidence to justify that practice. It is easy therefore to have sympathy with Pressley, Levin, and McDaniel (1987) who confessed themselves "puzzled at the positive regard afforded to students' use of external context cues as a vocabulary-remembering strategy" (p. 119); and with the conclusion of Nation (1987) that however attractive the idea of the efficacy of context for learning, this is a "statement of belief" (p. 137) rather than a principled judgement.

There are two main obstacles to the acceptance of the notion that the provision of a context at learning is a more effective means of vocabulary acquisition than word-pair learning. The first is the fact that "context" as a concept is in its present state undefined and unusable. The second is that however it *is* defined, there is little evidence to suggest that it is effective. On the question of definition, Nation (1987) pointed out that for Seibert (1930) context was a defining sentence with the L1 equivalent of the target word in brackets next to the target word; Holley (1973), Holley and King (1971), Morgan and Bailey (1943), and Morgan and Foltz (1944) used stories with various forms of glossary provided; Lado, Baldwin, and Lobo (1967) defined context as anything added to the word-pair, and this would include the use of cognates, derivatives, or a story. Adding to the confusion is the fact that there appears to have been little control over subjects' understanding of the context which was meant to lead to learning (Carroll, 1966; Gershman, 1970). Clearly, it is impossible in these circumstances to draw firm conclusions about the role of "context".

In any case, such evidence as there is appears not to support the notion of the effectiveness of learning in context. Seibert (1930) found learning in context inferior to learning in word-pairs. She argued that the task of learning word-pairs is less demanding since the link to be be made is more straightforward. Morgan and Bailey (1943) not only found no enabling effect from context, but argued that the provision of a (story) context interfered with the learning process since learners were inclined to guess the meaning of target words rather than establish their meaning from a dictionary. A similar conclusion was reached by Morgan and Foltz (1944). More recently, the issue has been addressed directly, and with the kind of control often lacking in the earlier experiments, by Mishima (1966) and Lado, Baldwin, and Lobo (1967). In both cases, no significant advantage attached to learning in context over learning word-pairs. Pickering (1982) found no advantage for context learning. Jenkins, Stein, and Wysocki (1984) found context learning to be less enabling than anticipated and concluded that the notion of learning from context is a "default explanation" (p. 769) since there is little evidence that people do learn from context. Curtis (1987) pointed out that the understanding of weaker learners is often bound by the context of learning and only stronger learners are able to decontextualise the items to be learned. A similar point was made by Elshout-Mohr and Van Daalen-Kapteijns (1987) and Sternberg (1987). Sternberg added that the skill of using context is one which needs to be taught; it is not "natural" in any simplistic sense. Pressley, Levin, and McDaniel (1987) saw remembering an item and inferring its meaning from context as complementary skills; inferring *per se* does not necessarily help learning. Their argument is that context does not operate directly on the association between the known L1 item and the unknown L2 item and that it is therefore contributing little to the process of remembering; any contribution that context makes can possibly be put down to the fact that more time is spent processing and this has been shown to have beneficial effects. Nation and Coady (1988) pointed out that beginning learners are often poor decoders even of L1 and are therefore not in a position to make use of context (see also Larsen-Freeman & Long, 1991; Perfetti & Lesgold, 1977, 1979). Summers (1988) argued that it is often not feasible to learn from context because meaning cannot be accurately deduced without help from L1.

Much of the evidence just reviewed must be received with a good degree of caution. As Nation (1987) pointed out, little control over materials or procedures was exercised in many of the experiments. This would not be true of the work of Gershman (1970) who conducted an experiment involving seven learning conditions including word-pairs and various forms of context. Her results showed no significant difference between any of the conditions and she concluded that all the tasks were equivalent; that is, all the tasks were made into word-pair learning tasks. She further argued that with beginners translation is inevitable. The task of learning vocabulary involves, initially at least, forming a word-pair association; context may provide more links, but it is more time-consuming and if time is held constant then paired-associate learning is highly effective. It should be noted, however, that even with word-pair learning, there is the possibility that learners provide their own context and engage in processing which, in the phrase of McDonough (1986), is "learning" rather than "memorisation" (p. 73). The point is that it is probably impossible to determine whether learners are engaged in "rote learning", "memorisation", or "learning proper"; the only distinction that can safely be made is between outcomes from different learning conditions.

Two reasons could be tentatively offered for the continued preference for context over word-pair learning among practitioners. The first is that word-pair learning was often closely associated with Behaviorism and was rejected with it. Krashen's (1977) so-called "Natural Approach" with its stress on acquisition as opposed to learning seems to have struck an appropriate chord at the time (see McLaughlin, 1987). The second reason, suggested by Levenston (1979), is a perceived relationship between "learning" and compound bilingualism. Weinrich (1953) made a distinction between three kinds of bilingualism: co-ordinate, compound, and sub-ordinate. Sub-ordinate bilingualism appears to be an accurate description of the incipient bilingualism of the learner for whom mediation through L1 is essential. This initial division was changed by Ervin and Osgood (1954) with sub-ordinate bilingualism being subsumed under compound bilingualism. The importance of the change is that compound bilingualism has always been seen as a less desirable form of bilingualism by those who accept the distinction (see Carroll, 1963; Saville & Troike, 1982). Whatever the reason, there does seem to be a case for a return to first principles in an attempt to clarify the issue.

Psychological theories on the use of context at learning

From a psychological point of view there are a number of reasons for thinking that the provision of a context at learning might be advantageous to the learner. In relation to the previous discussion on encoding specificity it would mean that the learning context and the testing context would be similar; it would mean the same kind of context (in the more general sense of learning environment) would obtain even though the same sentence was not used at learning and testing. Again, in terms of learning strategy, higher-ability learners may well take a clue from the learning context and adopt a more flexible strategy than if they were led to expect a list test (Hyde & Jenkins, 1969, 1973; McDaniel & Masson, 1977). However there is an additional possibility. That is, that the provision of a sentence context at learning, rather than a list of words, engages subjects in a different kind of learning; in other words, it is advantageous not just because subjects more or less consciously adopt a strategy with a particular kind of test in mind but because a different and more elaborated kind of learning is undertaken.

It has to be said that the literature on elaborated learning is notable for being helpful at the heuristic level rather than for being worked through convincingly at the experimental level. The reason for this is mainly that it has proved impossible to define the relevant factors and thus to provide a suitable metric for further investigation. A roughly historical approach will be adopted in an attempt to cover the main issues in the discussion: organisation; depth; elaboration; encoding distinctiveness; and transfer-appropriate processing.

Organisation Theories

Two main phases can be readily identified in memory study and they can be related to the work of two pioneers in the area, Ebbinghaus and Bartlett. Ebbinghaus (1885) was interested primarily in the mechanisms of memory and attempted as far as possible to eliminate from his experiments prior knowledge and skill on the part of the subject. Bartlett (1932) criticised the unnaturalness of these procedures and concentrated on the complex activity which is remembering. The Ebbinghaus approach was particularly well suited to Behaviorist methodology and a lot of useful work was done on lists and word-pairs; some of the conclusions have been referred to above. However, Bartlett's approach came into its own with the demise of Behaviorism and the advent of Organisation Theories of memory in the late 1960s.

Organisation Theories were developed in response to the perceived inadequacies of theories based merely on association and in acknowledgement of the role of the subject in the process of remembering. Indeed, Wood (1969) argued that a word-pair association constitutes a higher-order unit necessarily involving subject activity (see also Asch, 1968;

Battig, 1968, Deese, 1968; Mandler, 1968; Postman, 1971). A useful distinction made at the time was between primary and secondary organisation (Tulving, 1968). Primary organisation is something which "happens" because of modes of presentation and task characteristics; it accounts for phenomena such as primacy, recency, and lag effects. Secondary organisation is the organisation imposed by the subject and is observed in free recall experiments where the output order is to some extent governed by semantic or phonetic relations among items and by subjects' prior encounters with these items. It was secondary organisation which Organisation Theories concerned themselves with. Organisation Theories will be dealt with in more detail later (see Chapter 5) but in brief the principle behind them was that given the known limits of short-term memory (Miller, 1956), chunking of items into higher-order sets was essential to memory performance (Mandler, 1967; and see Postman, 1972, for a survey). The typical experimental paradigms were those of cued recall and, more usually, the observation of clustering in free recall. It was established that providing subjects with category names at the time of testing enhances recall for categorised lists (Tulving & Pearlstone, 1966; Tulving & Psotka, 1971; Weist, 1970) and a good deal of work was done on the relative effectiveness of different kinds of organisation and cueing techniques. Clustering at recall was taken to be indicative of the way in which items had been organised by subjects in the remembering process. As Bellezza, Cheesman, and Reddy (1977) pointed out, whatever the fate of organisation theories as such, organisation is still an important factor in remembering and there are useful heuristics associated with ideas of categorisation and the use of higher-order units. Organisation Theories as explanations, however, are not finally satisfactory. They have inherent methodological problems because there is no way of establishing the relationship between experimentally-defined categories, subject-learning categories, and subject-recall categories. In other words, it is impossible for these theories to establish a convincing metric (Wood, 1972).

Elaboration Theories

One of the consequences of dissatisfaction with Organisation Theories was the development of Elaboration Theories of remembering which began with the levels of processing theory of Craik and Lockhart (1972). These theories are of particular interest here since they are endorsed by Nation (1987) whose work is influential in the field. Craik and Lockhart envisaged a number of different levels of processing ranging from low level analysis of the stimulus through to deep level semantic analysis. Their argument was that those items are better remembered which are processed to a deeper level. They made a distinction between maintenance rehearsal and elaborative rehearsal. Type I, maintenance rehearsal, is simply a way of keeping items activated at the same level of processing; it has no impact on memory because it involves repeating analyses that have already been carried out. Type II, elaborative rehearsal, does enhance memory because it involves a deeper analysis of the stimulus. The total-time hypothesis, the notion that practice makes perfect, only obtains in the case of Type II rehearsal.

A central experimental paradigm was that based on the notion of incidental learning. It is part of the theory that intention to learn is irrelevant provided the right kind of (semantic) processing takes place. So, for example, Hyde and Jenkins (1973) showed that subjects engaged in a semantic orienting task were as successful in an unexpected free recall task as were members of a control group who had received instructions to learn. Hyde and Jenkins (1973) also claimed to show a correlation between levels of processing measured by different kinds of task and free recall. Again, although the idea of levels of processing has heuristic value, it is not satisfactory at the level of explanation; indeed, Lockhart and Craik (1978) stated that it was never their intention to offer a theory of memory as much as a framework for research.

The arguments against it have been extensively rehearsed. At the theoretical level, Eysenck (1978) argued that the notion of "depth" is unsatisfactory because there is no independent way of measuring it. In similar vein, Nelson (1977) argued that it is impossible to claim that semantic processing is more important than phonemic processing when there is no way of quantifying what "the same amount of semantic and the same amount of phonemic processing" can actually mean. On the distinction between Type I and Type II rehearsal, Jacoby and Craik (1979) pointed out that although it is possible for the experimenter to attempt to direct subjects' attention towards certain features of the stimulus there is no guarantee that one kind of processing rather than another is actually taking place (see Battig, 1968; Crothers & Suppes, 1967; Jenkins, 1979; Kolers; 1979; Madigan, 1969). At the empirical level, the claims made for depth of processing were shown to be too general even within the categories set by Craik and Lockhart. Craik and Lockhart (1972) stated that only Type II rehearsal could lead to improved memory performance. However, Nelson (1977) and Rundus (1977) showed that repeated presentation of an item does lead to enhanced recall even without Type II rehearsal and Maki and Schuler (1980) argued that Type I rehearsal assists delayed recall independently of Type II rehearsal. Bradley and Glenberg (1983) attempted to clarify the discussion about the relative merits of Type I and Type II rehearsal by determining whether duration, attention, or relations between items are responsible for strengthening associations between word-pairs. They concluded that association formation is a function of the number of different relations processed rather than being solely dependent on either sheer duration of processing or the amount of cognitive capacity devoted to processing. This led them to make the useful distinction between rehearsal and repetition where rehearsal is the temporary maintenance of items in short-term memory whereas repetition involves re-presentation of the material such that a new encoding takes place. With this new encoding additional relationships are established. It could certainly be argued that a sentence context provides a richer environment for learning, in this sense, than does list learning *per se* if only because sentence processing may involve subjects in considering more aspects or features of target words (e.g., syntactic and semantic features) than need otherwise be the case. Finally, on the question of incidental learning, Jacoby and Goolkasian (1973) showed that intention to learn can facilitate recall, thus the incidental learning effect is not robust; and McDaniel and Masson (1977) argued that intentional learners show better retention over time whereas incidental learning is short-lived in its effect.

Craik and Tulving (1975) eventually moved away from the notion of depth of processing to that of elaboration (Craik, 1979; Moscovitch & Craik, 1976). Effective elaboration is taken to be that process which adds attributes to the item to be remembered which are salient and which specify the item uniquely. This process can include structural, phonemic, and semantic encodings. It is worth noting that the idea of elaboration is not incompatible with the notion of depth of processing. As Anderson and Reder (1979) pointed out, a word such as *chair* has only one phonemic representation but at the semantic level it has a number of features which could be encoded. The increased amount of information could be a reason why "deeper" processing is more effective than "shallow" processing. The experiment which led to the change of mind of Craik and Tulving (1975) is interesting. The original idea in the experiment was that depth of processing could be controlled by the type of question asked of subjects in the learning condition. In increasing order of complexity, these were: structural (upper or lower case); phonemic (rhyme); category membership; semantic questions (appropriateness of the item for a given sentence slot). The notion was that the time taken to respond would be an indicator of the depth of processing, deeper processing taking more time. Their findings largely supported the predictions that longer latencies are consistent with depth and amount of processing and with probability of recall. However, an anomaly in the results was that items eliciting a positive response in the sentence completion task were better remembered than items eliciting a negative response although response times were equal and therefore, according to the theory, the same kind of processing had taken place in each case. They were led to the conclusion that the processing of an item could not be considered in isolation and the the complexity of the sentence in which it was embedded was a factor in memory. If it elicited a positive subject response it enabled the formation of a coherent unit of question and target which was presumably richer and more elaborate than that formed by the attenuated negative response. As with the notions of organisation and depth of processing, the notion of elaboration has been more successful as a heuristic than as an explanation; its lack of success as an explanation is due to problems of definition similar to those discussed previously. That is, it has proved impossible to devise a suitable metric for the notion of elaboration; and it is not possible to control the processes that subjects actually undertake.

Encoding distinctiveness

A different but related way of talking about facilitating recall has been in terms of encoding distinctiveness (Bransford, Franks, Morris, & Stein, 1979; Eysenck, 1979; Jacoby & Craik, 1979). The basic idea is that "deeper" encodings are more successful than "shallow" encodings because they are more distinctive. Whereas the notion of elaboration tends to stress the amount of processing, notions of "discriminability" and "distinctiveness" call attention to those particular characteristics of the items processed which enable those items to be distinguishable from other items. Thus Bransford, Franks, Morris, and Stein (1979) found that precisely elaborated similes were more successful as recall cues than more elaborate but less precise conceits. However, as before, it has proved impossible to devise any worthwhile definition of distinctiveness which is not circular. Distinctiveness is relative to both an individual subject and the task in hand. Thus distinctiveness will be more appropriate for recognition rather than for recall, and for episodic rather than semantic data. There seems to be a trade-off involved (Postman & Knecht, 1983). The more distinctive the processing, the more it is tied to a particular episode if only because the specific context will probably not arise again. Thus the encoding will be strong but relatively inaccessible (see Eysenck, 1984, p. 117). It will lead to good recognition, poor direct recall, but good indirect recall. The less distinctive the processing, the more adaptable it will be, the better suited to direct recall, but less suited to indirect recall should direct recall fail. In this sense, then, a parallel can be seen with the notions of encoding specificity on the one hand and encoding variability and multiple encoding on the other which also relate to different kinds of demands on memory.

Transfer-appropriate processing

A final concept to be considered which to some extent draws together the points just discussed is that of transfer-appropriate processing (Bransford, Franks, Morris, & Stein, 1979; Jenkins, 1974; Kolers, 1979; Kolers & Roediger, 1984; Morris, Bransford, & Franks, 1977; Tulving, 1979). Morris, Bransford, and Franks (1977) argued that there are some fundamental misconceptions in the theories just considered. Arguments about "superficial" as opposed to "semantic" processing are basically misguided because nothing is either superficial or meaningful in an absolute sense, but only in relation to goals. In experimental terms, what is required is transfer-appropriate processing; that processing is meaningful in a given experimental context which is adapted to the conditions of test or recall. The likelihood is that many "levels" contribute to memory processes and even "superficial" aspects of items are often well remembered (see Papagno, Valentine, & Baddeley, 1991). Even though experimental evidence seems to show in a general sense that semantic processing is more successful than, say, rhyming processing, it should be noted that this has only been shown to be true for semantic tests and largely with college students for whom semantic processing is, presumably, a normal way of processing material. In other words, semantic processing has not been shown to be more successful for different kinds of task and different groups of subjects. Bransford, Franks, Morris, and Stein (1979) developed this point. The importance of elaboration, particularly for cued recall, is that it preserves the relations between items. However, number, quantity, and redundancy of elaborations are not as important as elaboration which is appropriate to the learner and the retrieval context. It is in this context that they see a resolution of the encoding specificity debate; encoding specificity is "transfer-appropriate processing" for certain tasks. It has no inherent advantage attached to it.

Conclusion

It seems reasonable to concur with Nelson (1979) that although semantic processing is one factor among many which can influence memorability, it is particularly appropriate for learning verbal items because our normal contact with verbal items *is* semantic rather than episodic. Eysenck (1984) made a similar point. Semantic encoding is, of its nature, deeper, more elaborate, and more distinctive than non-semantic encoding and this remains true despite the fact that there is no obvious way of unconfounding the factors involved.

Provision of a context at learning could encourage semantic processing if only because subjects may well tend to understand the English sentence and to compare the meaning of its elements with those of the parallel French sentence thus making the process more elaborate than that of merely learning isolated items.

To this extent, therefore, the preference for learning vocabulary in a context reviewed at the beginning of this chapter seems to have some theoretical support from the psychological literature. The fact remains that there is a lack of firm experimental evidence from the second language learning domain. If learning vocabulary in a context does not lead to better retention, then explanations must be sought in the special circumstances obtaining in this domain.

EXPERIMENT 4

Experiment 4, then, was designed to test the effect of the provision of a context at learning on performance. Three possibilities were considered. The first was that the very similarity between learning condition and testing condition would be advantageous for reasons discussed previously in relation to encoding specificity. The second possibility was that learners might well take a clue from the learning context and adopt a more flexible strategy than when learning in a list (Bransford, Franks, Morris, & Stein, 1979; Hyde & Jenkins, 1969, 1973; Jenkins, 1974; Kolers, 1979; Morris, Bransford, & Franks, 1977). The third was that a context at learning would lead to a different and more effective kind of learning (Bransford, Franks, Morris, & Stein, 1979; Craik & Lockhart, 1972; Craik & Tulving, 1975; Eysenck, 1979; Jacoby & Craik, 1979).

A distinction between higher-ability subjects and lower-ability subjects can be made either in terms of the difference in performance between schools as in previous experiments or in terms of overall performance across experiments if results are pooled. Relative to higher-ability and lower-ability performance in previous experiments a working criterion would be above 45% items recalled for higher-ability learners and below 35% items recalled for lower-ability learners. It would be expected that for the higher-ability learners any improvement due to the use of context at learning would be more marked in the case of generation; higher-ability learners were inhibited in context testing in Experiment 2 by their having adopted what turned out to be an inappropriate task-specific strategy. In the case of comprehension, it could be that the advantage would be subsumed in their ability, already demonstrated in Experiment 3, to exploit the possibilities of the context to advantage.

In the case of lower-ability learners, there might be little effect. In the generation task, if they do not adopt a strategy based on the clues given by the learning condition then the presence or absence of context at learning will be irrelevant and performance would be much as in Experiment 2. In the comprehension task, failure to exploit the context will not be changed by the change in learning condition and a decrement would be expected with the change to context testing as in Experiment 3. Indeed, the change in the learning condition may well be further detrimental to lower-ability learners who are not familiar with this style of learning and their performance may suffer as a result.

The first hypothesis to be tested is, therefore, that higher-ability subjects who learn in a simple context and are tested in a simple context would perform more effectively than higher-ability subjects who learn in a list and that this effect would be more marked in the generation task than in the comprehension task. The second hypothesis is that lower-ability learners will derive no advantage from learning in a context where generation is concerned and that the unfamiliarity of the learning context coupled with the use of the backward-association will inhibit their recall in comprehension.

Method

Design

The experiment was a 2 x 2 x 2 factorial design. The two between-subjects factors were the learning condition and the test condition. The learning condition had two levels: learning in a list or learning in a context. The test condition had two levels: the generation task (producing an L2 response to an L1 cue) or comprehension task (producing an L1 response to an L2 cue). The within-subjects factor was time. There were two test days; the first test took place on the same day as learning; the second test took place five days later.

Materials

Lists of word-pairs 20 items long were prepared. The same list was used for Group 1 and Group 2 and a different list for Groups 3 and 4; this was because subjects in Groups 3 and 4 came from the same pool of subjects who had previously been members of Groups 1 and 2. There was no significant difference in overall frequency between the English items in the two lists. The French components, by definition, were all equally unfamiliar to the subjects.

As has been discussed previously, there is no agreement as to the definition of "context". In this experiment, maximum opportunity was given for the exploitation of information other than that contained in the word-pair by shadowing each L2 sentence with its L1 equivalent, with the target items highlighted:

There was **a monkey** in the circus.
Il y avait **un singe** dans le cirque.

Perhaps it will be fine tomorrow.
Peut-être fera-t-il beau demain.

The old man was very **kind.**
Le vieil homme était très **gentil.**

(Full lists can be found in the Materials Appendix).

Subjects

Subjects were taken from the same two schools as in the previous experiments but the pool of subjects was different from that used in the previous experiments. The pool consisted of

45 subjects, 22 girls and 23 boys, from School A4, the mixed-sex school, and 43 subjects from School B4, the all-girls school. Subjects were aged between 11 and 13.

Groups 1 and 2 were tested in the Spring term. Groups 3 and 4 were tested in the Summer term. Subjects had had, therefore, six months of formal French teaching when they began the experiment.

Procedure.

Instructions and time constraints were as in previous experiments. The test involved the completion of a sentence by filling in a slot appropriately. The order in which cue words appeared was randomised across groups; this was designed to reduce possible effects of list-dependency. The cue word was highlighted and positioned directly above the slot to be filled. Subjects had two minutes to fill in the blank spaces after which test papers were collected. A subsequent test was carried out five days later and was not announced to subjects in advance; this was intended to avoid encouraging relearning.

Table 4.1. Experiment 4.

Arrangement of Groups.

Group	Learning condition	Testing condition
1	English-French list	Generation in context
2	English-French context	Generation in context
3	English-French list	Comprehension in context
4	English-French context	Comprehension in context

Four experimental groups were formed in each school. The arrangement of the groups is shown in Table 4.1. Group 1 learned in a list and was tested in a context for generation (English-French). Group 2 learned in a context and was tested in a context for generation (English-French). Group 3 learned in a list and was tested in a context for comprehension (French-English). Group 4 learned in a context and was tested in a context for comprehension (French-English).

Membership of the four groups related to the between-subjects factors in the following way. Groups 1 and 3 learned in a list; Groups 2 and 4 learned in a sentence context. Groups 1 and 2 were tested for generation; Groups 3 and 4 were tested for comprehension.

In School A4, Group 1 was comprised of 20 pupils (10 boys, 10 girls); Group 2, 25 pupils (13 boys, 12 girls); Group 3, 20 pupils (10 boys, 10 girls); Group 4, 25 pupils (13 boys, 12 girls). In School B4, Group 1 was comprised of 19 pupils; Group 2, 21 pupils; Group 3, 23 pupils; Group 4, 20 pupils.

Results and discussion

An initial ANOVA with Schools as a factor revealed a significant difference in performance between School A4 and School B4, School B4 with a mean score of 46.17% correct items being more successful than School A4 with a mean score of 30.6%; F (1, 165) = 51.90, $p < 0.01$. There were also significant interactions between School x Group, F (3, 165) = 7.24, $p < 0.01$, and School x Day, F (1, 165) = 8.10, $p < 0.01$. For these reasons, separate analyses were conducted on the data for the two schools.

Results for School B4: Higher-ability subjec\ts

The provision of a list or a context at learning had a significant effect on performance, F_1 (1, 79) = 28.29, $p < 0.01$, F_2 (1, 76) = 9.85, $p < 0.01$. The mean percentage of items

Table 4.2. Experiment 4. School B4.

Mean percentage scores for items recalled: List or context at learning and generation or comprehension task.

List or context at learning		**Generation or Comprehension**	
List		**Generation**	
Group 1	30.39	Group 1	30.39
Group 3	44.78	Group 2	64.28
Mean	37.58	Mean	47.33
Significance of difference (Tukey test)	$p < 0.05$	Significance of difference (Tukey test)	$p < 0.01$
Context		**Comprehension**	
Group 2	64.28	Group 3	44.78
Group 4	45.25	Group 4	45.25
Mean	54.76	Mean	45.01
Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	Not significant
Difference between list or context at learning significant at: $p < 0.01$		Difference between tasks not significant: $p > 0.47$	

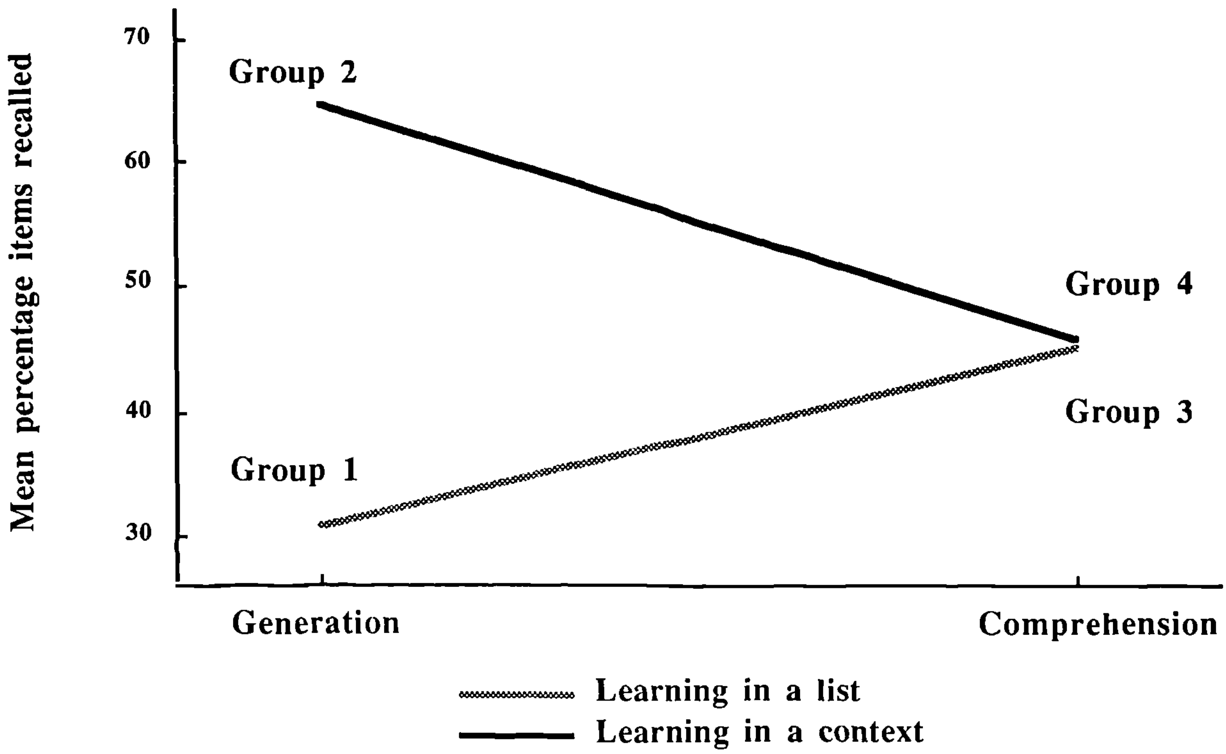

Figure 4.1. Experiment 4. School B4. Interaction: Learning condition x Task.

recalled for subjects learning in a list was 37.58% and for subjects learning in a simple sentence context was 54.76%. Mean percentage scores are detailed in Table 4.2. It appears therefore, that, as predicted, higher-ability subjects who learn in a simple context and are tested in a simple context perform more effectively than higher-ability subjects who learn in a list and are tested in a context.

The type of test did not have a significant effect on performance, F (1, 79) = 0.51, p > 0.47. Those engaged in the generation task (Groups 1 and 2) averaged 47.33% correct responses and those engaged in the comprehension task (Groups 3 and 4) averaged 45.01%.

Table 4.3. Experiment 4. School B4.

Mean percentage scores for items recalled by group and day of testing

Group	Day 1	Day 2	Mean
1. List learners	40.52	20.26	30.39
2. Context learners	75.47	53.09	64.28
3. List learners	53.04	36.52	44.78
4. Context learners	52.75	37.75	45.25
Mean	55.44	36.90	46.17

It had been anticipated that there would be an interaction between the learning condition and the type of test undertaken, with generation being more advantaged than comprehension by learning in a context. This prediction was confirmed (see Figure 4.1). There was a clear interaction between the learning condition and the type of test undertaken, F_1 (1, 79) = 26.77, $p < 0.01$, F_2 (1, 76) = 9.33, $p < 0.01$. Within this interaction, where generation was concerned, subjects in Group 1 who learned in a list averaged 30.39% correct responses whereas subjects in Group 2 who learned in a context averaged 64.28%. A pairwise comparison (Tukey test) shows that this difference is significant at $p < 0.01$. Where comprehension is concerned, the difference between Group 3 who learned in a list (44.78%) and Group 4 who learned in a context (45.25%), was not significant. In other words, the more difficult task, generation, is significantly enhanced by provision of a context at learning (at least for these higher-ability learners); in the case of comprehension the effect of the provision of a context at learning is less clear.

Where generation is concerned, it appears to be the case that, as in Experiment 3, higher-ability list learners are misled by the learning condition to adopt an inappropriate strategy. When the task is comprehension, subjects are able to use the information available in the context at testing to overcome the potential disadvantage.

Subjects in Group 2 (64.28%), who learned in a context, not only performed significantly better than Group 1 (30.39%), the other generation group, but also better than either of the comprehension groups, (Group 3, 44.78%; Group 4, 45.25%), $p < 0.01$ (pairwise comparison, Tukey test). The advantage over Group 1 must be put down to the provision of a context at learning since this is the only condition differentiating Group 1 and Group 2. Group 4 also used a context at learning and yet did not perform significantly better than the list learners in Group 3. However, if subjects in both Group 3 and Group 4 used a strategy well adapted to *generation,* as the learning condition might have encouraged them to do, and if both made good use of the context at testing, then in the comprehension task context learners would not necessarily be in a better situation than list learners. It seems to be the case as in Experiment 3, that constructive use of the context at testing can override the encoding specificity effect. In this case, list learning transferred well to the comprehension task because of the possibilities offered by the context at testing.

The day of testing had a significant effect on performance (see Table 4.3). The difference between Day 1 (55.44%) and Day 2 (36.90%) was significant F_1 (1, 79) = 125.82, $p < 0.01$, F_2 (1, 76) = 41.53, $p < 0.01$. However, there was no significant interaction between this effect and the learning condition, F (1, 79) = 0.08, $p > 0.92$, or the test condition, F (1, 79) = 2.83, $p > 0.09$. In other words, the rate of decay is not affected by more or less elaborated learning.

Results for School A4: Lower-ability subjects

The main feature of the results for School A4 was the generally low level of performance with an overall average recall of 30.60%. As with lower-ability subjects in

Table 4.4. Experiment 4. School A4.

Mean percentage scores for items recalled: List or context at learning and generation or comprehension task.

List or context at learning		**Generation or Comprehension**	
List		**Generation**	
Group 1	28.50	Group 1	28.50
Group 3	34.62	Group 2	37.20
Mean	31.56	Mean	32.85
Significance of difference (Tukey test)	Not significant	Significance of difference (Tukey test)	Not significant
Context		**Comprehension**	
Group 2	37.20	Group 3	34.62
Group 4	22.10	Group 4	22.10
Mean	29.65	Mean	28.36
Significance of difference (Tukey test)	$p < 0.01$	Significance of difference (Tukey test)	$p < 0.05$
Difference between list or context at learning not significant: $p > 0.51$		Difference between tasks not significant: $p > 0.12$	

previous experiments, there was a general levelling out of performance which tended to obscure possible discriminations.

The expectation had been that lower-ability learners would derive no advantage from learning in a context where generation is concerned and that the unfamiliarity of the learning context coupled with the use of the backward-association would inhibit their recall in comprehension.

Mean percentage scores are detailed in Table 4.4. The provision of a list or a context at learning had no significant effect on performance overall, $F\ (1, 86) = 0.43, p > 0.51$. The mean percentage of correct responses for subjects learning in a list was 31.56% and for subjects learning in a simple sentence context was 29.65%.

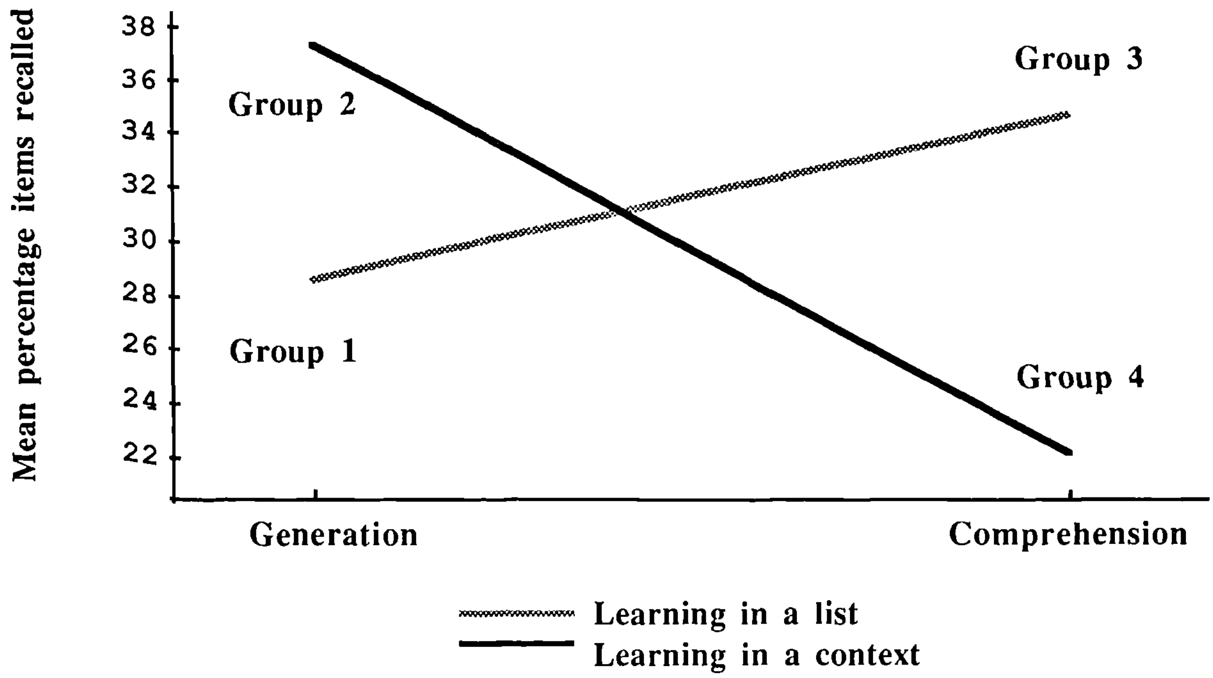

Figure 4.2. Experiment 4. School A4. Interaction: Learning condition x Task.

As with School B4, the type of test did not significantly affect performance overall, F (1, 86) = 2.4, $p > 0.12$. Those engaged in the generation task (Groups 1 and 2) had an average of 32.85% items recalled and those engaged in the comprehension task (Groups 3 and 4) had an average of 28.36%. However, as expected there was a significant interaction between the learning condition and the testing condition, F_1 (1, 86) = 13.46, $p < 0.01$, F_2 (1, 76) = 4.02, $p < 0.05$ (see Figure 4. 2). Where comprehension is concerned, the difference between Group 3 who learned in a list (34.62%) and Group 4 who learned in a context (22.10%) was significant at $p < 0.05$ (pairwise comparison, Tukey test). The only difference between the two groups was the learning condition and it does appear that the lower-ability learners are particularly susceptible to changes in "normal" learning practices. It is worth noting, however, that subjects in Group 4 (22.10%) were also significantly less successful than Group 2 (37.20%) with whom they shared the learning condition but from whom they differed in the test condition, $p < 0.01$ (pairwise comparison, Tukey test). This seems to suggest that the reason for the poor performance of subjects in Group 4 is to do with the combination of the use of the backward-association and the unfamiliarity of the learning condition in addition to their inability to use the possibilities of the context at testing to offset these problems. Where generation is concerned (see Table 4.4 and Figure 4.2), Group 1 who learned in a list averaged 28.50% and Group 2 who learned in a context averaged 37.20% correct responses. Pairwise comparison (Tukey test) shows that this difference was not significant. It seems that weaker learners do not pick up the clues offered by the provision of a context so as to adopt a strategy well suited to testing in a context.

The day of testing had a significant effect on performance, F_1 (1, 86) = 56.01, $p <$ 0.01, F_2 (1, 76) = 69.37, $p < 0.01$ (see Table 4.5). As with the higher-ability learners, there was no significant interaction with either learning condition, F (1, 86) = 0.30, $p >$ 0.58, or with test condition, F (1, 86) = 1.78, $p > 0.18$. Given the generally flat performance of subjects in School A4, this result is not surprising.

Table 4.5. Experiment 4. School A4.

Mean percentage scores for items recalled by group and day of testing.

Group	**Day 1**	**Day 2**	**Mean**
1. List learners	31.75	25.25	28.50
2. Context learners	43.80	30.60	37.20
3. List learners	44.25	25.00	34.62
4. Context learners	26.60	17.60	22.10
Mean	36.60	24.61	30.60

CONCLUSION

These results confirm the prediction that where higher-ability learners are concerned the learning condition is taken to be an indicator of the test condition. This in turn leads subjects to adopt what they consider to be an appropriate strategy. In other words, the learning condition becomes a guide to transfer-appropriate processing (see Bransford, Franks, Morris, & Stein, 1979; Hyde & Jenkins, 1969, 1973; Jenkins, 1974; Kolers, 1979; Morris, Bransford, & Franks, 1977). Where the expectation that testing will be in a list is not fulfilled, the inappropriate nature of the processing results in a decrement in performance as compared, in this case, with learners in a context. This effect is particularly evident in the generation task. Presumably it also occurs in the comprehension task, but any potential disadvantage is counterbalanced by subjects' ability to exploit the possibilities offered by the provision of a context in this task. Where lower-ability subjects are concerned, no observable advantage accrued from learning in a context where the generation task was concerned. The unfamiliarity of the context mode of presentation inhibited their learning and no advantage was gained in the comprehension task from the context available.

Part of the advantage of context learners over list learners, when testing is in a context, is probably due to subjects' engaging in more elaborate processing (Bransford, Franks, Morris, & Stein, 1979; Craik & Lockhart, 1972; Craik & Tulving, 1975; Eysenck, 1979; Jacoby & Craik, 1979). If the context sentences are read, a certain amount of both syntactic and semantic processing will take place. In more general terms, subjects are therefore more likely to treat the task as a semantic task than as an episodic task (see Eysenck, 1984).

It has been argued that beginners are more likely to be capable of and inclined to associative learning rather than conceptual or amodal processing (Chen, 1990; Chen & Ho 1986; Chen & Leung, 1989; Larsen-Freeman & Long, 1991; Nation & Coady, 1988; Perfetti & Lesgold, 1977, 1979). Where lower-ability subjects are concerned, the learning condition does not seem to result in the adoption of any particular strategy and the provision of a context at testing seems not to be exploited. It is probably true to say that neither their list learning nor their context learning is adapted to list testing and context testing respectively; they do not undertake transfer-appropriate processing. For this reason they do not show a decrement when only the test condition is changed; it is the combination of an unfamiliar test condition with the use of the weaker backward-association in the comprehension task (as in Experiments 3 and 4) that performance is impaired (see Cohen & Aphek, 1980).

Although a strict comparison across experiments is, of course, impossible, it is interesting to look at the pattern of results across Experiments 2, 3, and 4 (see Table 4.6). It seems that higher-ability learners are characterised by their ability to develop a suitable strategy, to use a context at testing where it can offer any assistance, to develop a stronger bond between the two items in the word-pair which can operate successfully even with the weaker backward-association. Lower-ability learners on the other hand do not adapt their strategy according to the perceived task, do not make effective use of the context at testing when this offers assistance, develop a weaker word-pair bond which is particularly vulnerable to the combination of a change of condition between learning and testing and the use of the backward-association. The issue of higher- and lower-ability students will be revisited in Chapter 9.

In the experiments to date, performance over time has been observed. It is clear that in all cases a significant decrement in performance occurred between Day 1 and Day 2 (see Table 4.7 and Table 4.8) but further than that there is no discernible pattern to the results. Long-term remembering is the aim of language learning but a substantial amount of forgetting takes place within a few days irrespective of the modes of learning or testing used in these experiments. These results are in line with the conclusion of Nation (1982) that most forgetting takes place immediately after initial learning. This finding has given rise to many studies on the phasing of repetition which are beyond the scope of this study (see Bahrick & Phelps, 1987; Pimsleur, 1967). The present data suggests that individual differences, differences in materials, and task demands are going to make precision in this respect difficult.

Table 4.6. Experiments 2, 3, 4.

A summary of the performance of higher-ability and lower-ability learners according to learning and task conditions. (All learning was English-French.)

		Higher-ability subjects
Experiment	**Test condition**	**Performance by learning/test condition**
2	Generation	List to list > list to context
4	Generation	Context to context > list to context
3	Comprehension	List to list = list to context
4	Comprehension	Context to context = list to context
		Lower-ability subjects
2	Generation	List to list = list to context
4	Generation	List to context = context to context
3	Comprehension	List to list > list to context
4	Comprehension	List to context > context to context

Table 4.7. Experiments 1-4. A summary of percentage decrement between Day 1 and Day 2.

Experiment	Learning and testing condition	Day 1	Day 2	Percentage decrement
Higher-ability subjects				
1. School A1	List-List (generation)	46.66	41.32	5.34
1. School A1	List-List (comprehension)	59.53	49.24	10.29
2. School B2	List-List (generation)	74.21	53.94	20.27
2. School B2	List-Context (generation)	36.57	19.21	17.36
3. School B3	List-List (comprehension)	68.15	47.63	20.52
3. School B3	List-Context (comprehension)	57.50	44.31	13.19
4. School B4	List-Context (generation)	40.52	20.26	20.26
4. School B4	Context-Context (generation)	75.47	53.09	22.38
4. School B4	List-Context (comprehension)	53.04	36.52	16.52
4. School B4	Context-Context (comprehension)	52.75	37.75	15.00
Lower-ability				
1. School B1	List-List (generation)	30.18	17.28	12.90
1. School B1	List-List (comprehension)	50.03	35.44	14.59
2. School A2	List-List (generation)	49.76	28.09	21.67
2. School A2	List-Context (generation)	39.80	22.00	17.80
3. School A3	List-List (comprehension)	39.75	35.25	4.50
3. School A3	List-Context (comprehension)	36.00	23.60	12.40
4. School A4	List-Context (generation)	31.75	25.25	6.50
4. School A4	Context-Context (generation)	43.80	30.60	13.20
4. School A4	List-Context (comprehension)	44.25	25.00	19.25
4. School A4	Context-Context (comprehension)	26.60	17.60	9.00

Table 4.8. Experiments 1-4. A summary of percentage decrement between Day 1 and Day 2. Rank order.

Experiment	Learning and testing condition	Day 1	Day 2	Percentage decrement
Higher-ability				
1. School A1	List-List (generation)	46.66	41.32	5.34
1. School A1	List-List (comprehension)	59.53	49.24	10.29
3. School B3	List-Context (comprehension)	57.50	44.31	13.19
4. School B4	Context-Context (comprehension)	52.75	37.75	15.00
4. School B4	List-Context (comprehension)	53.04	36.52	16.52
2. School B2	List-Context (generation)	36.57	19.21	17.36
4. School B4	List-Context (generation)	40.52	20.26	20.26
2. School B2	List-List (generation)	74.21	53.94	20.27
3. School B3	List-List (comprehension)	68.15	47.63	20.52
4. School B4	Context-Context (generation)	75.47	53.09	22.38
Lower-ability				
3. School A3	List-List (comprehension)	39.75	35.25	4.50
4. School A4	List-Context (generation)	31.75	25.25	6.50
4. School A4	Context-Context (comprehension)	26.60	17.60	9.00
3. School A3	List-Context (comprehension)	36.00	23.60	12.40
1. School B1	List-List (generation)	30.18	17.28	12.90
4. School A4	Context-Context (generation)	43.80	30.60	13.20
1. School B1	List-List (comprehension)	50.03	35.44	14.59
2. School A2	List-Context (generation)	39.80	22.00	17.80
4. School A4	List-Context (comprehension)	44.25	25.00	19.25
2. School A2	List-List (generation)	49.76	28.09	21.67

CHAPTER 5

List position and serial order effects and word-pair presentation

In the experiments carried out up to this point, regardless of the form of presentation at learning and the type of testing undertaken, some items in lists were learned and others were not. It seems reasonable to conclude, therefore, that whatever the effects of the direction of learning, the use of forward- or backward-association, the relative ease of generation and comprehension, and encoding specificity, there are other factors which affect learner performance. One such factor may be the list format itself. The assumption has been made in previous experiments that list-dependency is a possible effect of list learning. The intention here is to test this assumption. In this chapter, list position refers to the primacy and recency effect; serial order refers to organisation based on the order of items through the list.

A related question concerns the form of presentation used. Intuitively, it would seem that learning L2 vocabulary items as part of an L1-L2 word-pair is a "natural" way of proceeding, at least in the early stages of language learning. As discussed in the previous chapter, Mishima (1966) and Lado, Baldwin, and Lobo (1967) found presentation modes as such to be relatively unimportant. Gershman (1970) suggested that subjects make all tasks into word-pair learning tasks. It was argued above that even in word-pair learning, learners could provide their own context. Either way, presentation becomes relatively unimportant. Albert and Obler (1978) claim that: "There is no doubt that a word and its translation equivalent are connected in a nonrandom way" (p. 69). It could be therefore that word-pair learning is what subjects do whatever the mode of presentation; if this is the case, then it would seem that more attention should be paid to word-pair learning as such and less effort made to devise alternative methods of presentation.

Organisation and memory

There is a good deal of evidence to suggest that list learners use some form of organisation in attempting to memorise items (e.g., Banks & White, 1982; Brown, 1979; Mandler, Worden, & Graesser, 1974). A useful way of envisaging the process is to see subjects as clustering items together in some way which facilitates recall (Weist, 1972). Two ways of clustering items are categorisation based on perceived semantic relationships between items, and clustering based on the position of items in a list. Where list position is concerned, the most obvious manifestation would be primacy and recency effects which have been observed in many experimental conditions. Categorisation and list position effects could potentially conflict because it would appear that different processes are involved. Categorisation in a random list is normally going to result in a different form of organisation from one based on list position. Categorisation is effortful whereas serial organisation is in a sense "given" in list presentation. Categorisation is taken to involve "deeper" processing and serial organisation "shallower" processing.

Categorisation

Categorisation as an aid to memory has been extensively studied and, in general, it can be stated that categorisation does assist recall. However, beyond that it is difficult to go because it is probably impossible to determine what form of categorisation is being used by individuals either on a given occasion or between occasions. Baddeley (1976) reviewed the evidence. One of the early studies was that of Bousfield (1953) who found that if subjects were presented with a list made up of randomised categorial items, they tended to recall items in clusters by category. A large group of experiments followed this early work and these experiments appear to confirm the finding that subjects group together words which share a common category, whatever the order of presentation. Categorisation strategies can be shown in several ways. An early procedure was to examine cluster patterns in free recall. Baddeley (1976) and Bower, Clark, Lesgold, and Winenz (1969) detailed some of the results of research in this area. Thus the categorisation effect is greater for exhaustive categories (e.g., the points of the compass) than it is for non-exhaustive categories (Cohen, 1966); it is greater for high-frequency associates of the category name than it is for low-frequency associates (Cofer, Bruce, & Reicher, 1966; Deese, 1959;

Jenkins, Mink, & Russell, 1958); it is greater when items are presented in category blocks than when items are randomly presented (ibid.); and it can be found in cued recall (Tulving & Pearlstone, 1966). It is claimed that the categorisation effect is partly due to inter-item associations (Deese, 1959; Jenkins & Russell, 1952; Rothkopf & Coke, 1961) but also partly due to items' sharing a common superordinate category (Baddeley, 1976; Bousfield, 1953).

The problem with the procedure used to study categorisation was that in most of the early experiments, categories were dictated by those conducting the experiments and due consideration was not given to the possibility that categorisations used by subjects could be quite different from those envisaged by the experimenter. In making this point, Mandler (1977) included in a list of possible categories: associative categories, syntactic categories, semantic categories, and the cover-all of "idiosyncratic categories". Others have suggested further categories for inclusion (Bower, 1967; Voss, 1972; Wickens, 1970). This range of possibilities calls in question the assumption that experimenter-defined categories are the same as subject-defined categories and therefore the robustness of the conclusions drawn on the basis of that assumption.

A second procedure for observing categorisation effects arose from problems associated with the paradigms responsible for the evidence just considered and from a dissatisfaction with associationist explanations for categorisation effects. Tulving (1968) made a distinction between effects due to what he called primary organisation and those due to what he called secondary organisation. Primary organisation is something which "happens" independently of any strategy adopted by the subject or of any prior familiarity with the material to be learned. Secondary organisation is that form of organisation which is related to perceived semantic, or phonemic, or orthographic relations in the items to be remembered. Tulving was aware of the possibility of a claim that the fact that subjects recall items in a certain order, and that clusters of items correspond to what the experimenter expected, could be a reason for thinking categorisation effects were a form of primary organisation rather than secondary organisation and therefore explicable within the associationist approach to memory. However, Tulving (1962) contended that active subjective organisation was necessary for free recall learning and that mere repetition was not sufficient for learning as the associationist explanation would imply. He claimed to have established a metric for subjective organisation (SO). The basic idea was to compare the number of times any two items were recalled together on successive trials with what would have been expected on a chance basis.

As a demonstration of the impossibility of an associationist explanation based on mere repetition of items, and consequently as a demonstration of the effect of SO, Tulving (1966) reported the following experiment. Prior to a learning-testing experiment, one group of subjects read a list of twenty two letters paired with nouns; another group read a list of the same twenty two letters paired with what would become the target nouns in the learning and testing condition. Even after reading through the lists six times, subjects who had read the target nouns learned no more effectively in the new learning-test condition than those who had read the "irrelevant nouns". In another experiment, subjects learned a list of nine words which was then incorporated into a longer list of 18 words. The performance of this group was actually inferior to that of a control group which learned all 18 words as a new set of words. Tulving's explanation was that the subjective organisation suitable for the shorter list was not appropriate in the case of the longer list and recall was impaired despite the fact that the material had been encountered more frequently. Although these results are interesting, the part-to-whole negative transfer effect is not in itself the convincing evidence of subjects' organisation that Tulving took it to be as Postman (1972) argued. Roberts (1969) showed that subjects spent a "disproportionate" amount of time studying new items when confronted with the composite list and this could be a source of interference with the already learned items. Some support for this view came from Novinski (1969) who reported that old items showed the greater degree of impairment in the part-to-whole transfer experiments. Wood and Clark (1969) showed a much reduced negative effect when subjects were informed about the composition of the new list and later Slamecka, Moore, and Carey (1972) showed that when subjects were told to include in their response items about which they were not sure, the part-to-whole negative transfer effect disappeared without loss of accuracy. In other words, subjects may not have included already-learned items in their response because they felt these items were intrusions.

Mandler (1967) adopted a different approach to demonstrating categorisation effects based on instructions to subjects. He found that when subjects were instructed to categorise words on cards into conceptual categories three effects were clearly established. First, the instruction to categorise was as effective as an instruction to learn the items. Second, the more thorough the categorisation, indicated by the number of categories, the better the recall. Third, the larger the category size, the less items were recalled from that category. Again, these results are interesting but they do not necessarily show what they are claimed to show since there is no way of establishing the relationship between the *post hoc* categorisation used by the subject and the processes which have actually been undertaken in the process of remembering (Slamecka, 1968).

List position effects on organisation

The intractability of the problem of determining categorisation processes was one of the causes of loss of interest in organisation theories and the move towards levels of processing and elaboration theories as explanations of memory performance. A reasonable conclusion would seem to be that one would have to assume that categorisation takes place, but that there is no clear way of establishing just what process of categorisation is undertaken by individuals on different occasions. This is a justification, therefore, for concentrating on list position effects and for attempting to find whether there is evidence for list position effects irrespective of the presumed categorisation process undertaken by individual learners.

Where list learning is concerned, the most well-established list position effects are the primacy effect and the recency effect though the caveat needs to be entered that the effects were mainly (though not exclusively) studied in relation to oral presentation of lists of items to be remembered. Murdock (1962) showed that when a list of items is learned and freely recalled, the probability that a given item will be recalled is dependent on its position in the learning list. Subjects tend to recall first items at the end of the list and the accuracy of recall of these items is high compared with items recalled from earlier in the list; this is the recency effect. The primacy effect means that items at the beginning of the list are next to be recalled. Initially, these effects were considered to exhibit different characteristics which led to their being widely studied in discussions about differences between so-called short-term and long-term memory (STM and LTM).

The recency effect can be enhanced in relation to the primacy effect by a faster rate of presentation. It is very robust and is apparently unaffected by the characteristics of the words concerned (Glanzer, 1972). The recency effect can be removed by a delay between presentation of the list and the start of testing where the lag is filled by a counting task. For this reason it was taken to be largely acoustically based and therefore associated with short-term memory (Glanzer & Cunitz, 1966; Peterson & Peterson, 1959; Postman & Phillips, 1965; Rundus & Atkinson, 1970; Shiffrin, 1970).

The primacy effect increases with a slower rate of presentation (Bousfield, Cohen, & Whitmarsh, 1958; Deese, 1957; Murdock, 1962). It can be affected by a number of variables as, for example, similarity of sound or meaning between items (Craik & Levy, 1970); the number of syllables in the items (Craik, 1968); the number of times items are repeated in the list (Glanzer & Meinzer, 1967); the number of languages involved (Tulving & Colotla, 1970); word frequency (Raymond, 1969); the age of the subject (Craik, 1968). In other words, it is affected by many of the variables which affect memory generally and for this reason it is taken to be a long-term memory phenomenon. According to Murdock (1962) the primacy effect is more precipitous than the recency effect but smaller in magnitude; it extends over the first three or four items in the list and the recency effect over the last eight serial positions.

Neat though the attribution of the effects to separate memory stores is, it can probably not be sustained, even if the dubious distinction between short-term memory and long-term memory is accepted as other than a merely logical distinction (see Anderson, 1985; Crowder, 1982; Kintsch, 1974). Baddeley and Hitch (1974, 1977), Bjork and Whitten (1974), and Tzeng (1973) all showed that recency effects can be obtained in free recall under conditions which are taken to eliminate short-term memory performance.

Tzeng (1973) for example showed that the recency effect can be obtained even with a time lag and a distractor task and this effect was retained in a final free-recall test.

Negative recency effects have also been observed. Maskarinec and Brown (1974), for example, examined the claim that there was a positive recency effect only in immediate free recall and that this became a negative recency effect in final free recall; that is, that when at the end of an experimental session subjects were asked to recall what they could of all lists encountered, words at the end of all lists were badly remembered. Although, as mentioned, this result had been put down to the fact that the recency effect was essentially a short-term memory phenomenon, Craik (1970) and Jacoby and Bartz (1972) had argued that it was the result of a strategy adopted by learners who in the expectation of an immediate test did not undertake to learn the final items in other than a temporary manner. Maskarinec and Brown pointed out that it could equally be the result of a retrieval problem rather than an encoding problem; that is, it was not a question of the recency effect being dependent on (temporary) short-term memory but that subjects encoded the material without adopting at the same time an adequate retrieval strategy. Maskarinec and Brown (1974) adopted the experimental paradigm of exposing subjects to 10 lists of items with the expectation that lists would be 21 items long; one list, however, was curtailed after 12 items. This list showed no negative recency effect in final free recall thus suggesting that the negative recency effect was due to subject strategy rather than initial encoding being transitory.

The work of Baddeley and Hitch (1974) is particularly interesting because the paradigm used was less "artificial" than many and so would seem to be more applicable to the present domain. They found that there was a marked recency effect in an incidental recall task although the end of the task had been followed by a minute-long discussion of methods of solving anagrams, an activity which would certainly eliminate any so-called short-term memory effects. On the question of subject strategy, Baddeley and Hitch (1977) argued that the fact that recency effects are found in incidental learning tasks indicates that although the effect can be enhanced by subject strategy, to some extent it is independent of it. However it is not possible to claim that the independent component is to be found exclusively in short-term memory. Although the short-term memory is taken to be acoustically based, an articulatory suppression technique had no influence on the recency effect thus leading to the conclusion that the recency effect was due to a retrieval strategy which was equally usable in long-term memory.

The question remains of why recency effects are not always evident. The suggestion of Baddeley (1976) was that the strategy of retrieval by using short-term recency effects (though not necessarily STM effects) is only adopted in the absence of better retrieval strategies; better strategies might involve categorisation or some other kind of elaborated processing as discussed earlier. This would explain why recency effects tend to be less with categorised lists than they are with random lists, and with lists where there is little possibility of inter-item organisation. Thus in the Tzeng (1973) experiment the requirement to count backwards before and after each item would make inter-item organisation impossible and thus make subjects reliant on primary organisation. Ordinal cues therefore are used in the absence of alternative means of recall.

Serial order effects

Serial order organisation, that is organisation based on the order in which items appear, is another aspect of the effect of list presentation which needs to be considered. Intuitively, it appears to be a powerful mnemonic device in some restricted cases such as memory for the days of the week, the months of the year, and the letters of the alphabet. In addition to these rather restricted applications, there is experimental evidence for the importance of serial input as an organising principle in other kinds of lists. The early work of Tulving and Patkau (1962) on serial ordering was supported by Jung and Skeebo (1967), Lachman and Laughery (1968), Mandler and Dean (1969), Postman, Burns, and Hasher (1970) who all showed that recall is more effective when a list is re-presented in the same order as learning than it is when presentation is randomised. Mandler and Dean (1969) made the strong claim that seriation is the preferred method of subjects for organising lists, even when other modes of organisation based on semantics, for example, are available. Kintsch (1970) showed that for categorised as well as for non-categorised items there is a correlation between the serial position of items in a list at learning and at testing in free recall; this is roughly in line with the findings of Murdock (1976) who showed that loss of order information increases the probability of loss of item information. In the present experiments, list order at testing is different from list order at learning; however, a strategy based on list order would be revealed if the pattern of recall were continuous from any primacy effect. In other words, the decline in recall after recall of items from early positions in the list would be consistent.

Specifically on word-pair learning, Lesgold and Bower (1970) demonstrated the possibility of interference when subjects chose an inappropriate strategy on the basis of instruction given at learning. Two groups learned sets of paired-associates. Subjects in Group 1 were told that the paired-associates were taken from a list; subjects in Group 2 were not instructed in this manner. At recall of the paired-associates, it was clear that Group 1 was disadvantaged as compared with Group 2 and Lesgold and Bower (following Jensen, 1962) argued that this was because the information they had received encouraged subjects in Group 1 to engage in serial learning which interfered with the process of paired-associate learning.

It is not clear, however, that serial learning is necessarily inhibitory. Segal and Mandler (1967) argued that when subjects are presented with a list of word-pairs, they learn not only "horizontal" word-pair associations but construct some form of "vertical" organisation for items in the list. They showed that when word-pairs are learned in both directions (A-B and B-A) then horizontal associations between the two halves of the word-pairs are the basis of recall. However, when word-pair lists are learned in one direction only, a vertical organisation is set up within items in the stimulus position in the list and within items in the response position in the list; this is in addition to the horizontal associations. It appears, therefore, that a stimulus set and a response set is perceived and as a result free recall is higher when all items come from one vertical set than when the list is mixed. Segal and Mandler (1967) found little evidence of actual inter-item associations between vertically organised sets and their suggestion was that facilitation of recall due to vertical organisation should be attributed to "right-side" and "left-side" cues. It is not entirely clear what this means; however, the point remains that some form of vertical dependence can be formed in list learning which could be facilitatory. Even if the test order is different from the learning order, subjects could still use an indirect technique whereby they first recovered items by free recall then identified the appropriate response; something of this sort is seen when the alphabet is run through in order to ascertain, for example, what letter follows *h*.

Conclusion

If subjects base at least part of their learning strategy, whether consciously or unconsciously, on list position or serial order, this would have implications for the use of lists in vocabulary learning. If subjects relied on list position, words in the middle of the list would be difficult to recall. If subjects relied on serial order, changing the serial order could also impede recall.

However, most of the research on list position effects has been concerned with lower-level tasks than the cued recall of vocabulary items. These tasks allowed little scope for indirect retrieval involving reconstruction of target items (see Chapter 3, p. 64 above). Experiments 5 and 6 were therefore designed to find out whether there is evidence for list position and serial order effects in the present domain when list order at learning was different from that at testing. A second question to be examined is whether similar effects are to be found when items are presented in "normal" word-pair lists compared with presentation of items in simple sentence contexts of different degrees of complexity.

EXPERIMENT 5

Experiment 5 was designed to test whether list position and serial order effects influence performance, and whether the extent of their influence is affected by the form of presentation used, when the task is generation. On the assumption that subjects make use of the most available form of organisation, in the case of English-French learning it might be expected that list position and serial effects would be more likely to be in evidence in this task than in comprehension because in generation the ostensible targets, the L2 items, are unknown and no obvious basis exists for clustering these items semantically.

Of the list position effects, the evidence suggests that the recency effect is normally robust and unaffected by the items used in testing, provided testing immediately follows presentation (Glanzer, 1972). There is evidence to suggest that the recency effect is not just a short-term phenomenon (Baddeley & Hitch, 1974, 1977; Bjork and Whitten, 1974; Tzeng, 1973). Even so, it could be that the effect is mainly acoustically based, and for that reason mainly a short-term memory phenomenon (Glanzer & Cunitz, 1966; Postman & Phillips, 1965). If this were the case, it is not clear whether recency would be influential for written language presentation with a time lag of some minutes between learning and testing. Such a time lag would normally be taken to eliminate short-term memory effects, as would verbal and written instructions to subjects intervening between the learning and testing sessions. The recency effect is taken to be dependent to some extent on subject strategy (Jacoby & Bartz, 1972; Maskarinec & Brown, 1974) though Baddeley and Hitch (1977) show that it occurs also in incidental learning. It is not clear, therefore, whether it is to be expected in the present learning and test conditions.

The primacy effect is taken to be more volatile and to be affected by a range of factors which affect memory for language items generally such as: similarity of sound or meaning between items (Craik & Levy, 1970); the number of syllables in the items (Craik, 1968); the number of times items are repeated in the list (Glanzer & Meinzer, 1967); the number of languages involved (Tulving & Colotla, 1970); word frequency (Raymond, 1969); the age of the subject (Craik, 1968; Craik & Levy, 1970; Glanzer & Meinzer, 1967; Raymond, 1969; Tulving & Colotla, 1970). On the other hand, the primacy effect increases with a slower rate of presentation (Bousfield, Cohen, & Whitmarsh, 1958; Deese, 1957; Murdock, 1962) and the present learning condition which is based on simultaneous rather than serial presentation might be expected to be favourable to the primacy effect.

Finally, there is evidence to suggest that serial order overall, and not just in relation to the first few items and the last few items in the list, is a factor in learning lists (Jung & Skeebo, 1967; Kintsch, 1970; Lachman & Laughery, 1968; Murdock, 1976; Postman, Burns, & Hasher, 1970; Tulving & Patkau, 1962) and it is argued that serial order can be a preferred method of organisation even when other methods are available (Mandler & Dean, 1969). This would manifest itself in clustering by serial position other than or in addition to the primacy and recency positions.

On the question of presentation, there is an argument, as discussed above, that presentation as such is relatively unimportant either because learners tend to make the process into one of word-pair learning, or because they provide their own context, or both (see Gershman, 1970; Lado, Baldwin, and Lobo, 1967; Mishima, 1966).

Given this evidence, it was anticipated that some serial position and serial order effects would be observed, whatever the form of presentation. However, it was also predicted that the full range of these effects was unlikely to be seen because of the range of other possible ways of organising the materials open to subjects.

Method

Design

The experiment was a 2 x 3 x 3 factorial design. The preliminary analysis revealed no significant difference between the two schools, so School was included as a between-subjects factor. The other between-subjects factor was type of presentation. It had three levels: presentation in a word-pair list; presentation in a simple sentence context; presentation in a disrupted context (see below for details). The within-subjects factor was list position. Its three levels were: the beginning of the list, taken to be the first six items; the end of the list, taken to be the last six items; the middle of the list, taken to be the remaining eight items.

Materials

Groups were given lists of word-pairs 20 items long. The pattern of categories was as in previous experiments. All groups used the same items but three different forms of presentation were used. One form of presentation was a "normal" list of word-pairs. The second presented items in a simple sentence context (as in Experiment 4). The third form of presentation disrupted the word-pair and list formats as far as possible in order to test the importance of an overt list presentation for list position and serial order effects to occur. The cue word was placed in the left-hand margin and was separated from the target word by sentence contexts of various lengths. Although the cue word had to be repeated within the sentence, it was not emboldened and therefore the word-pair was not clearly displayed as in Group 1 and, to a lesser extent, in Group 2:

the tyre	The tyre of his car was flat. **Le pneu** de sa voiture était dégonflé.
despite	Despite his illness he was happy. **Malgré** sa maladie il était content.
smooth	The surface of the mirror was smooth. La surface du miroir était **lisse.**

Four versions of the list were produced for learning purposes and all results refer to original learning order. (Full lists are contained in the Materials Appendix).

Subjects

Subjects from two schools were involved in the experiment, 54 from School A5 and 55 from School C5. Both were comprehensive schools; both were mixed sex schools. Different subjects from School A5 had been involved in previous experiments. The experiment took place in the Spring term of the school year and subjects had already experienced, therefore, six months of formal French teaching when they began the experiment. Subjects were aged between 11 and 13.

Table 5.1. Experiment 5.

Arrangement of Groups.

Group	Learning direction	Learning condition	Testing condition
1	English-French	List	English-French (context)
2	English-French	Context	English-French (context)
3	English-French	Disrupted list	English-French (context)

Procedure

Three experimental groups were formed in each school. In School A5, Group 1 consisted of 18 pupils (9 boys, 9 girls); Group 2 consisted of 19 pupils (9 boys, 10 girls); Group 3 consisted on 17 pupils (10 boys, 7 girls). In School C5, Group 1 consisted of 22 pupils (10 boys, 12 girls); Group 2 consisted of 14 pupils (7 boys, 7 girls); Group 3 consisted of 19 pupils (10 boys, 9 girls).

Membership of the groups related to the between-subjects factor in the following way. Group 1 learned in a normal word-pair list. Group 2 learned in a simple sentence context. Group 3 learned in the disrupted format. All subjects were tested in a simple context since the ability to recall words in a "natural" context was central to the experiment. The arrangement of the groups is shown in Table 5.1.

For purposes of analysis, performance of subjects was compared by using a mean percentage established for each of the components.

In other respects, the procedure was as in previous experiments.

Results and discussion

An initial ANOVA revealed no significant difference between the performance of the two schools, School A5 averaging 30.33% correct responses and School C5, 35.93%, F (1, 104) = 2.11, $p > 0.14$. The results were therefore pooled with School as a between-subjects factor.

List position had a significant effect on recall, F (2, 208) = 92.38 $p < 0.01$ (see Table 5.2. and Figure 5.1). Pairwise comparison (Tukey test) shows a clear primacy effect. The mean percentage score for the first six items in the list was 46.12% items correctly recalled. This was significantly better than for items in the middle position of the list, 34.71%, and for items in the final position 18.55%, $p < 0.01$. In addition, items in the middle position in the list were significantly better recalled than items in the final position in the list, $p < 0.01$. These results suggest therefore a strong serial element in

Table 5.2. Experiment 5. School A5 and School C5 pooled. Mean percentage scores for items recalled.

Group	List position			Overall
	Beginning	Middle	End	
1	44.93	27.08	16.68	29.56
2	49.21	42.20	22.59	38.00
3	44.23	34.85	16.40	31.82
Overall	46.12	34.71	18.55	33.12

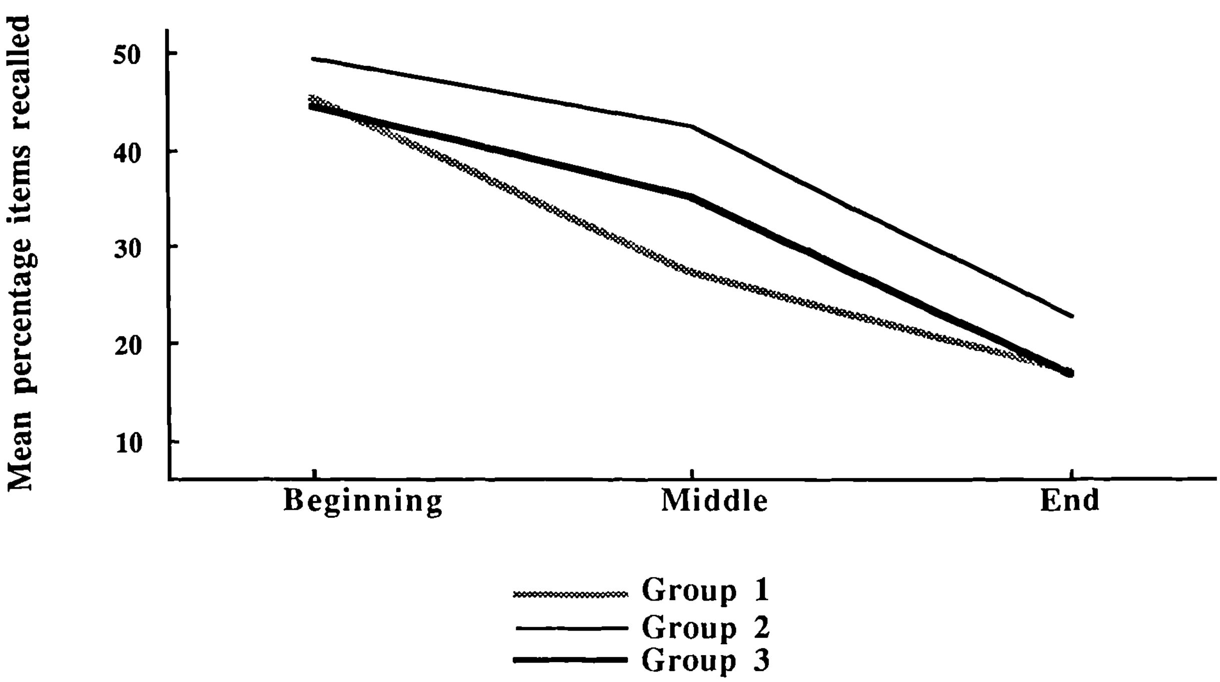

Figure 5.1. Experiment 5. Mean percentage scores for items recalled: Learning condition x List position. Group 1 learned in a normal word-pair list; Group 2 in simple sentence contexts; Group 3 in a disrupted format.

learner organisation beyond the primacy effect, as such, since Murdock (1962) had talked about a primacy effect extending only over the first three or four items in the list. Here items in the middle position in the list were learned well relative to items at the end of the list. This suggests that subjects used the list order as a principle of organisation and that whatever the possibilities for other forms of item organisation they were heavily influenced by the serial position of items at learning.

There was no significant interaction between the learning condition and the position of items in the list, F (4, 208) = 1.55, $p > 0.18$. The effect of serial position was thus consistent across the different forms of presentation.

The form of presentation did not significantly affect performance, F (2, 104) = 1.70, $p > 0.18$. Group 1, who learned in a "normal" list of word-pairs averaged 29.56% correct responses; Group 2, who learned in a simple sentence context, averaged 38.00%;

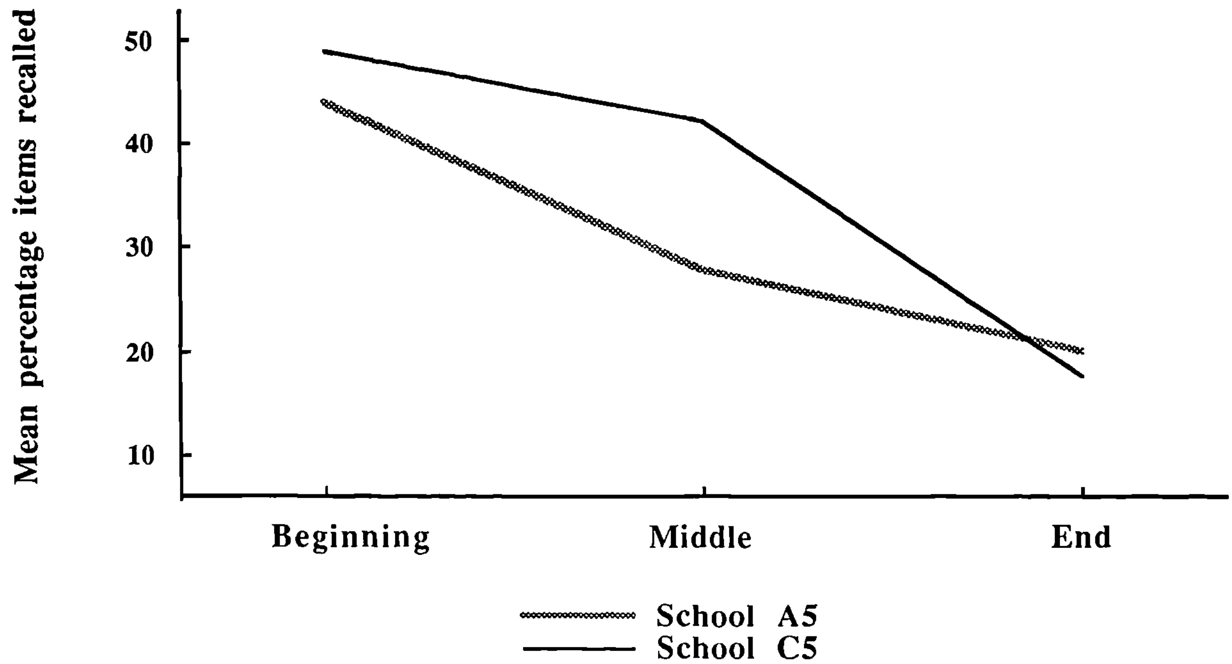

Figure 5.2. Experiment 5. Mean percentage scores for items recalled: School x List position.

and Group 3, who learned in the disrupted list, averaged 31.82%. This would seem to indicate that similar patterns of learning took place in each case and the most likely explanation is that subjects saw the task as one of learning word-pairs (see Gershman, 1970). Given the generally low level of performance, and the importance of list position effects, it seems unlikely that any of the subjects undertook more elaborate processing than the basic word-pair learning. These results reinforce the notion that there may well be a discrepancy between items as presented by the experimenter and items as perceived by subjects.

It was argued previously that lower-ability subjects (defined by the criterion used previously, i.e., < 35% items recalled) show little evidence of adopting a list-based strategy because in previous experiments they showed no decrement in the generation task when the test condition was different from the learning condition. Here the lack of difference between performance over different conditions of learning supports this argument. What seems to happen is that lower-ability subjects start at the beginning of the list and work their way through until time runs out; this could be called a strategy of sorts but presumably the strategy undertaken by higher-ability learners is rather more sophisticated than this.

There was a significant interaction between School and Position, $F\ (2, 208) = 8.38$, $p < 0.01$. A simple effects analysis identifies the source of the interaction as being at Position 2, $F\ (1, 104) = 9.51, p < 0.01$. Figure 5.2. shows the divergence in performance on items in the middle position in the list. Whereas subjects in School A5 averaged 27.59% items recalled in this position, School C5 averaged 41.83% items recalled. There is no obvious explanation for this result but it does serve to highlight the difficulty of generalising about learning strategies even within a relatively homogeneous population.

EXPERIMENT 6

Experiment 6 was similar in design to Experiment 5, but with comprehension as the task rather than generation. In general terms, it might be expected that list position and serial order effects would be less in evidence in a comprehension task where the context could be used, as discussed previously, to assist recall, thus obscuring any such effects. On the other hand since the target items are in L1 there is the possibility of subjects' establishing vertical relationships between target items at learning and of their exploiting those relationships at testing.

Where mode of presentation is concerned, in addition to the argument of Gershman (1970) considered previously, the availability of the sentence context in the easier comprehension task would help subjects of higher-ability and this would be expected to subsume any effects from different modes of presentation. Lower-ability subjects might show a decrement due to the unfamiliarity of the task and the use of the backward-association.

Method

Design

The experiment was a 3 x 3 factorial design. The between-subjects factor was type of presentation. It had three levels: presentation in a word-pair list; presentation in a simple sentence context; presentation in a disrupted context . The within-subjects factor was list position. Its three levels were: the beginning of the list, taken to be the first six items; the end of the list, taken to be the last six items; the middle of the list, taken to be the remaining eight items.

Materials

The set of word-pairs used in the experiment was different from that used in Experiment 5 but apart from the change of language for cue words at testing all other aspects of presentation were as in Experiment 5. (Full lists are contained in the Materials Appendix). Four versions of the list were produced for learning purposes and all results refer to original learning order.

Subjects

59 subjects from the same mixed-sex comprehensive school, C6, took part in the experiment. None of the subjects had taken part in previous experiments. The experiment took place in the Summer term of the school year. Subjects had already experienced, therefore, nine months of formal French teaching when they began the experiment.

Procedure

Three experimental groups were formed. Group 1 consisted of 18 pupils (9 boys, 9 girls); Group 2 consisted of 21 pupils (10 boys, 11 girls); Group 3 consisted of 20 pupils (10 boys, 10 girls).

Membership of the groups related to the between-subjects factor in the following way. Group 1 learned in a normal word-pair list. Group 2 learned in a simple sentence

Table 5.3. Experiment 6.

Arrangement of Groups.

Group	Learning direction	Learning condition	Testing condition
1	English-French	List	French-English (context)
2	English-French	Context	French-English (context)
3	English-French	Disrupted list	French-English (context)

context. Group 3 learned in the disrupted format. All subjects were tested for comprehension in a simple context since the ability to recall words in a "natural" context was central to the experiment. The arrangement of the groups is as in Table 5.3. In other respects, the procedure was as in previous experiments.

Results and discussion

An analysis of variance was performed. The between-subjects factor was the modes of presentation; list position was the within-subjects factor.

List position had a significant effect on performance, $F(2, 112) = 21.21$, $p < 0.01$ (see Table 5.4. and Figure 5.3). In the first position in the list, 78.03% items were correctly recalled; in the second position, 58.75%; in the third position, 56.96%. Pairwise comparison (Tukey test) shows that there was a significant difference between items in the first position and other items in the list, $p < 0.01$ (see Table 5.4. and Figure 5.3.). These

Table 5.4. Experiment 6.

Mean percentage scores for items recalled.

Group	List position			Overall
	Beginning	Middle	End	
1	71.83	53.88	59.61	61.77
2	74.28	52.38	54.28	60.31
3	88.00	70.00	57.00	71.66
Overall	78.03	58.75	56.96	64.58

results indicate, therefore, a clear primacy effect. The difference between items in the middle of the list (58.75%) and items at the end of the list (56.96%) was not significant.

It was suggested previously that higher-ability learners develop a recall strategy based on the mode of presentation and that this is potentially inhibitory when there is a change from the expected mode of presentation as is the case for Group 1. When the task is comprehension in a sentence context, however, the advantages offered by the context compensate for any potential disadvantage. Here, in addition, the presence of a primacy effect suggests that the first few items in the list received special attention from subjects and that as a result memory for these items was particularly effective. The effect of list position was consistent across groups; there was no significant interaction between list position and modes of presentation, F (4, 112) = 1.65, $p > 0.16$.

As in Experiment 5, the mode of presentation did not have a significant effect on performance, F (2, 56) = 2.45, $p > 0.09$. The performances of list learners (61.77% items recalled), learners in context (60.31%), and learners in the disrupted context (71.66%) were not significantly differentiated. As suggested previously this could be because the use

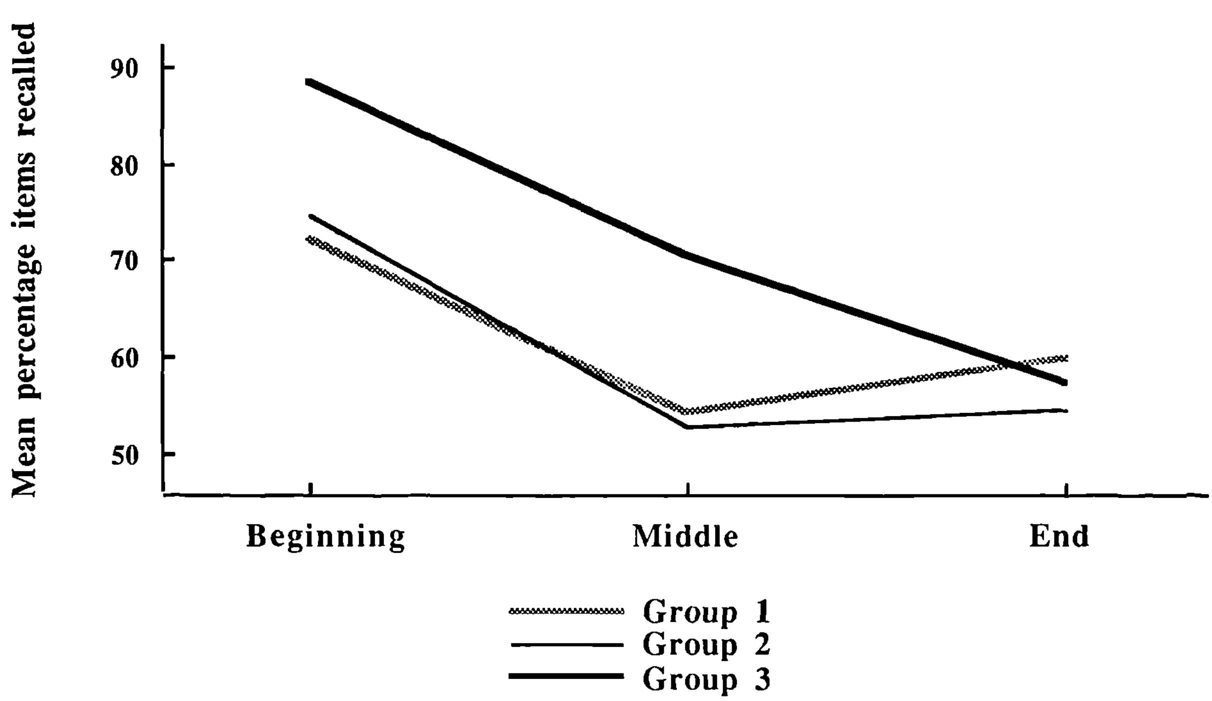

Figure 5.3. Experiment 6. Mean percentage scores for items recalled: Learning condition x List position. Group 1 learned in a normal word-pair list; Group 2 in simple sentence contexts; Group 3 in a disrupted format.

of context by higher-ability learners in the comprehension task obscures any effect of presentation. It could also be that subjects view the task as a word-pair learning task whatever the form of presentation (Gershman, 1970). Either way, a certain amount of support is indicated for the practice of presenting lists of word-pairs as a first step towards vocabulary learning.

Conclusion

There was, in these experiments, no evidence of a recency effect. This could be because recency effects are mainly acoustically based and therefore more effective when testing is immediate (see Glanzer, 1972; Glanzer & Cunitz, 1966; Postman & Phillips, 1965). On the other hand, with lower-ability subjects in Experiment 5 items at the end of the sentence were *less* well learned than other items in the list. This could be because subjects in Experiment 5 used a short-term learning strategy and the delay between learning and testing was too long for this to be effective (Baddeley, 1976; Jacoby & Bartz, 1972), or because they adopted an inappropriate recall strategy (Maskarinec & Brown, 1974). However, given the fact that items recalled seemed to be continuous from the first items in the list, the poor performance with the last few items appears to be the result of difficulty in covering the whole list at learning rather than due to any strategy adopted. Where the higher-ability subjects in Experiment 6 are concerned, it appears that subjects did cover more of the list but there was no significant difference between items at the end of the list and items in the middle of the list. In this case, the length of the delay between learning and testing could possibly account for there being no advantage for items at the end of the list.

There was a clear primacy effect in both experiments with items in the beginning of the list being recalled more successfully than items in any other position, and this despite the fact that the order at recall was different from the order at learning. It does seem that items at the beginning of the list received more attention from both lower-ability and higher-ability subjects than items elsewhere in the list.

In terms of the desirability of list learning, the results from Experiments 5 and 6 are inconclusive. If list position effects are taken to be undesirable because they set up list dependency, then there is some indication here that these effects do occur. On the other hand, list position effects appear to have done little to inhibit the performance of higher-ability learners in Experiment 6 and it could be argued that the list provides some rudimentary structure for lower-ability learners in Experiment 5 which would not otherwise be available.

It is interesting to note that there was no difference between groups, as such, in either experiment. It would appear that despite efforts to disguise the nature of the task, subjects treated all three forms of presentation in much the same manner as a word-pair learning task. It is difficult to envisage how the form of presentation could be further distorted than was attempted in these experiments, and the evidence seems to suggest that simple list presentation of word-pairs is just as effective as any other form of written presentation provided that higher-ability learners are in some way prevented from undertaking an inappropriate strategy for recall (for reasons discussed previously). This conclusion would be in line with that of Gershman (1970), Lado, Baldwin, and Lobo (1967), and Mishima (1966) who found little evidence of presentation effects despite manipulation of a range of possibilities.

In conclusion, there was some evidence in this experiment for list position and serial effects. Both lower-ability and higher-ability subjects appear to have given more attention to items at the beginning of the list than to items elsewhere in the list. Lower-ability learners were dependent on serial order whereas higher-ability learners appear to have adopted a more sophisticated strategy. The recency effect was not observed in either case.

CHAPTER 6

Word frequency and list position

Introduction

Up to this point, the main interest in the study has been in modes of presentation and their influence on learning. Next under consideration will be the notions of word difficulty and learnability which are topics of some interest in the second language learning literature, mainly in the context of seeking criteria for the order of teaching of new words. Three areas of discussion are word frequency, word category, and the relationship between L1 and L2 items to be learned. In this chapter, the influence of word frequency on learning and recall will be considered.

Word frequency: a review of the second language learning literature

In the second language learning literature, there is both a recognition of the importance of word frequency in determining word learnability and a certain amount of distrust of the reliability of objective word counts in this respect. Thus while Kellerman (1977, 1983) saw frequency as an indicator of learnability, in the sense that infrequent words are perceived to be psycholinguistically marked and in consequence difficult to learn, many writers concentrate on the problem of measuring word frequency in the first place. Nation (1987) made the point that many apparently common and useful words are not included in the first 1000 words in the count of Kucera and Francis (1967). He suggested that a whole range of factors in addition to objective word frequency needs to be taken into account. These factors include the range of situations in which a word can be used, the language needs of the individual, the availability of the word in the sense of subjective frequency, a word's coverage or the number of words which it can easily replace, its regularity, and its ease of learning. Polysemy presents another kind of problem where objective word counts are concerned. The Kucera and Francis (1967) word count does not take account of polysemy (even though slightly different forms of the same word are counted separately).

Curtis (1987) showed that while *last*, *box*, *store*, and *fire* are all high-frequency words, recommended for inclusion in first grade learning texts, *store* and *fire,* in the sense of *put away* and *dismiss* respectively, proved more difficult to understand than *last* and *box*. The explanation given is that these words require more knowledge about less familiar meanings than do the former two words. In this case, Kucera and Francis norms would be misleading.

However, alternatives have been difficult to come by. Richards (1974) argued that subjective estimates of word frequency are not reliable and that West's (1953) combination of subjective and objective criteria results in a list which is dated and idiosyncratic. On the other hand, objective counts are mutually inconsistent; are often counter-intuitive, especially where apparently common concrete words are concerned; and confuse frequency and utility. He referred instead to Michéa (1964) and the notion of availability. However, given the definition that: "An available word is a word which though not necessarily frequent, is always ready for use" (Richards, 1974, pp. 90-91), it is difficult to see that this can mean anything other than a subjective measure of frequency, which has already been rejected. The same must be said for his own preferred term of "familiarity" (Richards, 1970, 1971) which "refers to the subjective impression of words" (Richards, 1974, p. 77).

In summary then there does seem to be a strong element of unease in the second language learning literature about the use of objective word counts. This is based on a sense that subjective measures of word familiarity do not correlate well with objective measures of word frequency, particularly where apparently common nouns of low frequency are concerned.

A psychological approach to word frequency

From a psychological point of view, any discussion about the psychology of vocabulary list learning must take into account word frequency if only because word frequency effects are, in the words of Besner and McCann (1987), "probably the most powerful and ubiquitous effects in visual word recognition" (p. 202), or because, in the words of Whaley (1978), "it is now apparent that word frequency is by far the most powerful predictor" (p. 152) of response latencies to word stimuli. There are two stages in the discussion. First, there will be a review of the word frequency literature which will reveal a significant amount of disagreement about the nature and locus of the word frequency effect. Second, the implications of the word frequency effect will be considered in relation to the present domain.

The review of the word frequency literature is somewhat extensive, reflecting the complexity of the debate. However, it was felt that no purpose would be served by over-simplifying the issue since psychological theory can only be usefully applied when its full implications are worked through (see the discussion in Chapter 1).

The locus of word frequency effects

The word frequency effect refers to the fact that common words are responded to more quickly than uncommon words and there is no difficulty in pointing to experimental evidence for its influence. Humphreys, Besner, and Quinlan (1988) and Jacoby and Dallas (1981) reported word frequency effects in tachistoscopic reports; Forster and Chambers (1973) in word naming; Frederiksen and Kroll (1976), Rubenstein, Garfield, and Millikan (1970), Scarborough, Cortese, and Scarborough (1977) in lexical decision tasks; Savin (1963) in word identification with a degraded stimulus; Slowiaczek and Pisoni (1986) and Taft and Hambly (1986) in auditory lexical decision tasks; Monsell (1985) in categorisation tasks; and Rayner and Duffy (1986) in fixation times. Given its apparent ubiquity, it is no surprise that the word frequency effect has become something of a battleground over the years for competing models of the language processing system; each model must have its account of the effect, and it is often this account which is a distinguishing characteristic of that model. There has been a great deal of experimentation which in effect, if not in intention, has served to advance the cause of one model and hinder the advance of its competitors. The result is a mass of apparently conflicting data and a questioning in recent years of the usefulness of the notion of word frequency itself (as will be discussed below).

There are several reasons why data may conflict. One reason could be that like is not being compared with like because task demands, or materials, or subjects have not been sufficiently closely controlled; because variables which have not been taken into account confound results; or because methods of interpretation are based on assumptions which are not tenable. Examples of all three categories are to be found in the literature.

On the question of task demands, lexical decision tasks and naming tasks result in different latencies. Where materials are concerned, Seidenberg (1985a, 1989) pointed out that the notion of a "common" word has no absolute sense. Although word frequency can be averaged, as in the work of Kucera and Francis (1967), frequency is relative for any given individual and for groups of individuals. Thus Seidenberg found that among undergraduate subjects the fastest readers named lower-frequency words more quickly than the slowest readers read higher-frequency words. In this sense, "frequency" is a continuous variable and in using the term, reference must be made to the state of knowledge of subjects; likewise, "common" must be related to the nature of other items in test materials.

On the question of variables which have not been taken into account, it has to be said that even if word frequency can be isolated as an effect (see Gardner, Rothkopf, Lapan, and Lafferty, 1987; Whaley, 1978), many factors may closely interact with that effect. Thus Landauer and Streeter (1973) referred to grapheme and phoneme distribution patterns; Whaley (1978) pointed to word length, imagery, concreteness, meaningfulness, and letter frequency; the influence of the neighbourhood effect is discussed in Andrews (1989); Coltheart, Davelaar, Jonasson, and Besner (1977); Forster (1987); Forster, Davis, Schocknecht, and Carter (1987); Grainger (1990); Grainger, O'Regan, Jacobs, and Segui (1989); Laxon, Coltheart, and Keating (1988); Luce (1986), Luce, Pisoni, and Goldinger (1990); McCann and Besner (1987); Meyer, Schvaneveldt, and Ruddy (1974); Scheerer (1987).

Finally, where methods of interpretation are concerned, many predictions from models have been made on the basis of Sternberg's (1969) additive factors logic which depends on the identification of discrete stages in a serial process. It is not at all clear that this logic is appropriate if the language processing system is parallel and interactive rather than serial, and if discrete stages cannot be clearly identified (Besner and McCann, 1987,

p. 204). It is in the context of this somewhat confusing situation that the following discussion will take place, looking first at the evidence for the existence of a word frequency effect and secondly at the issue of subjective frequency.

Despite the apparently universal agreement about the existence of a word frequency effect, there are basically two quite different meanings given to the term. On the one hand, it means that the lexical representation is modified by use and familiarity with the result that more familiar words are accessed more quickly than less familiar words. Using the terminology of Besner and McCann (1987), it is criterion bias models which understand word frequency in this sense. On the other hand, there is an argument, represented by Balota and Chumbley (1984, 1985) and Besner and McCann (1987), that word frequency effects are the result of post-access processes and, in effect, distributed through the system rather than located in the lexical representation itself.

Criterion bias models derive from the work of Morton (1969). In such models, logogens or word-detectors are sets of neurons, with an adjustable threshold of activation, which collect evidence from both word input and context and which "fire" when the build-up of evidence causes them to exceed their threshold of activation. When the logogen representing a particular word fires then that word is said to have been identified. Criterion bias models account for the word frequency effect by the assumption that the threshold for a logogen is inversely related to word frequency; this means that less evidence is required for a logogen for a high-frequency word to be activated than for a logogen for a low-frequency word.

Various predictions have been made on the basis of criterion bias models and results have been conflicting (see Table 6.1). However, it has to be said that predictions have often been made by opponents of this class of model rather than by its proponents and there is not even agreement about what the models would predict. So, for example, it is Becker and Killion (1977), proponents of the Verification Model, who claimed that criterion bias models predict an interaction between word frequency and stimulus quality and then proceeded to show this not to be the case with reference to Becker (1976) and Stanners, Jastrzembski, and Westbrook (1975). However, not only did Norris (1984)

Table 6.1.

Predictions of criterion bias models compared with experimental evidence.

Variables	Predicted outcome	Outcome
stimulus quality and congruity	interaction	interaction (1)
word frequency and stimulus quality	interaction	additive (1)(2)(3)
		interaction (4)
word frequency and congruity	additive	additive (5)
		interaction (6)

References:

(1) Becker and Killion (1977) (2) Becker (1976) (3) Stanners, Jastrzembski, and Westbrook (1975)

(4) Norris (1984) (5) Schuberth and Eimas (1977) (6) Becker (1979)

show that word frequency does interact with stimulus quality under certain conditions but Besner and McCann (1987) contended that criterion bias models would predict an additive effect in any case. There is similar confusion surrounding the relationship between word frequency and contextual congruity. Schuberth and Eimas (1977) found that word frequency and congruity were additive; Becker (1979) finds that the variables interact.

There are various ways of coping with this unsatisfactory state of affairs. One approach is to call into question the word frequency effect itself as understood within this class of models. Balota and Chumbley (1984, 1985) argued that the lexical decision task, the main experimental paradigm used to locate the effect of word frequency on lexical access, is not an accurate measure of the effect; instead, the word frequency effect is produced by a combination of factors and located throughout the system rather than at lexical access. Balota and Chumbley (1984) found no evidence for the word frequency effect in a category verification task. Although this was not surprising where positive responses were expected, due to priming effects, the finding which was of particular interest was that in trials involving a negative answer there was no word frequency effect either. The category verification task, they argued, certainly involves lexical access; therefore the word frequency effect should have manifested itself. Although great care was taken to isolate any possible word frequency effect by controlling other variables (such as word length, instance dominance effects, category dominance effects) their finding was that frequency did not have a significant unique effect on reaction time in the category verification task although a lexical decision task, with the same set of words, did show a significant frequency effect. They also found that category dominance and instance dominance as factors were significant predictors of the outcome of the same lexical decision task. In other words, either information normally considered to be available only after lexical access appeared to influence lexical decision, or some other component of the lexical decision task, post access, is affected by semantic variables. Given that Chumbley and Balota (1984), James (1975), and Whaley (1978), also reported an effect of semantic variables on lexical decision, Balota and Chumbley (1984) suggested that the lexical decision task is not an effective way of studying lexical access. Forster (1979), Theios and Muise (1977), and West and Stanovich (1982) all suggested that the naming task is a better method of assessing lexical access. Balota and Chumbley therefore used the same word items in a naming task. Here word length and word frequency were both good predictors of latencies. The only impact of a semantic variable was category dominance unlike in the lexical decision task where both category dominance and instance dominance were predictors. This led Balota and Chumbley to the conclusion that instance dominance is the semantic factor which affects lexical decision in this case.

Balota and Chumbley (1984) argued, on the basis of results such as those just discussed, that the so-called word frequency effect is inflated in the lexical decision task because in this task word familiarity is particularly important at the decision stage as opposed to the access stage. Their explanation for lexical decision effects is based on the notion of familiarity/meaningfulness (FM) judgements rather than on word frequency. The FM value is based on the orthographic and phonological similarity of the input to actual words. According to the kind of items in the test, subjects can set a high or low criterion as a basis for their lexical decisions. Items which are clearly above or below the criteria are responded to rapidly. Other items are subjected to further examination and this takes time. The result is that items which are clearly nonwords, below the low criterion, are rejected quickly; items which are clearly words because they are familiar will be accepted rapidly; low-frequency words, on the other hand, will take longer than either high-frequency words or un-wordlike nonwords. In the category verification task, it is meaning rather than familiarity which is of importance; hence frequency will have little effect, *per se*, because lexical access is only a small part of the decision process. Although they claimed that they were not arguing that word frequency has no impact on lexical access, it is clear that Balota and Chumbley were attempting to reduce its importance; and by introducing the notion of judgements made on FM values, they were effectively suggesting an alternative to the criterion bias explanation for word frequency effects.

This impression is further strengthened by Balota and Chumbley (1985). Transferring their attention to the naming task, they argued that if naming involves lexical access, then testing response times over a range of delay intervals should "place" the word frequency effect. More specifically, any frequency effect should only be seen with delays up to 400 milliseconds and should disappear once the stimulus has been recognised. Frequency effects connected with response production should, on the other hand, be clear at much longer delays. Their results showed a frequency effect through the range of delays and up to the maximum of 1400 milliseconds. Some of this effect could have been because subjects chose to delay their response in anticipation of a delay. When the delay was varied in length, the word frequency effect was evident up to 900 milliseconds but not beyond. Taking the most likely explanation as being intervening rehearsal, Balota and Chumbley conducted further experimentation involving a distractor task; this time they found that the frequency effect at longer delays reappeared. Their conclusion was that the word frequency effect in the naming task has a substantial production component. Again, they insisted that they were "not arguing that word frequency has no impact on lexical access" (Balota and Chumbley, 1985, p.104) but were arguing against the attribution of the effect exclusively to lexical access. Despite their disclaimers, repeated again in Balota and Chumbley (1990), Balota and Chumbley (1984) and Balota and Chumbley (1985) taken together do present problems for the generally accepted notion of word frequency. In the first case, this is because they suggest an alternative explanation, based on familiarity, of so-called frequency effects; in the second case because they suggest that any lexical access account is confounded by production factors. In effect, they are saying that word frequency effects exist, but that there is no way of measuring them. It is difficult to see what this might mean.

Besner and McCann (1987) took the argument further. They collated a number of predictions relating to pattern distortion, word frequency, and task from four classes of models; these are: pattern analysing operations (e.g., Kolers, 1985); criterion bias models; serial search models (e.g., Forster, 1976); verification models (e.g., Becker, 1976, 1979). Table 6.2. lists these predictions. Besner and McCann's results showed that none of the models adequately cope with the data. Words were more impaired by pattern distortion in the lexical decision task than in the naming task and this was against the predictions of all the models. Pattern distortion impaired words more than nonwords in the lexical decision task and only the shape sensing model would make this prediction. The effect of pattern distortion on the latency for low-frequency words was greater than that for high-frequency words in the naming task which only verification models would predict. Finally, word frequency and pattern distortion were additive in the lexical decision task which was contrary to the predictions of the verification class of models and the shape sensing model. The conclusion of Besner and McCann (1987) seems appropriate, *mutatis mutandis*, to the word frequency debate in general and not just to this manifestation of it:

> "The results of these experiments can be summarised as follows; they were inconsistent with three out of four predictions for each of the four classes of models. It appears from these results that we do not, on the basis of the models discussed here, have an adequate theoretical understanding of why pattern

Table 6.2.

Predicted effects of format distortion on lexical decision and naming.

Factors	Word frequency x format		Lexical status x format	Format x task (words)
Task	Lexical decision	Naming	Lexical decision	Lexical decision and naming
Models				
Shape sensing	HF > LF	HF > LF	NW > N	Additive
Criterion bias	Additive	Additive	Additive	Additive
Serial search	Additive	Additive	Additive	Additive
Verification	LF > HF	LF > HF	N > NW	Additive

Note: > = longer latency than.

(See Besner & McCann, 1987, p. 207).

> distortion affects one task more than the other, or why pattern distortion is additive with word frequency in one task, but interacts with word frequency in another task." (Besner and McCann, 1987, p. 209).

Besner and McCann (1987) questioned the manner in which Sternberg's additive factors methodology is applied to the language process when the assumptions upon which Sternberg's methodology is based would not be accepted by the cascade model of the language process (McClelland, 1979), and the interactive activation model of McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982), (to which could probably be added the interactive cohort model of Marslen-Wilson, 1989). In these models, the letter identification process and the word identification process overlap and indeed interact. However, Besner and McCann (1987), McCann and Besner (1987), and McCann, Besner, and Davelaar (1988) were broadly in agreement with Balota and Chumbley and indeed were much more forthcoming about the implications of their argument for "traditional" accounts of the word frequency effect. Besner and McCann (1987) suggested that it was time "to abandon the assumption that word frequency effects reflect the operation of a unitary mechanism" and accept instead the assumption that "the effects of word frequency are distributed throughout the word recognition system" (p. 215). Indeed, they suggested at least three possible loci for these effects: the visual familiarity mechanism; word detectors; and connections between input and output lexicons. McCann and Besner (1987) claimed that their results show "that the phonological lexicon *is not sensitive to word frequency*" (p. 20; their italics) and it is not plausible to suggest that one lexicon (orthographic) is frequency sensitive whereas the other (phonological) is not; the clear implication is that both are not. McCann, Besner, and Davelaar (1988) concluded that it is "no longer tenable to regard the effects of word frequency in various tasks as reflecting something unitary, such as lexical access operations" (p. 705). As discussed previously, the rejection of the lexical decision task as a measure of word frequency, and the conclusion that word frequency can no longer be seen as a largely unitary effect, are both problematical, particularly for criterion bias models.

Monsell (1991) conducted a review of the evidence to date and made a robust defence of the traditional locus of word frequency effects, that is in lexical access, within a connectionist or parallel distributed processing (PDP) framework. An important preliminary to his detailed discussion is the point that PDP learning models (e.g., McClelland, Rumelhart, & the PDP Research Group, 1986) are different from other types of language processing model in one crucial respect; they account for the system's acquisition of new words, and in so doing also account for the frequency effect. Search and activation-verification models are both, in effect, serial search models; that is, word frequency is accounted for by search order. In connectionist learning models to describe a word as having high- or low-frequency is to describe the number of exposures of that item to the system. In other words, connectionist explanations of the word frequency effect are inherent in the model and not arguments constructed expressly for the purpose of explaining the effect. This means that connectionist explanations are contained within a framework which has widespread application for psychological phenomena whereas the relationship of other kinds of explanation to wider issues of mechanisms and processes is rarely covered. The main point at issue is whether or not familiar words are more easily identified than unfamiliar words and whether the effects of word frequency are located, mainly at least, at the identification stage of the language process.

For a definition of "identification" it is necessary to make a distinction, in connectionist models, between identification and transcoding. In the simplest form of connectionist network (see Figure 6.1a) the system generates an appropriate output to a given input; it is engaged in a transcoding task. However, a more complex system (see Figure 6.1b) will also need to learn to identify patterns as well as transcode them. That is, it will need to recognise or identify elements which co-occur in input patterns. To do this, it will form connections between input units which then represent those patterns. The more often a pattern is input, the more stable its representation becomes because units previously coactivated tend to activate each other, and units not previously coactivated tend to inhibit each other. Due to the connections between input and output units, better (more rapid) identification necessarily results in better (more rapid) transcoding. This account implies that the effect of word frequency, that is the effect of the number of times the system has been exposed to a given input, is a property of the model which is necessary rather than contingent. In practice, the more developed connectionist models (e.g., Seidenberg & McClelland, 1989; Sejnowski & Rosenberg, 1986) incorporate so-called hidden units as illustrated in Figure 6.1c. It is the hidden units, not responsible directly for either input or output, which detect and store useful input patterns. These patterns range in size, it is presumed, from bigrams through to whole word patterns. Within this model, lexical decision can be made on the basis of identification without transcoding into output units. Here again, though, word frequency effects would be accounted for by the patterns stored in the hidden units. Frequency effects are inherent in the system and are accounted for by

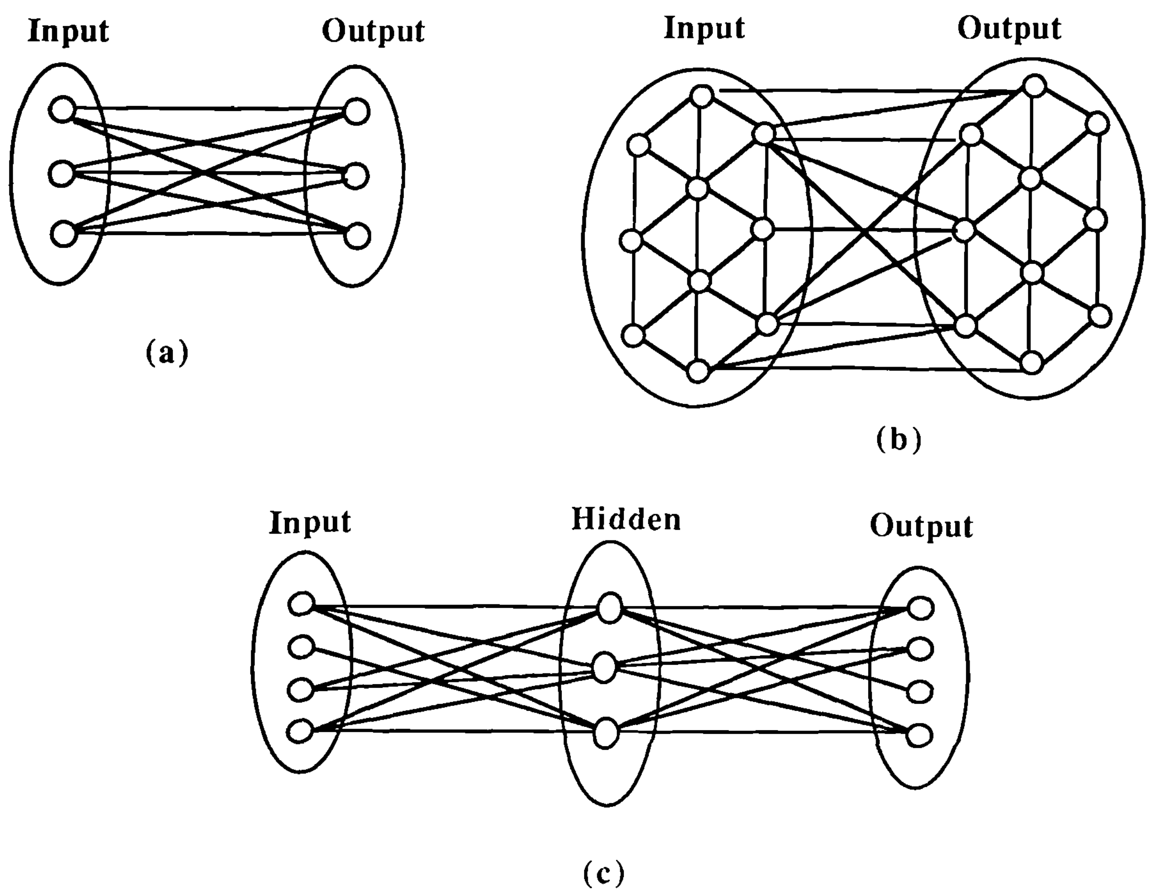

Figure 6.1. Three PDP architectures for lexical transcoding (see Monsell, 1991, p. 157)

the learning process itself. Identification and transcoding are necessarily frequency sensitive in connectionist models of this type. Monsell (1991) acknowledged that while (or perhaps because) "conventional" word-detector, serial search, and activation-verification models are all built round an identification process which is sensitive to frequency, they have tended to ignore possible post-identification word frequency effects (see also McRae, Jared, & Seidenberg, 1990). This is to some extent associated with the predominance of the use of the lexical decision task as a supposed measure of lexical access and therefore as a direct measure of lexical identification. As has already been discussed, in recent years

Balota and Chumbley (1984, 1985), Besner and McCann (1987), McCann and Besner (1987), McCann, Besner, and Davelaar (1988), have all challenged the reliability of measures of the word frequency effect which are confined within what Monsell (1991) called the identification and transcoding stage of the language process. The result of this challenge has been to leave "very *little* effect of frequency to be localized in either identification or access to meaning" (Monsell, 1991, p. 167). The bulk of Monsell's argument is taken up with a refutation of this conclusion.

Monsell argued that Balota and Chumbley's FM dimension, used to explain word frequency effects in the lexical decision task, means nothing unless it is correlated with frequency in the process of obtaining an FM value. Balota and Chumbley (1984) stated that a value on the FM dimension "is based primarily on its orthographic and phonological similarity to actual words" (p. 352). Unless this process is equivalent to evaluating the frequency of the input string, then there can be no place in the process where frequency has an effect because it is clear from their account that although items not exceeding either the high or low criteria are subject to further analysis, frequency evaluation is not part of this process. This being so, the only genuine alternative to the standard account of word frequency effects in the lexical decision task is one based on a strictly extra-lexical source.

Monsell went on to reconsider the evidence for and against word frequency effects from categorisation tasks and naming tasks. Categorisation is a difficult task to control since any computing of membership properties can clearly result in increased latencies and therefore in the swamping of possible frequency effects. Accordingly, Monsell, Doyle, and Haggard (1989) used a different kind of categorisation task from that used by Balota and Chumbley (1984). It was based on what they took to be prestored and highly available material, namely, the classification of items as a "Person" or an "Inanimate Thing". Their results were in direct conflict with Balota and Chumbley (1984) showing strong and significant effects for frequency and word class, and no interaction between either of these effects and task. In other words, the effects of frequency were consistent over tasks. The explanation of Monsell, Doyle, and Haggard (1989) was that the task is different in the two experiments. Balota and Chumbley were effectively asking subjects to conduct a semantic comparison over a number of attributes and this task was inherently more difficult than the Monsell, Doyle, and Haggard (1989) task. Thus the complexity of the post-transcoding decision effectively masked the effects of word frequency earlier in the process. Whether or not this is so, the more general point to be made is that the absence of strong frequency effects in a particular instance does not allow the conclusion that lexical access is not frequency sensitive because this conclusion would leave the Monsell, Doyle, and Haggard (1989) result without explanation.

Word frequency effects are more marked in the lexical decision task than they are in the naming task (see Balota & Chumbley, 1984; Forster & Chambers, 1973; Frederiksen & Kroll, 1976). However, Monsell (1991) argued that this generalisation conceals what he calls a second order generalisation: the size of the frequency effect in naming depends on the amount of lexical and sublexical transcoding from orthography to sound involved. The more transcoding there is at lexical level, the larger the frequency effect. Whether lexical and sublexical transcoding from spelling to sound are carried out by separate "pathways" (see Coltheart, 1985), or by a single transcoding module (see Seidenberg & McClelland, 1989; Shallice & McCarthy, 1985), the assumption is that information from both these sources is pooled until the system has sufficient information to generate pronunciation. This being so, it is clear that sublexical transcoding will not be sensitive to word frequency, although it may be sensitive to frequency patterns in sub-lexical components; lexical transcoding, on the other hand, will be sensitive to whole-word frequency. It is to be expected therefore that, in general, naming, which will involve pooled information rather than "pure" lexical information, will show less of a frequency effect than the lexical decision task. Again, where so-called regular words are concerned the two pooled sources of information will be conveying compatible "messages" and a response may be possible before full lexical identification has taken place. In the case of exception words, the two sources of information may well be in conflict, and this will take time to resolve. It would be expected, therefore, that word frequency effects would be more marked with exception words than with regular words. Bringing these points together, Monsell, Doyle, and Haggard (1989) conducted a naming experiment which involved a mixture of regular and exceptional (stress-final) words. Although overall the word frequency effect was much less than that observed for a comparable lexical decision task, the word frequency effect for the exception words was equal to that for the same words in a lexical decision task. Again, these results are in conflict with those of Balota and Chumbley (1985). At the very least, Balota and Chumbley's conclusion cannot be generalised if the data are to be accounted for.

The final possibility considered by Monsell (1991) was that of a genuinely extra-lexical source of familiarity as an explanation of the word frequency effect in lexical decision tasks. The most likely candidates here are those suggested as explanations for repetition effects: response learning; and recovery of an episodic trace. However, as discussed in Chapter 3, Monsell (1985) showed that learned responses and episodic traces are a highly unlikely main source of repetition priming and it is therefore implausible that they are a major source of frequency effects in lexical decision tasks either.

Monsell's conclusion was that word frequency effects observed in lexical decision tasks and naming tasks are not task specific. Although it is not possible to say that word frequency effects are restricted to the identification stage, there are no grounds for saying that word frequency effects do not occur at the identification stage. Indeed, in PDP models the distinction between identification and retrieval stages, and the logic based on this distinction, is of dubious relevance.

Subjective word frequency

A subject of increasing interest in the context of the discussion on word frequency effects is the relationship between "objective" word frequency, derived from samples of written English (as for example Kucera & Francis, 1967, whose work has been used in this thesis), and "subjective" word frequency or the actual available reading vocabulary of a subject. Clearly, if there is a great discrepancy between these two measures, results derived from experiments using objective word frequency as a baseline may be of questionable validity. The issue has been approached in several ways.

Gernsbacher (1984) made a distinction between word familiarity and word frequency. Whereas the familiarity of a word has been shown to affect speed and accuracy of recognition, familiarity has normally been taken to be synonymous with word frequency derived from three standard indices: Carroll, Davies, and Richman (1971); Kucera and Francis (1967); Thorndike and Lorge (1944). However, Landauer and Streeter (1973) pointed out that there may be factors affecting the reliability of using these indices for low-frequency (as opposed to high-frequency) words. These factors include letter and phoneme distribution (Landauer & Streeter, 1973); concreteness (Paivio, Yuille, & Madigan, 1968); polysemy (Glanzer & Bowles, 1976; Reder, Anderson, & Bjork, 1974; Schnorr & Atkinson, 1970). Experimental evidence seems to suggest that these variables are relatively unimportant for high-frequency words; where low-frequency words are concerned, interactions between low-frequency words and these variables have been inconsistent and paradoxical. As others have also pointed out (e.g., Gordon, 1985; Nation, 1987; Wallace, 1982) printed word frequency indices suffer from potential drawbacks: they are by definition based only on samples from the printed word; they are dated; and discrimination at the bottom end of the frequency scale is sparse or non-existent. (In relation to the final point, it is worth noting that in Kucera & Francis, 1967, for example, around 45% of the total corpus share a frequency of 1 per million). On the other hand, Gernsbacher (1983) found that experiential familiarity was both a reliable indicator of performance with low-frequency words and provided a wide range of discrimination within the word frequency category of 1 per million.

Gernsbacher (1984) explored these inconsistencies using where possible materials used in previous experiments. Her argument was that bigram frequency effects, concreteness effects, and effects from polysemy can all be accounted for by experiential familiarity. In other words, when subjects' prior familiarity with the materials is controlled, effects attributed to the three variables disappear. Her conclusion was that word frequency is a reliable indicator of the availability of high-frequency words and experiential familiarity is a reliable indicator of the availability of low-frequency words.

Gordon (1985) pointed to the evidence for an overall adequate correlation between experiential familiarity and objective frequency counts (e.g., Carroll, 1971; Shapiro, 1969; Underwood, 1966). However, Howes (1954) found that familiarity ratings correlated better with performance for low-frequency words than did word frequency counts. Carroll, Davies, and Richman (1971) also found a good correlation between familiarity ratings and performance with low-frequency words even when objective frequency was partialled out. A possible drawback with these studies was that they only used a small set of words, the length and inflectional complexity of the words were not controlled, and the range of frequencies was not controlled. Gordon's experiments were designed to address these deficiencies and resulted in the following conclusions. The rating reliability among subjects, that is the extent of agreement between them, was higher for low-frequency words ($r = +0.76$) than for high-frequency words ($r = +0.54$). Overall, the correlation between the word frequency measures of Kucera and Francis (1967) and familiarity measures was satisfactory ($r = +0.79$) although it was not as great as that between Kucera and Francis on the one hand and Carroll, Davies, and Richman (1971) on the other. This apparent agreement obscures some important differences. Thus when the list of words examined was made up of mixed low-frequency and high-frequency words, inter-subject estimates of low-frequency words correlated well ($r = +0.85$) but collectively they differed significantly from word frequency values ($p < 0.01$). In other words, word frequency counts for low-frequency words are somewhat problematical. According to Gordon (1985), familiarity measures account well for lexical decision task data with low-frequency words. He is in agreement with Gernsbacher (1984) that the summed frequency values of Kucera and Francis (1967) account for only 7.5% of the latency variance for low-frequency words whereas familiarity measures account for 29-30% of the variance; the difference is significant ($p < 0.05$). It is worth noting in passing that in relating his findings to different models of lexical access, Gordon (1985) made a clear distinction between his use of the term "familiarity" and that of Balota and Chumbley (1984); indeed, Gordon insisted on a role for word frequency in lexical access in opposition to the argument of Balota and Chumbley (1984).

In their review of the topic, Graves, Ryder, Slater, and Calfee (1987) found a wide range of data reported over the years (see Table 6.3). In the light of evidence of this kind, some have rejected objective word frequency as an indicator of either word difficulty or of

Table 6.3.

Survey of correlations between word frequency and word knowledge.

Graves, Ryder, Slater, and Calfee (1987)

Source	**Correlation: word frequency and word knowledge.**
Davis (1944)	0.05 - 0.19
Guilford and Zimmerman (1948)	0.87
Kirkpatrick and Cureton (1949)	0.47 - 0.56
Kibby (1977)	0.22 - 0.58
Graves, Boettcher, Ryder, and Peacock (1980)	0.44
Graves, Ryder, and Slater (1983)	0.46 - 0.94

subjective word frequency (e.g., Davis, 1944; Goodman & Bird, 1984); however others such as Carroll (1972) have argued for its usefulness, and Anderson and Freebody (1981) pointed to its success in predicting the likelihood of a subject's knowledge of a word. Graves, Ryder, Slater, and Calfee (1987) attempted to clarify the issue by comparing six metrics of frequency derived from Anderson and Freebody (1981), Carroll (1968), and Carroll, Davies, and Richman (1971). The work of Carroll would suggest that the relationship between word frequency and word knowledge is logarithmic rather than arithmetic. The work of Anderson and Freebody suggested that a good predictor of word knowledge is not individual word frequency but family frequency based on the total frequency of the root word and all its compounds and derivatives. Six metrics in all were tested. Two were arithmetic, applied to individual words and families of words. The other four metrics were logarithmic. Two were straightforwardly logarithmic, again applied to individual words and families of words. The final two used logarithmic groupings but with a logitized scale (created by applying logit transformations to the proportions of right answers) applied to individual words and families of words. The conclusion of Graves, Ryder, Slater, and Calfee (1987) was that provided words are grouped using logarithmic groupings, and provided groups of words rather than individual words are considered, then objective word frequency is an excellent predictor of subjective word knowledge. Grouping in families, as such, is only important when arithmetic groupings are used; with logarithmic groupings, family frequencies are no better as indicators than groupings of individual words.

Conclusion

It has not been practicable in the present study to establish measures of subjective word frequency though this may be more important with children than with adults. As Carroll (1971) pointed out, familiarity measures (such as his Subjective Frequency Index) are a measure of perceived frequency and to that extent are potentially more useful than word frequency counts as such. On the other hand, familiarity measures are themselves not without problems since they depend on conscious activity on the part of subjects and are thus open to influences not present in objective word counts. It must be borne in mind, nevertheless, that what word counts measure and what subjects know, and particularly relatively young subjects, may be somewhat different entities.

It is clear that the debate about word frequency effects is complex and it is difficult to anticipate or predict the effects of word frequency in a complex task such as vocabulary learning when automatic and strategic processes are to be taken into account. The experiments to follow will add to the data which need to be taken into account in the discussion and they will also give an indication of whether objective word frequency is a reliable indicator of performance in this domain.

EXPERIMENT 7

Experiment 7 was designed to examine the role of word frequency in a second language generation task. Despite the strong evidence for word frequency effects across a range of tasks involving lexical access, it is difficult to predict the effect of word frequency on subject performance when a more complex process than word identification or word naming is involved because of the range of variables involved.

If the word frequency effect results from a state of the lexical representation, then it is possible that a word-pair has a lexical status whose frequency is determined by the frequency of the L1 component. If this is the case, then word frequency could have a direct effect even in the generation task. The more frequent L1 component of the word-pair would be more quickly accessed by an automatic process, and word-pair completion would take place in much the same way as with recall of individual words.

It is also possible to envisage an indirect effect of word frequency on recall whether or not the word-pair has some kind of lexical status. If word frequency has its effect on lexical access (see Marslen-Wilson, 1989; McClelland, 1985, 1986; McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Seidenberg, 1989; Seidenberg & McClelland, 1989) and is connected with the state of the logogen, then it could be that subjects would be more inclined to learn word-pairs where the English component was more frequent simply because these words would be more rapidly accessed. At testing, the frequency of the English word cue might also influence performance in the sense that more attention would be given to attempting to recall the L2 equivalent for the readily accessed high-frequency word. Conversely, ease of access to the English component might encourage subjects to consider a word-pair easily learned and this could result in their giving less attention to it.

If the word frequency effect is the product of a familiarity judgement (Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984), or of a word production process, and therefore having very little relationship to lexical access (Besner & McCann, 1987; McCann & Besner, 1987; McCann, Besner, & Davelaar, 1988), it is difficult to see how word frequency could affect performance in the generation of L2 items when neither a familiarity judgement nor time constraints apply.

Thus although the level of description possible with this experiment is too high to allow detailed predictions about the effect of word frequency in this domain, it should be possible to draw conclusions about whether or not the influence of objective word frequency on performance is worth pursuing or whether the issue needs further specification.

The experiment was also designed to consider further the effect of list position at learning on performance and its importance relative to word frequency. Previous experiments averaged word frequency over the whole list or sections of the list; here a more accurate estimate of list position influences should be possible due to a tighter control of word frequency and presentation of materials in word-pair lists. In addition, in Experiments 5 and 6, all subjects were tested in a context. In Experiment 7, some subjects were tested in a list, thus allowing a comparison to be made between list position and serial effects in list testing, and list position and serial effects in context testing.

Method

Design

The experiment had a 2 x 2 x 3 factorial design. The between-subjects factor was the test condition. It had two levels: generation in a list; and generation in a context. The within-subjects factors were word frequency and list position. There were two levels of word frequency: high-frequency words and low-frequency words (as defined below). List position had three levels: the beginning of the list, taken to be the first six items; the end of the list, taken to be the last six items; the middle of the list, taken to be the remaining eight items.

Table 6.4. Experiment 7.

Arrangement of groups.

Group	**Learning condition**	**Test condition**
1	English-French word-pair list	List: generation
2	English-French word-pair list	Context: generation

Materials

A list of 20 items was presented for learning. Half of the words were of high-frequency (frequency range 207-589 occurrences per million; mean, 316.5 occurrences per million). The other half were of low-frequency (frequency range 2-37 occurrences per million; mean, 17.5 occurrences per million). All measures of word frequency were taken from Kucera and Francis (1967). High-frequency and low-frequency words were intercalated. Full lists are to be found in the Materials Appendix.

Subjects

The pool of 59 middle-ability subjects, aged 11-13, came from School C7, a mixed-sex comprehensive school; none of the subjects had taken part in previous experiments. The experiment took place in the Summer term of the school year. Subjects had already experienced, therefore, nine months of formal French teaching when they began the experiment.

Procedure

Two experimental groups were formed to represent the between-subjects factor. Group 1 consisted of 29 pupils (14 boys, 15 girls); Group 2 consisted of 30 pupils (15 boys, 15 girls). Subjects in Group 1 were tested in a word-pair list in randomised order. Subjects in Group 2 were tested in sentence contexts in randomised order. The arrangement of the groups is as in Table 6.4. In other respects, the procedure was as in previous experiments.

Results and discussion

An analysis of variance was performed. The mode of testing (group membership) was the between-subjects factor. Within-subject factors were word frequency and list position. Mean percentage scores are shown in Table 6.5.

Table 6.5. Experiment 7.
Mean percentage scores for items recalled.

	High frequency	Low frequency	Overall
Group 1	48.25	38.93	43.59
Group 2	49.22	36.58	42.90
Overall	48.73	37.75	43.24

Word frequency had a significant effect on performance, $F\ (1, 57) = 19.86, p < 0.01$. The percentage mean of items recalled for high-frequency words was 48.73% and for low-frequency words 37.75% (see Table 6.5). In broad terms, therefore, the frequency of the English components of the word-pairs appears to have influenced subjects' performance.

List position also had a significant effect, $F\ (2, 114) = 4.79, p < 0.05$. As Table 6.6. shows, percentage scores of items correctly recalled for the three list positions were

Table 6.6. Experiment 7.

Mean percentage scores for items recalled: List position and word frequency.

Frequency	List position			Overall
	Beginning	Middle	End	
High	46.82	39.44	59.95	48.73
Low	32.75	42.71	37.81	37.75
Overall	39.78	41.07	48.88	43.24

39.78% for the beginning of the list; 41.07% for the middle of the list; and 48.88% for the end of the list. A pairwise comparison (Tukey test) shows that the score for the end of the list was significantly higher than that for the beginning of the list and for the middle of the list, $p < 0.05$. Thus there is evidence here for a recency effect which had not been seen in previous experiments. There was no significant difference between recall of items from the beginning of the list and items from the middle of the list.

There was an interaction between word frequency and list position, F (2, 114) = 11.68, $p < 0.01$ (see Table 6. 6 and Figure 6. 2). A pairwise comparison (Tukey test) shows that high-frequency words from the beginning of the list (46.82%) were more successfully recalled than low-frequency words from the beginning of the list (32.75%), $p < 0.05$, and high-frequency words from the end of the list (59.95%) were more successfully recalled than all other items, $p < 0.01$. However, in the middle of the list, the recall of low-frequency words (42.71%) was not significantly different from from that of high-frequency words (39.44%).

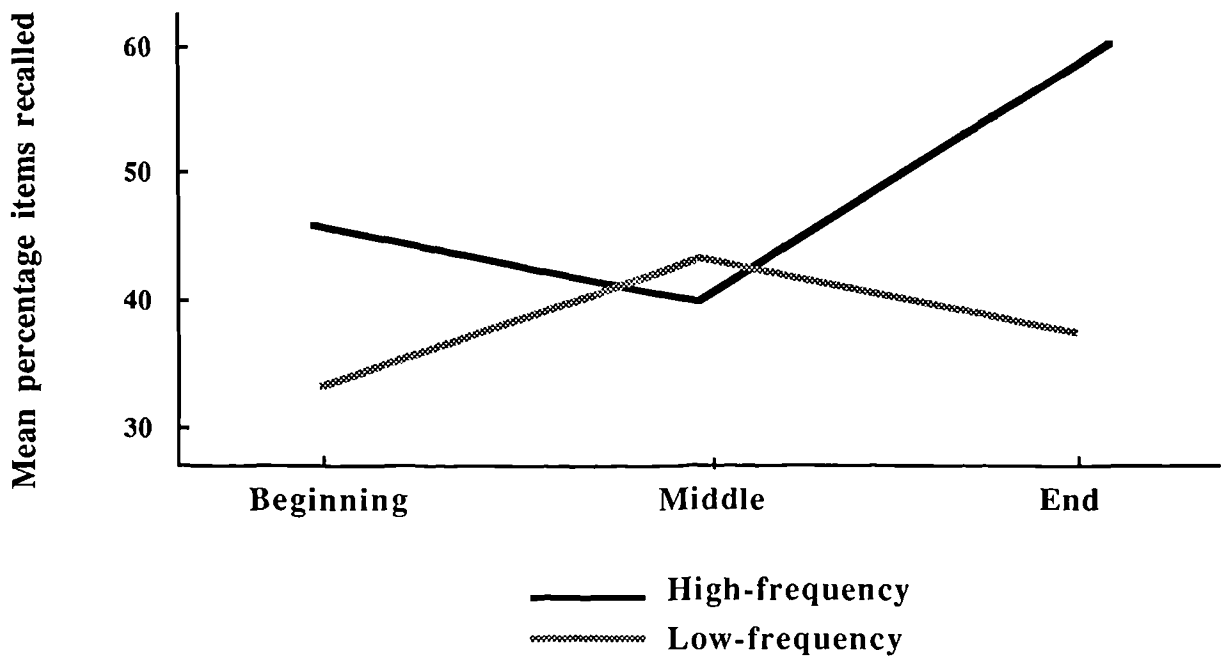

Figure 6.2. Experiment 7. Interaction: Word frequency x List position.

The difference in performance due to test condition was not significant, $F\ (1, 57) = 0.01, p > 0.91$. Group 1, tested in a list, had a mean percentage score for items recalled of 43.59%. Group 2, tested in a context, averaged 42.90%. Whatever serial effects occurred, they applied equally to being tested in a list and being tested in a context. It is worth noting, with respect to previous experiments, that with an overall performance of 43.24% items recalled, subjects in Experiment 7 fall somewhere between the higher-ability and lower-ability categories. It could be therefore that learning undertaken transferred equally well to list testing and context testing because neither group of subjects, as middle-ability learners, developed a strategy which was heavily task-dependent and therefore the use of a context at testing was not particularly disruptive.

There are three main points of interest here. There was a clear effect of word frequency in this experiment although the level of description of the experiment does not make it possible to say whether this was due to automatic processes or strategies adopted by subjects. There was a list position effect but it was noticeably different from that observed in Experiments 5 and 6 in that there was a recency effect but no primacy effect. Finally, there was a significant interaction between word frequency and list position with items in the middle of the list being treated differently from items at the beginning and end of the list.

In relation to the word frequency effect, if subjects learned those items to which they responded more quickly, that is the high-frequency items, then a learning strategy could have been based, in part at least, on frequency. It would be as if two lists were perceived - a high-frequency list and a low-frequency list.

Where list position is concerned, perhaps the first point to note is that the effect is not as clear as in the previous experiments, being significant only at $p < 0.05$ rather than at $p < 0.01$. Nevertheless, there is a significant recency effect. In Experiment 5, with lower-ability learners, subjects appear to have worked their way serially through the list with special attention being paid to items at the start of the list and with performance tailing off with later items. In Experiment 6, with higher-ability learners, subjects still gave particular attention to items at the beginning of the list but otherwise their attention was spread evenly throughout. In Experiment 7, subjects appear to have concentrated on items at the end of the list with their attention otherwise spread evenly (if more thinly than for high-ability subjects in Experiment 6) throughout. There is no obvious explanation for this other than to say that concentration on either the beginning or the end of the list appears to be a "natural" strategy adopted by learners.

Where the interaction between word frequency and list position is concerned, it should be noted that in all three relevant experiments, items in the middle position in the list were never the least successfully recalled. It seems, therefore, that subjects treated items in the middle of the list differently from those in the beginning and end positions which are normally associated with the clearest list position effects. It could be that in the absence of obvious serial cues for remembering these items, subjects use a different kind of processing. Whatever the form this processing takes, it would appear that this it is to some extent independent of the effect of word frequency in addition to being independent of list position effects.

EXPERIMENT 8

Experiment 8 was designed to examine the role of word frequency and serial effects in a second language comprehension task under two different conditions.

Here, the application of the word frequency effect would appear to be more straightforward. If the state of the logogen is a product of word frequency (see Marslen-Wilson, 1989; McClelland, 1985, 1986; McClelland & Elman, 1986; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982; Seidenberg, 1989; Seidenberg & McClelland, 1989) then it is possible to envisage word frequency as having a facilitatory effect on performance when the target is an English word. Put simply, less information would be needed to access a higher-frequency English target word than it would a lower-frequency word.

As before, if the word frequency effect is the product of a familiarity judgement (Balota and Chumbley, 1984, 1985; Chumbley and Balota, 1984), or a word production process, and therefore having very little relationship to lexical access (Besner and McCann, 1987; McCann and Besner, 1987; McCann, Besner, and Davelaar, 1988), it is difficult to see how word frequency could affect performance in comprehension any more than it would do in the generation task with the exception that a familiarity judgement might be made concerning subject-generated L1 response candidates.

Where list position and serial effects are concerned, it might be anticipated that serial order effects would be more in evidence than in the previous experiment if only because vertical links (that is links between items in the cue position and between items in the response position) are more likely to be established between the L1 items in the list, which are here the target, than between the L2 items in the list. This is because L1 items

Table 6.7. Experiment 8.

Arrangement of groups.

Group	Learning condition	Test condition
1	Word-pair list	List: comprehension
2	Word-pair list	Context: comprehension

are to some extent familiar to subjects whereas all L2 items are unfamiliar. On the other hand, if there were a high level of performance, as in Experiment 6, this effect could be more obscured in the list test condition by an overall spread of recall, and in the context test condition because any such influences would be obscured by the enabling effect of the context at testing.

Method

Design

The experiment had a 2 x 2 x 3 factorial design. The between-subjects factor was the test condition. It had two levels: comprehension in a list; and comprehension in a context. The within-subjects factors were word frequency and list position. There were two levels of word frequency: high-frequency words and low-frequency words (as defined below). List position had three levels: the beginning of the list, taken to be the first six items; the end of the list, taken to be the last six items; the middle of the list, taken to be the remaining eight items.

Materials

A list of 20 items was presented for learning. Half of the words were of high-frequency (frequency range 207-589 occurrences per million; mean, 300.1 occurrences per million). The other half were of low-frequency (frequency range 2-37 occurrences per million; mean, 20.3 occurrences per million). All measures of word frequency were taken from Kucera and Francis (1967). High-frequency and low-frequency words were intercalated.

Subjects

A pool of 58 subjects, aged 11-13, from School A8 the mixed-sex comprehensive school took part in the experiment. The experiment took place in the Summer term of the school year. Subjects had already experienced, therefore, nine months of formal French teaching when they began the experiment.

Procedure

Two experimental groups were formed to represent the between-subjects factor. Group 1 consisted of 28 pupils (13 boys, 15 girls). Group 2 consisted of 30 pupils (15 boys, 15 girls). Subjects in Group 1 were tested in a list in randomised order. Subjects in Group 2 were tested in sentence contexts in randomised order. The arrangement of the groups is as in Table 6.7.

In other respects, the procedure was as in previous experiments.

Results and discussion

An analysis of variance was performed. Modes of testing (group membership) was the between-subjects factor. Within-subject factors were word frequency and list position. Word frequency had a significant effect on performance, but in the opposite

Table 6.8. Experiment 8.

Mean percentage scores for items recalled.

	High frequency	Low frequency	Overall
Group 1	50.78	58.26	54.52
Group 2	43.84	57.98	50.90
Overall	47.31	58.12	52.71

direction from the one anticipated, $F\,(1, 56) = 20.73, p < 0.01$ (see Table 6.8). The mean percentage of high-frequency words recalled was 47.31% and the mean percentage for low-frequency words was 58.12%. There are a number of possible explanations for this surprising result.

One possibility is that subjective frequency is quite different from objective frequency for this group of subjects. However, the reversal of the expected effect is so clear-cut that this can only be put forward as a partial explanation; it is not just a case that high-frequency words were not advantaged, but they were clearly disadvantaged in subject performance.

It could be that subjects simply paid less attention at learning to those items which were readily accessed and concentrated instead on word-pairs with less familiar English components. Measured by performance, subjects in Experiment 8 were clearly of higher-ability with an overall performance of 52.71% items recalled. Higher-ability subjects in experiments reported here are characterised in part by their use of strategies; concentration on the ostensibly more difficult items could be one such strategy. In more general terms, Gorman (1961) reported that words of very high-frequency were less well recognised than words of lower-frequency. He suggested that this was because there is more candidate

Table 6.9. Experiment 8.

Mean percentage scores for items recalled: Word frequency and list position.

Frequency	List position			Overall
	Beginning	Middle	End	
High	43.13	56.19	42.61	47.31
Low	54.02	58.18	62.18	58.12
Overall	48.57	57.18	52.39	52.71

interference with high-frequency word recognition; the comprehension task has much in common with the recognition task in that candidates can be generated prior to recognition.

It could also be that some other factor affected performance which had not been taken into account; word category is one such factor which will be discussed in the next chapter. Whatever the explanation, it is evident that objective word frequency counts cannot, on this evidence, be taken as a straightforward predictor of ease of learnability of word-pairs measured by recall.

The effect of list position fell short of significance, $F\ (2, 112) = 2.74, p > 0.06$ (see Table 6. 9). There were 48.57% of items correctly recalled from the beginning of the list as compared with 57.18% from the middle of the list, and 52.39% from the end of the list. Recall in this case, therefore, was evenly spread throughout items in the list. It was argued previously that concentration on the beginning or end of the list is a "natural" strategy adopted by learners. It is clearly a strategy available to learners, but these results show that

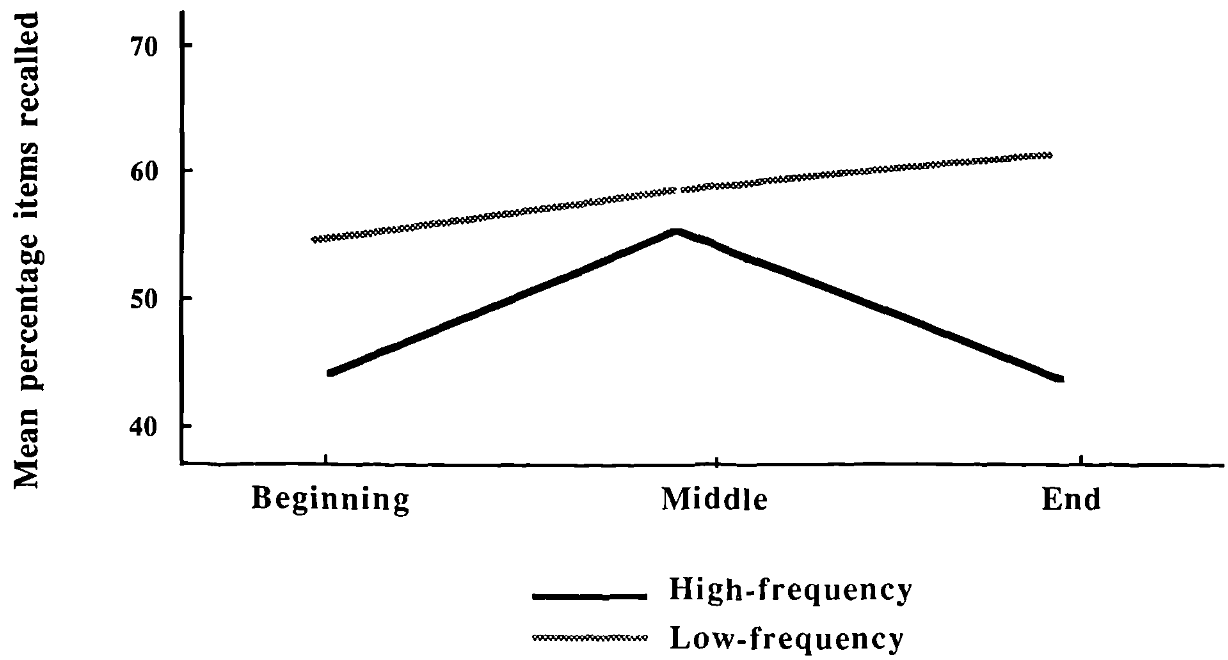

Figure 6.3. Experiment 8. Interaction: Word frequency x List position.

it is a strategy which need not be used. Interestingly, though, variations within the list positions associated with word frequency occurred only at the beginning and end of the list. This is shown by the significant interaction between word frequency and list position, $F\ (2, 112) = 3.92, p < 0.05$ (see Table 6.9. and Figure 6.3). Low-frequency items at the end of the list (62.18%) were better recalled than high-frequency items at the end of the list (42.61%), $p < 0.01$. Items in the middle of the list again appear to have been treated differently from items elsewhere since high-frequency items at the beginning of the list (43.13%) were less well recalled than either high-frequency (56.19%) or low-frequency items (58.18%) in the middle of the list, $p < 0.05$.

The difference in performance due to test condition was not significant, $F\ (1, 56) = 0.64, p > 0.42$. Group 1 had a mean percentage score for items recalled of 54.52% and Group 2 had a mean percentage score for items recalled of 50.90%. It was suggested previously that higher-ability learners develop a strategy based on the form of presentation at learning and that any adverse effect of a change of condition between learning and testing is offset by the advantages offered by a context at testing. These results are not incompatible with that suggestion.

The interesting aspect of this experiment is that it appeared to offer the most favourable circumstances for word frequency effects to manifest themselves given that the task was comprehension, the target L1 items, and the subjects higher-ability subjects. In the event, higher-frequency words were less well recalled than lower-frequency words, even when a context was provided at testing. If this surprising result is due to higher-ability subjects adopting a strategy based on concentration on more difficult items, as discussed above, the results are not in themselves evidence against the influence of word frequency on lexical access. Nevertheless they do indicate that objective word frequency counts are not necessarily a reliable indicator of performance in complex tasks which go beyond lexical access and identification.

Conclusion

It is clear from these results that objective word frequency values are inconsistent indicators of subject performance (see Table 6.10). Word frequency may affect performance; but it may affect it positively as in Experiment 7, or negatively, as in Experiment 8. If subjects perceive word-pairs with readily accessed L1 word components as easy to recall, then concentration on the more difficult items would be a reasonable strategy. This could explain why it is higher-ability learners, who throughout have shown greater versatility in terms of strategy, who show a negative word frequency effect rather than less able learners. However, a more likely explanation is offered in Chapter 8 with regard to word category effects. Either way, it would be inadvisable, on the basis of these results, to predict future performance by objective word frequency values alone.

Table 6.10. Experiments 5, 6, 7, 8.

Summary of word frequency and list position data.

Experiment	Task	List position effect	Word frequency effect	Percentage items recalled
5	Generation	Primacy effect and serial effect	n/a	33.12
6	Comprehension	Primacy effect	n/a	64.58
7	Generation	Recency effect	Positive effect. Interacts with list position.	43.24
8	Comprehension	No significant effect	Negative effect. Interacts with list position.	52.71

In terms of the more general psychological discussion on word frequency effects, in Experiment 7 objective word frequency appeared to have a positive effect on performance. This could be put down to the indirect use of word frequency effects on lexical access, as in criterion bias models. Paradoxically, the negative effect of word frequency in Experiment 8 could also be put down to the indirect use of word frequency effects in that these effects influenced learner strategy. However, it would be rash to draw any firm conclusions without controlling further subjective word frequency and word category.

It is possible at this point to review the evidence on the influence of list position at learning over four experiments (see Table 6.10). Direct comparisons cannot be made across experiments, but it is clear that list position is no more reliable an indicator of performance than is objective word frequency. The effect of list position varies between: a primacy effect and serial effect in Experiment 5; a primacy effect without a serial effect in

Experiment 6; a recency effect in Experiment 7; and no significant list position effect in Experiment 8. There is evidence for a serial effect only with the lower-ability group of learners in Experiment 5; that is, this group of subjects appear to have simply started at the beginning of the list and worked their way through. Otherwise, list order seems to be only one factor affecting subject strategy. One consistent result was that items in the middle of the list were on no occasion either best recalled or worst recalled in relation to the other two list positions. This suggests that these items present particular problems to learners who attempt to compensate for the absence of list position information by adopting alternative forms of processing which appear not to depend to any great extent on either serial information or the frequency of the items concerned.

Objections to list learning based on list dependency would seem to have little support from this data. The main drawback of list learning is for higher-ability learners in the generation task who are misled by the list format in their choice of strategy (see Chapter 3) which is not a case of list dependency as such. Otherwise, list position effects are not obligatory and other variables can interact with them to render them of relatively little importance in the learning process.

CHAPTER 7

The effects of word category on recall

It is reasonable to conclude from the experiments just considered that objective word frequency and list position effects alone are not consistent indicators of performance in second language vocabulary tasks and it was particularly difficult to explain the negative effect of word frequency on recall in Experiment 8 without invoking learner strategy. It could well be that another as yet unconsidered factor is involved. Word category effects are widely discused in both the second language learning and the psychological domains and in this chapter there is an examination of the possible significance of word category effects on learner performance.

A review of the second language learning literature on word category

In the second language learning literature there is a clear sense that certain categories of word present particular difficulty to learners; the precise nature of those categories is a subject of debate. Higa (1965) included the part of speech of the L1 form of the item to be learned as an aspect of word difficulty. Rodgers (1969) saw a hierarchy of categories in ascending order of difficulty: definite pronouns; colours; concrete nouns; possessive pronouns; prepositions; adjectives; numbers; abstract nouns; interrogatives; adverbs; conjunctions; verbs. Blum and Levenston (1978) and Levenston (1979), in the context of the notion of lexical simplification, argued that learners will avoid words which present phonological, grammatical, or semantic difficulty; conversely, they will prefer words which can be used in many different contexts. Levenston (1979) included both morphological and syntactic factors under the heading of grammatical difficulty and it must be presumed that function words are more liable to be syntactically demanding than, for example, concrete nouns. Meara (1980) expressed caution about syntagmatic and paradigmatic categorisation while accepting that "it is quite likely that major differences could be found for words of different types within an individual learner" (p. 239). Carter and McCarthy (1988, p. 16) reported the work of Kellerman (1977, 1983) who used the notion of words' being "psycholinguistically marked". Kellerman (1983) claimed that words "perceived as infrequent, irregular, semantically or structurally opaque, or in any other way exceptional" (p. 117) will present particular difficulty to the learner. Nattinger (1988) was somewhat unusual in claiming that function words can easily be remembered along with concrete nouns. He argued that function words (a loose category which normally includes all words other than nouns, verbs, adjectives, and adverbs) are few in number and of high frequency; concrete nouns are easily imageable. It was not explained why imageability should affect learnability.

The association of concreteness and imageability with ease of learning is not uncommon; indeed, it is one of the complaints against written word frequency lists that they appear not to account for the apparent "availability" of low-frequency concrete nouns (see Michéa, 1964; Nation, 1987; Richards, 1970, 1974). Carroll (1963) claimed that concrete words are more easily learned than abstract words or function words. Strick (1980) argued that learners depend on the concreteness and pronounceability of words in the early stages of learning and only later develop an ability to process them in abstract and semantic terms (see also Henning, 1973).

A review of the psychological literature on word category effects

In terms of the psychology of vocabulary learning, the examination of word category effects is prompted in part by aspects of research into deep dyslexia where word category effects have been consistently in evidence. Although deep dyslexia is normally studied in relation to acquired dyslexia associated with some degree of brain damage, it could be that dyslexic behaviour is not related to brain damage *per se* but to certain characteristics of lexical items which make for difficult processing unless supplemented by other forms of information. If these forms of information were not available to normals, for some reason, then normals too could be said to behave in a dyslexic manner (see Margolin, Marcel, & Carlson, 1985). Without going into detail about the dyslexias *per se*, an analogy is being drawn between the problems encountered by deep dyslexics in processing certain categories of words and possible difficulties encountered by L2 learners in generating certain categories of L2 items when cued with an L1 item, or in understanding certain L1 items when cued with L2 items.

The word category effect refers to the fact that certain categories of words are more easily learned or in some cases more easily pronounced than other categories. This is a somewhat vague definition but its vagueness reflects the present state of the discussion in the psychological literature. For example, the categories used in the discussion vary. In some cases, the discussion is in terms of the concrete or abstract nature of the words to be processed; in other cases it is in terms of the syntactic category of those words. However, there does seem to be an assumption that notions of concreteness and abstractness map well on to categorisation by syntactic function so that these are taken to be more or less equivalent terms. The perceived order is usually nouns, followed by adjectives, verbs, and function words (e.g., Coltheart, 1987a; Cohen & Aphek, 1980; Crothers & Suppes, 1967; Jones, 1985; Marcel, 1987; Marshall & Newcombe, 1987; Marshall, Newcombe, & Holmes, 1975; Morton, 1987; Morton & Patterson, 1987; Patterson, 1979, 1987; Raugh & Atkinson, 1975; Shallice & Warrington, 1975). It should be noted that where nouns are concerned, the status of abstract nouns is often not clear; as nouns it is anticipated they will be easily learned; as abstract, that they will be learned with more difficulty. Generally speaking, though, it seems to be the case that abstract nouns are taken to be of equivalent difficulty to verbs.

The reasons for the effect of word category on processing are the subject of much debate. Indeed, this is not surprising if words possess 51 properties (Rubin, 1980) which could all presumably contribute to the effect. Attempts have been made to find a single measure of word difficulty; that is, one factor which would account for the other measures either as being closely correlated with them or causally connected with them. However, the review which follows will show that these attempts have been inconclusive and that it is unlikely that a unitary explanation for the phenomenon will be forthcoming. In the present study, the factors to be discussed include word frequency, imageability and concreteness, the dual-code hypothesis, and the nature of the lexical representation. The last factor includes discussion of context availability, the branching of nodes in memory, ease of predication, and hemisphere differences.

Word frequency, imageability and concreteness

One possibility would be that word category effects, imageability effects, and concreteness effects could all be manifestations of word frequency. In other words, the frequency of a word would correlate highly with its syntactic category, its imageability, and its concreteness. The first problem with this explanation is that the correlations have not been demonstrated; the second problem is that there is often not sufficient correlation between the other measures themselves to enable a generalisation to be made.

The correlations have not been demonstrated because while there are well-established written-word frequency counts (e.g., Carroll, Davies, & Richman, 1971; Kucera & Francis, 1967; Thorndike & Lorge, 1944), imageability and concreteness counts have only been carried out with a relatively small number of words (e.g., Klee & Legge, 1976; Noble, 1961; Paivio, Yuille, & Madigan, 1968) possibly due to the painstaking methods for eliciting this information.

As to the question of correlation between the various factors, Galbraith and Underwood (1973), for example, called attention to the fact that when college students were asked to assess the relative frequency of paired concrete and abstract nouns, matched for objective frequency, they invariably chose the abstract words as more familiar (see Ghatala & Levin, 1976, for a discussion). Galbraith and Underwood (1973) argued that while there is a good correlation between subjective frequency measures and the objective frequency measures of Thorndike and Lorge (1944), the correlation does not hold when only nouns are used and when these are clearly distinguished by their being abstract or concrete. Further, a comparison between Thorndike and Lorge (1944) and Kucera and Francis (1967) shows a satisfactory correlation between the two counts for abstract words (+0.84), whereas for concrete words the correlation is only +0.61; in addition, relative ordering varies considerably. The Kucera and Francis count appears to overestimate the frequency of abstract words relative to concrete words with the result that the difference between the subjective estimate of the frequency of abstract words and the objective frequency of those same words is less marked when Kucera and Francis values are used than when Thorndike and Lorge values are used. Galbraith and Underwood (1973) argued that their subjects, who were college students, were probably more exposed to abstract words than would be the case for other populations, and that they estimated word frequency on the basis of the number of different contexts in which a word appeared. In this way abstract words would be rated higher than would be the norm. Although Galbraith and Underwood (1973) did not address the issue, their conclusion does point to the possibility that concrete words are easier to learn because they are less varied in their application than abstract words and that they have a more established set of features. The overall conclusion of Galbraith and Underwood (1973) was that concreteness and abstractness as concepts can easily be confounded with subjective or phenomenal frequency; what has been attributed to imagery in the past might be a function of contextual variety which leads subjects to overestimate the frequency of nouns.

L. G. Richards (1976), following Galbraith and Underwood, found that word frequency was a good indicator of performance for both short words and long words, with one exception; for long words in the 10-30 per million frequency range (based on Paivio, Yuille, & Madigan, 1968) concreteness appeared to interact with word frequency. In a second experiment restricted to long words in this critical frequency range Richards found that there was a clear advantage for concrete words over abstract words (although concreteness was still not as good an indicator of performance as word frequency). He therefore concluded that concreteness is in some sense a manifestation of word frequency; since concrete words are learned earlier in life (Anglin, 1970; Brown, 1957), their effective frequency is higher than that of abstract words.

The dual-code hypothesis

One aspect of this discussion has been dominated by the work of Paivio and colleagues (e.g., Paivio 1969, 1971, 1983; Paivio & Csapo, 1969; Paivio & Desrochers, 1979; Paivio & Foth, 1970; Paivio & Rowe, 1970; Paivio & Yuille, 1969; Rowe & Paivio, 1971, 1972). Paivio has argued strongly over many years that word category effects, word frequency effects, and word concreteness effects are all to be explained in terms of imageability. The theory is that there are two forms of semantic representation, one verbal, the other imaginal. This means that for words of high imageability, there is simply more information available from the second (visual) source than is the case for words with lower imageability. Because there is a high correlation between imageability and the other predictors of performance, imageability is taken to be the relevant measure. However, apart from the direct criticisms of the model offered by Pylyshyn (1973, 1981), it has proved almost impossible to dissociate imageability from other factors affecting performance. Richardson (1975a) for example found significant effects in a free recall paradigm not only for imageability, but also for complexity, rate of presentation, serial position as well as a significant interaction between rate of presentation and serial position for early and middle items in the list of items presented. He argued that imageability is always confounded with concreteness. Indeed, Paivio (1969) assumed that imageability and concreteness are alternative measures and Paivio, Yuille, and Smythe (1966) claimed that there is a high correlation between imageability and concreteness. However, even from the data of Paivio and Yuille, this assumption seems to be possible only because Paivio (1969, 1971) and Yuille (1968) simply dismissed as peculiar items which differentiate between the two measures. In fact the correlation between the two measures is normally around +0.80 but both Paivio and Yuille provided evidence counter to their own assumptions in their experimental data. Yuille (1968) for example showed that concreteness is positively correlated with performance in paired-associate learning when imageability is held constant. Paivio, Yuille, and Smythe (1966) also showed that imageability in the stimulus position has no effect on performance if concreteness is high, in spite of which Paivio (1971) claimed that imageability in the stimulus position is normally influential. In the second experiment reported by Richardson (1975a), concrete nouns did not show an effect of imageability in free recall whereas abstract nouns did. Therefore it has to be concluded that imageability and concreteness are not alternative measures. The dual-code hypothesis of Paivio would probably claim that high imageability does not affect concrete words because these words prompt the formation of an image anyway. The opposite is also possible; no words may be "imaged" in the sense in which Paivio uses the term if the dual-coding hypothesis is rejected. The issue remains unresolved.

Further reservations about the dual-code hypothesis were expressed in Winograd, Cohen, and Barresi (1976). They noted that concreteness and imagery have powerful effects on memory and that concrete words are taken to be easier to recall and recognise than abstract words (Paivio, 1971). Paivio's explanation for this can be tested with reference to bilinguals. Many argue for separate language lexical representations in bilinguals. Although there is evidence against the existence of wholly separate stores for different languages, Winograd, Cohen, and Barresi (1976) pointed out the possibility that there may be greater lexical independence between languages for concrete representations than for abstract ones. The argument would be that if the success of subjects in recalling concrete words is due in part to their use of visual imagery then subjects should be less aware of the language of learning when identifying a concrete target word than when identifying an abstract word. The counter argument would be that of Anderson and Bower (1973) and Pylyshyn (1973). They argued for a common store for visual and verbal input; if this is the case, then subjects should have equal facility in recalling the language of learning with both concrete and abstract words. The data of Winograd, Cohen, and Barresi (1976) clearly showed that subjects' memory for the language of learning was better for *concrete* words than for abstract words. Their argument was therefore that a lexical event is encoded as a set of features, including images and the language of the event. For some reason, imagery features are particularly effective in aiding recall and this advantage extends even to the language of the original event. The process is therefore that the activated image in turn activates the set of features for the lexical event, including information about the language of learning.

By the time of the publication of Coltheart, Patterson, and Marshall (1987) there appeared to be little sign of an end to the discussion about the relative merits of word-category, imageability, and concreteness as indicators of performance. Morton and

Patterson (1987), for example, did not attempt to distinguish between imageability, concreteness, and operativity and refer only to the "abstractness effect" (p. 117), whereas Coltheart (1987a) continued to claim that concreteness and imageability are highly correlated, that word-category and imageability correlate, and that therefore the syntactic category effect could be simply due to imageability.

The nature of the lexical representation

Context availability

Running parallel to arguments based on the relative importance of concreteness and imageability have been explanations based on the nature of word representations themselves. Schwanenflugel and Shoben (1983) introduced the notion of context availability. They noted that performance for concrete words is consistently better in learning (Paivio, 1971), recall (Paivio, 1971), recognition (Begg & Paivio, 1969), and comprehension (Holmes & Langford, 1976; Moeser, 1974). However, they pointed out that Paivio's dual-coding explanation elides two different explanations. The first is that there is simply more information available as Paivio would argue. The second depends on the notion of context availability. Either the stimulus itself, or the subject's world knowledge, can provide a context which will make remembering, recall, and comprehension easier. Abstract words and abstract sentences are more difficult to remember and understand because it is difficult for the subject to determine an appropriate context to serve as a useful cue. If this is so, then abstract words should be as easily recalled as concrete words if a context is provided. Thus Pezdek and Royer (1974) found that an abstract word embedded in a paragraph was as well remembered as a concrete word. Bransford and McCarrell (1974) found that in paired-associate learning the abstract word *hindrance* is as well remembered as the concrete word *wheelchair* when both are paired with the concrete word *stairway*. In other words, the concrete associate serves as an effective context for the abstract word. Schwanenflugel and Shoben (1983) argued that context availability is a better predictor of performance than ratings of concreteness in lexical decision tasks.

Two different kinds of explanation for the predictive power of this notion are given. Information is more weakly associated for representations of abstract concepts than for representations of concrete concepts. The amount of spreading activation to a particular link from a particular representation is inversely related to the number of links from that representation (Anderson, 1976). Therefore the amount of activation available to traverse the link is dependent on the degree of fan-out; the more fan-out, the less likely the connection is to be made. Abstract words have more fan-out because they occur in more contexts. Provision of a specific context would select part of the network and thus concentrate the flow of activation rendering performance on abstract words as good as that with concrete words. An alternative explanation is that of Kieras (1978). He argued that abstract words have less information stored with them than concrete words which are both more familiar and have more propositions associated with them pertaining to perceptual aspects. Context therefore provides information for abstract words which would not otherwise be available. At first sight it is difficult to envisage how this explanation could be accommodated by the network model of memory. The availability of more information would seem to imply that it is more difficult to access that portion of information which is relevant; therefore it would seem to predict that relevant contextual information is more difficult to access for concrete words due to the large number of associative links. However, this problem can be overcome. Information associated with abstract concepts is only weakly associated whereas some information for concrete concepts is very highly associated; it could be that this applies to perceptual information in particular. In other words, context primes a subset of the information available and the target, being highly associated, rapidly reaches threshold. This explanation looks similar to the dual-coding model because perceptual information is identified with concrete concepts. However, the dual-coding model would say that only an increase in imaginal information would improve access to abstract words whereas what is being said here is that any association would have that effect. This view is consistent with the idea that abstract words have more associative contexts so that it is difficult to retrieve a particular one.

Context-dependent and context-independent representations

This notion is also compatible with the ideas of Barsalou (1982). For any conceptual representation, some of the information will be context-dependent and some will be context-independent. The difference is that context-independent information is always automatically activated by the appropriate input; context-dependent information is not. Context-independent information would be features of high diagnosticity, such as *gills* in the case of the representation for *fish*, or features relevant to the typical use of a word, such as *edible* for *apple*. Context-dependent information is not activated by words as such but by contexts in which the words have been encountered. Some would argue that all the properties of a concept are context-independent; for example, Katz and Postal (1964) envisaged a concept as a fixed set of properties. Others say that all the properties of a concept are context-dependent (e.g., Olson, 1970). Barsalou's argument is that if there is something of truth in both extreme positions, in any given context all context-independent information should become available and only some of the context-dependent information and this is borne out by the experiments reported. Although the distinction between context-independent and context-dependent information bears some resemblance to Tulving's (1972) distinction between semantic and episodic memory, the difference here is that context-independent information is derived from context-dependent information and it is not necessarily "more semantic" in the sense intended by Tulving. The difference between context-independent and context-dependent information is based solely on the means of activation in each case.

De Groot (1989) offered another explanation based on the network model of memory (Anderson, 1976; Collins & Loftus, 1975) and brought together ideas from both the imageability discussion and from the discussion about the nature of concept representations. She argued that if there is more branching out from a given node, the node itself is relatively less activated; if there is less branching out from a node, the node is relatively more highly activated. Some argue that abstract representations have more nodes and this accounts for their being perceived as more frequent (Galbraith & Underwood, 1973; Schwanenflugel & Shoben, 1983). The counter argument is that abstract representations have less nodes; concrete representations not only have more nodes but some of the connections between these nodes are particularly strong. De Groot explored possible effects of these assumptions. She argues that given the measure of response availability *m* (Noble, 1952), based on the mean number of associations evoked by a word, if concrete representations have more nodes, *m* should be larger for these words than for abstract words and further if the links between these nodes are stronger, the time taken to evoke responses should be less than for abstract words.

Word frequency also has to be taken into account. Response times in a word-association task could be shorter for a high-frequency word than for a low-frequency word because the links between its nodes are stronger; but this advantage may be offset by the larger number of links from these nodes which could make the process of association slower. The question is, then, whether the strength of the links or the number of links has the greater effect. Her conclusions were that imageability (which here is taken to be totally confounded with concreteness) has a strong effect on word-association whereas word frequency has a weaker effect on word-association. There are larger *m* scores for concrete words because their concept nodes hold more information (Kieras, 1978; Simons, Vonk, & Noordman, 1989) and the information is less dispersed. The number of different responses by subjects to words, and the number of idiosyncratic (one-off) responses to words, is negatively correlated with the imageability of those words. Not only do people have expert knowledge of concrete words but this knowledge is commonly shared among subjects. Among the advantages shared by concrete words is that their representations contain perceptual information, though whether this is stored in a separate visual store (Paivio 1978, 1986) or in the form of propositional representations (Pylyshyn, 1973) is left an open question.

The relative unimportance of word frequency in predicting performance in this context contrasts with the findings of Cramer (1968). Cramer found that word-association reaction times were less with high-frequency words than with low-frequency words; that *m* response availability was greater; that heterogeneity of response decreased; that the association frequency of primary responses increased. Only the second of these findings was replicated by de Groot. De Groot argued that the most likely explanation for the discrepancy is that Cramer used a greater difference between frequency classes (0-9 per million for low frequency and 70-250 per million for high frequency). De Groot also argued on the basis of the data reported that link strength is a greater determiner of association speed than the amount of fan-out. The number of links as such has little effect since it is true that response availability is larger for concrete words and that the association time is shorter. One corollary of this is that if word frequency does not affect association retrieval, simply activating the link will not strengthen it. It is the activation of the target and its active association with the prime that affects the strength of the link. In other words, although priming studies show that *war* activates *peace*, the strength of the link between the two is not a function of the frequency of *war*. Only a direct association of *peace* with *war* will strengthen this link. Thus *m* scores for high-frequency and low-frequency words are virtually identical and this could be because free association taps only context-independent information and does not tap the context-dependent information (Barsalou, 1982) which could be what differentiates high-frequency from low-frequency words. Context-independent information is taken to be roughly equal for both high-frequency and low-frequency words, but high-frequency words occur in more contexts (cf., Jones, 1985, and the notion of ease-of-predication) and therefore have more context-dependent information associated with them than is the case with low-frequency words. So whereas concrete and abstract words can be differentiated by perceptual information in addition to context-dependent and context-independent information, there is no such obvious source of differentiation between high-frequency and low-frequency words in the absence of context-dependent information.

Ease of predication

Jones (1985) offers an explanation in terms of ease of predication with reference to both dyslexics and normals. He referred to the work of Paivio (e.g., 1971, 1983) who talked of imageability as the readiness with which *x* is encoded by the imagery-encoding component of a dual-coding system. Jorm (1979) suggested that the importance of imageability in developmental dyslexia is that it affects the direct route for visual word processing when the more normal phonological route is impaired. However, the evidence is against this hypothesis because it shows that subjects become less dependent on phonology as they become more expert at reading (Baddeley, Ellis, Miles, & Lewis, 1982; Doctor & Coltheart, 1980). It is probably better to look in the semantic domain for the source of the so-called imageability effect in deep dyslexia because it is at this stage in the language process that experience of mental imagery and ease of reading coincide. If it is accepted that the representation of a word is made up of a number of features or predicates, then it is possible to hypothesise that the variability in the ease of imagining is a function of the variability in the associated distribution of predicates for individual words. It is clear that imageability, as a measure, is related to the performance of deep dyslexics. However, Jones (1985) argued that this is not due to the use of imagery, as such, but due to the relationship between imageability and ease-of-predication. In deep dyslexia reading output is only possible when mediated by the cognitive system (Morton & Patterson, 1980). Input to the system seeks "its" predicates and the more easily these are found, the easier it is for the subject to name the word. Imageability therefore is closely related to ease-of-predication or the "ease of putting words into simple factual statements" (Jones, 1985, p. 4) . Thus although Paivio, Yuille, and Madigan (1968) found that imageability scores were higher than ease-of-predication scores, the correlation was satisfactory at +0.88. Further evidence for such a correlation is to be found in Anderson and Bower (1973), Kieras (1978), and Shallice and Warrington (1980). The advantage of this explanation is that it is not necessary to posit two kinds of semantic representation with only one of them damaged as in Morton and Patterson (1980) and it does constitute a good explanation for many symptoms in deep dyslexia including that of the word category effect. In summary, then, the explanation is that the deficit is not due to damage to an abstract-word linguistic processor but follows naturally from reliance of the subject on the cognitive route which may or may not yield matching predicates. For normals, the cognitive route is probably not accurate enough because the semantic system is inherently imprecise (Anderson & Ortony, 1975) and so the predication effect is not to the fore. Words vary according to the ease or difficulty with which they can be put into simple statements; in other words, they differ in their ease-of-predication. The argument is that not only individual words vary in this manner but syntactic categories of words also; these categories would be ordered in the same way as the order of difficulty encountered by deep dyslexics.

Hemisphere differences

Another approach to the question of different kinds of representation in the system comes from the study of brain-damaged patients (Marcel & Patterson, 1978; Marshall & Newcombe, 1966; Richardson, 1975b; Schwartz, Saffran, & Marin, 1977; Shallice & Warrington, 1975). A dissociation between the representations of concrete and abstract words has been observed in spontaneous speech (Goodglass & Geschwind, 1976) and Warrington (1975) describes a patient who unusually, and perhaps uniquely, had retained comprehension of abstract words but had impaired comprehension of concrete words. The difference in kind of representation could be related to the way in which words are acquired (Goodglass, Hyde, & Blumstein, 1969; Marcel & Patterson, 1978; Shallice & Warrington, 1975), or alternatively to the relative importance of content words in "reading for meaning" (e.g., Andreewsky, Deloche, & Kossanyi, 1987). A distinction is often made between errors based on semantic relationships (paradigmatic errors) and errors based on associative relationships (syntagmatic errors), although the distinction is somewhat difficult to sustain in practice (see Meara, 1980, 1982). An example of the former would be "happy" as a response to "merry"; an example of the latter would be "Christmas" as a response to "merry". Coltheart (1987b) argued that in word-association tasks, children tend to make syntagmatic responses, whereas adults tend to make paradigmatic responses. However, the latter tendency depends on the category of the stimulus; the less concrete the stimulus (in the order: noun, verb, adjective, adverb), the greater the tendency to give syntagmatic responses (Deese, 1962). With deep dyslexics something similar can be observed; nouns tend to produce paradigmatic errors whereas adjectives and verbs result in associative errors.

These differences could be related to hemisphere differences. Marcel and Patterson (1978) noted that phonemic dyslexics have a selective inability to read words of low imageability (Patterson & Marcel, 1977; Shallice & Warrington, 1975) and they discussed the possibility that this is a hemisphere effect which could be replicated in normals under certain circumstances. They found that if concreteness and imageability are co-varied independently, concreteness has no effect (see also Richardson, 1975a), whereas imageability interacts with the hemisphere used, affecting only the left visual field. In a further experiment, Marcel and Patterson (1978) varied imageability but held concreteness, word frequency, and length constant; here the findings were even more clear-cut again showing that imageability affects only the left visual field. There are three possibilities for the locus of the deficit: it could be pre-lexical (Ellis & Shepherd, 1974); or lexical (Richardson, 1975b); or semantic (Shallice, 1978). An attempt was made to remove production as a factor by using pattern masked semantic priming in a lexical decision task. The results show that if the need for production is eliminated, so is the imageability effect. In other words, imageability seems to be something to do with that part of the semantic representation which affects production. They hypothesise that initially language is not hemisphere specific until around the age of three. The influence of the time of acquisition of a word is not seen at lexical level because a number of high imageability words are only acquired relatively late in development and conversely function words which are difficult to process are acquired early. So the source of the effect appears to be in the semantic system. It could be that a distinction needs to be made between words of sensori-motor origin, S_1, and words of lexical or linguistic origin, S_2. S_1 and S_2 words acquired early may be stored in either hemisphere. However, only S_2 words in the left-hemisphere have access to production, therefore S_2 words in the right-hemisphere are produced via the left-hemisphere. Phonemic dyslexics have left-hemisphere problems which leave only S_1 words intact.

Warrington (1981) argued that there is a structural basis in the brain for performance with concrete and abstract words and that this is particularly evident in dyslexics. One explanation offered was that the phonological route is needed for abstract words because only concrete words can use the direct semantic route from orthography (e.g., Marshall & Newcombe, 1973; Shallice & Warrington, 1975). This is no longer tenable because some patients with little phonological capacity show no concreteness effect (e.g., Shallice & Warrington, 1980; Warrington & Shallice, 1979). An alternative explanation is based on a dual encoding hypothesis (Richardson, 1975b). The idea here is that concrete words are in the right-hemisphere and can be pronounced on the basis of the lexical entry alone; if the right-hemisphere stores concrete words and remains undamaged, then this would explain differences in the performance of some patients between abstract and concrete words (Coltheart, 1980; Marcel & Patterson, 1978; Saffran, Bogyo, Schwartz, & Marin, 1980). However, this explanation implies that all patients with an intact right-hemisphere should be able to read some written words; this is not the case (Warrington & Shallice, 1980). Further, the patient C.A.V. is evidence against the argument for the storage of concrete words in a right-hemisphere lexicon. This patient has a left-hemisphere lesion but a deficit for concrete words.

Saffran, Bogyo, Schwartz, and Marin (1987) discussed the word-category effect in terms of right-hemisphere use in normals. The argument is that if deep dyslexics use the right-hemisphere, then the same word-category effects would be expected in normals who are made to use the right-hemisphere. With tachistoscopic presentation it was found that overall performance was better in the right visual field but that there were no interactions between visual field and word category; in particular, abstract nouns and function words did not show a larger right visual field advantage than concrete nouns. It could be therefore that all verbal input is transferred from the left visual field to the left-hemisphere for analysis. On the basis of this data, Saffran, Bogyo, Schwartz, and Marin (1987) concluded that the imageability effect is not robust; it does not apply to all subjects under all circumstances. If the right-hemisphere is used, deep dyslexic behaviour results but this is not necessary in normals.

Coltheart (1987c) also advanced an explanation based on hemisphere representations. Ellis and Shepherd (1974) found with normals that concrete words had an insignificant left-hemisphere advantage whereas abstract words had a significant left-hemisphere advantage (see also Hines, 1976, 1977). In other words, it is consistent with the data that deep dyslexics use the right-hemisphere to read and that normals (perhaps) show that this hemisphere is less effective with abstract words. What seems to happen is that a lesion abolishes access from orthography to the left-hemisphere of the brain. However, spoken responses are derived from the left-hemisphere. So reading aloud requires access to information in the right-hemisphere, the transfer of information to the left-hemisphere, access to pronunciation in the left-hemisphere, and then pronunciation. A certain amount of information is lost in this process, but the more specific a semantic feature is, the more willing a patient is to tolerate a mismatch of that feature. The problem with function words is that the right-hemisphere is not effective with purely syntactic functions and this is true of many function words. The problem with abstract words is that when they are stored in the right-hemisphere they are semantically impoverished or simply have no representation at all. In either case, reading aloud becomes difficult.

Conclusion

In conclusion, it has to be said that there is rather more consensus about the existence of word category effects than there is about an explanation for them. Explanations based on the nature of the representations seem to offer the best way forward since there is an apparent circularity in explanations based on measures of abstractness, concreteness, imageability, and word frequency. Representations for concrete words appear to have more information attached to them and this information appears to be more strongly associated. Representations for abstract words have less information attached to them and this information is less strongly associated. Concrete words are less context-dependent and this in turn means that they are more easily predicated; abstract words are more context-dependent and are therefore less easily predicated.

The issue is of importance in the present thesis and in the second language learning domain generally because of the need to establish criteria for the concept of word difficulty when other criteria examined appear to be somewhat inconsistent.

EXPERIMENT 9

Experiment 9 was designed to test the effect of word category on subject performance. The notion of "word category" can be seen as a higher-level description for a number of phenomena which have proved difficult to define and explain otherwise. In general terms, it seems to be agreed that items in what can loosely be called more concrete syntactic categories are more easily processed than items in more abstract syntactic categories (Coltheart, 1987a; Coltheart, Patterson, & Marshall, 1987; Jones, 1985; Marcel, 1987; Marshall & Newcombe, 1987; Morton, 1987; Morton & Patterson, 1987; Patterson, 1979, 1987; Schwanenflugel & Shoben, 1983; Shallice & Warrington, 1975). The more concrete syntactic categories are taken to be concrete nouns and adjectives; the more abstract are taken to be verbs and function words; abstract nouns are seen as somewhere in between but tending towards the latter category. "Concreteness" is not a self-explanatory attribute of a category or a word and there is a range of explanations for the accessibility of items within categories as discussed.

In the present domain, there is a particular interest in a possible word category effect because of the unreliability of both objective word frequency and list position as consistent indicators of subject performance. Word category could fulfill this predictive role if it could be shown that the behaviour of L2 learners is subject to word category influences. The categories chosen for examination in this experiment were concrete nouns, abstract nouns, adjectives, verbs, and "other words". The first four categories occur in most discussions of the issue (see the introduction to this chapter for references); the category of "other words" is the equivalent of the often discussed but rarely defined category of "function words" which is avoided because it could be taken to exclude, for example, adverbs.

Given the evidence discussed above, it was predicted that concrete nouns would be better recalled than abstract nouns and verbs with adjectives and "other words" somewhere in between.

Method

Design

The experiment had a 2 x 5 factorial design. The between-subjects factor was the test condition with two levels, generation and comprehension. The within-subjects factor was word-category with five levels. The five levels were concrete nouns, abstract nouns, adjectives, verbs, and "other words".

Materials

A word-pair list of 20 items was prepared. Word frequency was controlled within the range of 50-70 occurrences per million (Kucera & Francis, 1967). There were five categories of word used with four items in each category: concrete nouns; abstract nouns; verbs; adjectives; other words. The imageability of the concrete nouns chosen was confirmed by independent judges and with reference to imageability ratings contained in Quinlan (1992). The category of "other words" included one adverb and three prepositions. Items from each category were intercalated. A full listing is provided in the Materials Appendix.

Subjects

A pool of 47 subjects, aged 11-13, from School C9 took part in the experiment; none of the subjects had taken part in previous experiments. The experiment took place in the Summer term of the school year. Subjects had already experienced, therefore, nine months of formal French teaching when they began the experiment.

Procedure

Two experimental groups were formed. Group 1 was tested for generation. Group 2 was tested for comprehension. Group 1 consisted of 23 pupils (11 boys, 12 girls); Group 2 consisted of 24 pupils (12 boys, 12 girls). All subjects were tested in a simple sentence context. In all other respects, the procedure was as in previous experiments.

Results and discussion

An analysis of variance was performed. The test condition (group membership) was the between-subjects factor. Word category, with five levels, was the within-subjects factor.

Word category had a significant effect on performance, $F\ (4, 80) = 17.68, p < 0.01$ (see Table 7.1 and Figure 7.1). A pairwise comparison (Tukey test) shows that three categories were relatively well recalled: concrete nouns (56.99% items recalled); adjectives (59.55% items recalled); "other" words (56.02% items recalled). Performance with these items was significantly better than with verbs (35.05% recalled) and abstract nouns (34.10% recalled), $p < 0.01$. These results are, therefore, broadly in line with the generally agreed

Table 7.1. Experiment 9.

Mean percentage scores for items recalled by word category and test condition.

Word category	**Group 1 Generation**	**Group 2 Comprehension**	**Overall**
Concrete nouns	60.86	53.12	56.99
Abstract nouns	36.95	31.25	34.10
Verbs	32.60	37.50	35.05
Adjectives	58.69	60.41	59.55
Other words	64.13	47.91	56.02
Overall	50.65	46.04	48.34

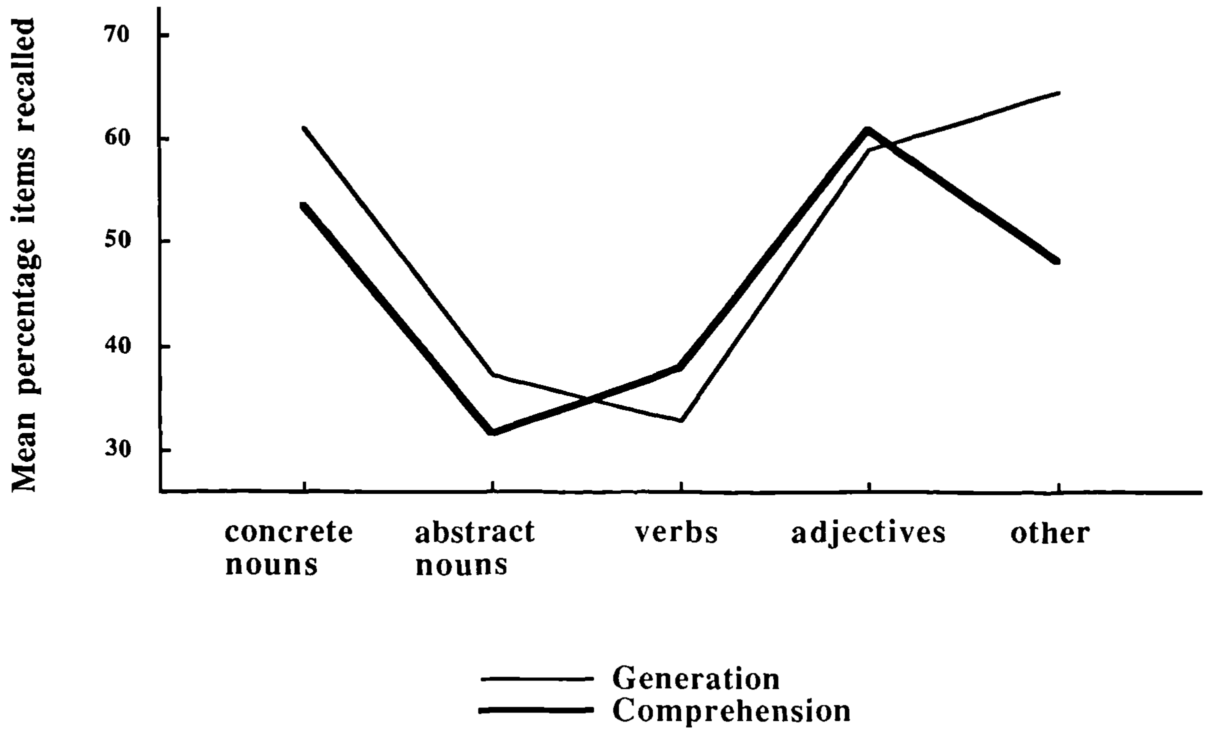

Figure 7.1. Experiment 9. Performance by word category and task.

order of ease of processing in the psycholinguistic literature (e.g., Coltheart, 1987a; Cohen & Aphek, 1980; Crothers & Suppes, 1967; Jones, 1985; Marcel, 1987; Marshall, Newcombe, & Holmes, 1975; Marshall & Newcombe, 1987; Morton, 1987; Morton & Patterson, 1987; Patterson, 1979, 1987; Raugh & Atkinson, 1975; Shallice & Warrington, 1975), and in the second language learning literature (e.g., Higa, 1965; Rodgers, 1969).

The one surprising result would be the relatively strong performance of subjects with the "other words" category. It could be that this somewhat loose categorisation of items needs to be discriminated further if it is to be of use. For example, it could be argued that words such as *beneath* and *towards*, included in the materials for this experiment, are more imageable than other words which could have been used such as *without* or *otherwise*. However, it is difficult to see how this line of argument would account for the recall of *equally* and *besides* which appear to be abstract by most definitions. Clearly there is scope for more investigation in this respect (see Rubin, 1980), possibly along the lines of Patterson (1981) who similarly argues for distinctions within the category of "function words" on the basis that some have semantic content (for example, *she* and *between*) whereas others do not (for example, *of* and *at*).

The task undertaken did not have a significant effect on performance. Subjects tested for generation recalled 50.65% of the items presented and subjects tested for comprehension recalled 46.04%, F (1, 45) = 0.47, $p > 0.49$. There was no significant interaction of word category with the task undertaken, F (4, 80) = 1.91, $p > 0.11$.

There is therefore a very clear word category effect evident in these results. The effect was independent of word frequency which was held constant and independent of the different tasks undertaken. It could well be that until a closer correlation is established between subjective and objective word frequency norms, and between word frequency, concreteness, imageability, and word category, it is word category which is the most reliable indicator of learnability, at least in this particular domain.

On the basis of these results, a *post hoc* examination of the materials used in Experiment 8 was undertaken. Here, it will be recalled, high-frequency words were significantly less well recalled than low-frequency words, contrary to expectations. As Table 7.2. shows, the high-frequency words with lower than average recall in this experiment included: four abstract nouns, two verbs, and one "other word". Conversely, the low-frequency words with higher than average recall included: three concrete nouns, two adjectives, and one abstract noun. It could be therefore that word category was the crucial factor in these results and that word frequency was relatively unimportant.

However, a certain amount of caution is required. What has become evident over the course of these experiments is that quite small changes in subjects, tasks, and materials can bring about quite large changes in performance. It should be noted that these results, and the results for Experiment 8, for example, both related to higher-ability subjects. It is also true that the lists in both cases were not the random lists used by students in the normal course of events, but controlled for frequency. Finally, while it is relatively easy to establish membership of categories such as concrete noun, verb, and adjective, the category of "other words" needs more development and discrimination. Hence, a considerable amount of work needs to be done in this area.

Table 7.2. Materials from Experiment 8. High-frequency words with below average recall and low-frequency words with above average recall arranged by word category. Experimental mean 51.54%.

Item	Word frequency	Word Category	Percentage recall
High-frequency tems with below average recall			
half	275	abstract noun	18.74
thought	515	abstract noun	32.80
the land	217	abstract/concrete noun	36.37
to place	571	verb	37.32
to believe	200	verb	40.10
quite	281	other	44.18
the rate	209	abstract noun	44.67
Low-frequency items with above average recall.			
relaxed	14	adjective	51.95
the herd	22	concrete noun	62.77
crazy	34	adjective	65.22
the sand	28	concrete noun	75.60
the mood	37	abstract noun	77.59
the bush	14	concrete noun	80.92

CHAPTER 8

The effect of embedded words on memory processes

The final aspect of word learnability to be considered in this thesis is the effect of embedding. That is, the effect on learning of an English word embedded in a French vocabulary item; for example, the French word *causer* has embedded in it the English words *cause, use,* and *user*. The effect of embedding can be considered in the context of the discussion about similarities between items in L1 and L2 in the second language learning domain and in the psychological literature.

Cross-language influences: a review of the second language learning literature.

A constant theme in the second language learning literature is that similarity between L2 words and their L1 equivalent is an important factor for ease of learning. Thus Anderson and Jordan (1928) discussed the importance of the similarity in form and meaning of the L2 target to an equivalent item in L1. Higa (1965) in his treatment of the notion of word difficulty envisaged an interaction between the target word and other words known to the learner; in brief, the meaningfulness and therefore learnability of an L2 word will in part be a function of its similarity to a known L1 word, the familiarity of the L1 equivalent to the learner, and the meaningfulness of the L1 equivalent to the learner. Blum and Levenston (1978) and Levenston (1979) in their explanation of "lexical simplification" saw phonological, grammatical, and semantic features of L2 words as influencing learnability and these features are to some extent "defined" against the learner's L1. Nation (1987) saw the regularity of an L2 item, in terms of a learner's L1, as influencing its ease of learning. Grainger and Beauvillain (1987) found interference effects between L1 and L2 in the case of a common orthographic pattern across languages; that is, when an L2 word (such as *lire* in French) forms an orthographically legal string in L1 (in this case English).

However, most of the studies in the literature envisage the influence of similarity in terms of L2 items and their L1 semantic equivalents. This similarity can range from exact cognates (for example, *Paris* and *Paris* where the only difference is in pronunciation) to near cognates where minor differences do not obscure a commonality of meaning (for example, *poste* and *post*). What is being investigated in this chapter is whether L1 items embedded in an L2 word will influence the learnability of an L2 word in the absence of any semantic connection between the two (for example, the English word *bran* has no semantic connection with the French *ébranler, to rattle*, in which it is embedded). In this context, therefore the notion of pronounceability is probably of more direct importance. Many writers identify pronounceability as affecting learnability in a general sense (e.g., Higa, 1965; Levenston, 1979; Nation, 1987). More specifically, Rodgers (1969) suggested that learners preparing for comprehension pronounce the stimulus word; match it to a known L1 word as a mediator; then associate the mediator to the English target word. Henning (1973) made the point that acoustic and orthographic clustering are normal early in the learning process. Meara (1982) discussed the phenomenon of "clang associates" in language learning. That is, second language learners, as do young children, tend to produce an L1 word which shares a sound relationship with an L2 stimulus even though the item produced lacks any obvious semantic relationship with the target. Meara gives the example of the response *vache* to the stimulus *mou* (*soft*) which is presumably elicited because of the sound of the stimulus item. Hatch (1983), referring to the work of Cohen and Aphek (1979, 1980), found that learners use a range of associations in the learning process and these include the use of sound similarities independently of any semantic association. It seems reasonable to suppose that an English word embedded in a French word *ipso facto* makes that word more easily pronounceable than would otherwise be the case and that this might affect learnability.

Verbal mediation

In psychological terms, the main reason for supposing that English words embedded in French words might affect subjects' memory for French words is that of verbal mediation, though it is not clear whether that effect would enhance or inhibit learning. If an English word embedded in a French word is accessed by subjects, then the embedded word could be available either as a mediating term between the known (the L1 item) and the unknown (the L2 item in which the word is embedded) to assist the performance of the subject; it could on the other hand lead to inhibition if subjects do not use the embedded word as a mediating term but misrecognise it as a semantic clue to the target in the manner of a false cognate or *faux ami* (see Banta, 1981).

The idea of verbal mediation, the use of a known term to mediate between a cue and a target, is well established in the literature on memory. Paivio (1971), after providing a very full review of the evidence, concluded that subjects "persistently find ways of coding items or mediating interitem associations, rather than learning by rote" (p. 300). The use of verbal mediation in word-pair learning is specifically addressed in Adams (1967), Battig (1966), Bugelski (1962), Bugelski, Kidd, and Segmen (1968), Horotwitz and Newman (1969), and Underwood and Schulz (1960). These studies showed clear evidence for at least a correlation between the use of verbal mediation and word-pair performance, although Paivio (1971) was prepared to argue for a causal relationship between them. It is interesting to note in the present domain that verbal mediation is particularly effective when one or both members of the word-pair is a nonsense word (Paivio, 1971, p. 301); it is at least possible that subjects will treat French language items as nonwords and that verbal mediation will be helpful to performance.

However, if embedded words are to be used for mediation, then they must first be identified. Therefore a preliminary issue to be resolved is whether English words embedded in French words are accessed when the French words are encountered. Reasons for supposing that this might be the case can be grouped into four, ranging from the general to the particular. They are: the neighbourhood effect; the tendency of the system to seek words, or the lexicalisation effect; the phenomenon of word-constituent priming; and cross-language effects.

The neighbourhood effect

The neighbourhood effect is usually discussed in relation to word frequency effects rather than in relation to the issue being discussed. However, what the neighbourhood effect seems to show is that response to a target word is in part determined by its orthographic neighbours, that is, words with which it shares orthography. A word's neighbour is variously defined, but the definition of Coltheart, Davelaar, Jonasson, and Besner (1977) is commonly adopted; a neighbour of a given word is any word which can be constructed by changing one letter of the target word. Thus most studies of the neighbourhood effect have been conducted with neighbours which differ from the target word by only one letter. However, there is no reason to suppose that neighbourhood effects are restricted to these "near neighbours" particularly when L2 items are involved which, by definition, will be surrounded by higher-frequency L1 words with which they share aspects of orthography.

The neighbourhood effect was outlined by Andrews (1989). The basic idea of the neighbourhood effect is that when a word is encountered, not only that word is activated but also its neighbours. Neighbours activated in this way can affect the speed of the recognition process. Coltheart, Davelaar, Jonasson, and Besner (1977) had found an inhibitory neighbourhood effect, but only with nonwords and not with the classification of words. Forster, Davis, Schocknecht, and Carter (1987) and Meyer, Schvaneveldt, and Ruddy (1974) showed that a word which differs from a target by one letter will prime the target word, even if the prime word is masked (Forster, 1987). McCann and Besner (1987) found a facilitatory effect of neighbourhood size on nonword naming latencies and therefore reversed the finding of Coltheart, Davelaar, Jonasson, and Besner (1977) for nonwords in the lexical decision task. Luce (1986) found that a large neighbourhood speeded the identification of masked words and Laxon, Coltheart, and Keating (1988) found that children and poorer readers were more successful in their recognition of words which had a large number of neighbours.

What is not clear in the literature is the conditions under which neighbourhood effects will be facilitatory or inhibitory. It appears to depend on the nature of the task and the precise characteristics of both the target word and the neighbourhood concerned. So, for example, Andrews (1989) argued that facilitatory effects are observed for low-frequency words but not for high-frequency words in all tasks except delayed naming. The effect is also more marked in the lexical decision task than in the naming task, suggesting that part, but only part, of the neighbourhood effect is located in a post-access decision process. Taking an explanation from the interactive activation group of models, the presence of neighbours could result in facilitation at the letter level but inhibition at the word level. Therefore whether the effect will be facilitatory or inhibitory depends on the relative contribution of letter level and word level effects to the recognition process. High-frequency words require less information to be activated, therefore they cannot "use" the extra activation provided by letter-level information. Low-frequency words are recognised more slowly, therefore they can take advantage of the letter-level activation provided by neighbourhood activation. Given her results, it appears to Andrews that the advantages provided by letter-level excitation outweigh the inhibitory effects of neighbourhood activation at word level where low-frequency words are concerned. It is worth noting that search models cannot easily cope with the notion that a large neighbourhood can actually facilitate word recognition since a large number of competitors should take longer to process than a small number of competitors.

Grainger, O'Regan, Jacobs, Segui (1989) and Grainger (1990) offered an alternative explanation. They argued that the effect is concerned not so much with neighbourhood size as with neighbourhood frequency. The failure of Andrews (1989) to take account of this might at least partly explain the differences between her results and those of Coltheart, Davelaar, Jonasson, and Besner (1977). Response to an item may be inhibited, rather than facilitated, if that item has at least one neighbour of higher frequency in competition with it. The significance of the phrase "at least one" is that the effect of more than one higher-frequency word is not cumulative; neighbourhood size, in this sense, is irrelevant. Low-frequency words tend to have high-frequency neighbours whereas high-frequency words tend not to have higher-frequency neighbours (Grainger, 1990), and this could result in neighbourhood effects being apparent with low-frequency words only. The argument that low-frequency words tend to have high-frequency neighbours whereas high-frequency words tend not have higher-frequency neighbours was not supported by Luce, Pisoni, and Goldinger (1990), although the overall argument is not seriously compromised. Luce suggested that high-frequency and low-frequency words have a similar number of neighbours, but that high-frequency words tend to have higher-frequency neighbours than do low-frequency words and that therefore high-frequency words are *potentially* more easily confused with each other than are low-frequency words. On the other hand, the difference between high-frequency words and their neighbours is greater than that between low-frequency words and their neighbours, therefore high-frequency words are more distinctive. The crucial factor in the neighbourhood effect is, therefore, the relationship between target frequency and neighbourhood frequency.

Grainger (1990) used the notion of neighbourhood frequency to argue that when target and neighbourhood frequency are taken into account, it can be shown that neighbourhood effects are to be found in both lexical decision and naming tasks. Thus, for example, if a stimulus is preceded by a low-frequency orthographically related word, lexical decision and naming are slowed compared with latency times obtained when the target is preceded by a neutral prime. In processing the low-frequency prime, the system must inhibit the activation of any high-frequency neighbours; this means that when the target is presented, it is in a state of inhibition which slows down its recognition.

Support for the argument of Grainger (1990) is to be found in Marslen-Wilson (1990). In a cross-modality lexical decision task, designed to test neighbourhood effects in spoken word recognition, Marslen-Wilson found an unexpected competitor frequency effect in the visual domain. The design of the experiment involved giving an auditory prime followed by a visual target. Although there was no competitor frequency effect in the auditory lexical decision task proper, in the baseline condition, where there was no connection between the prime and the target, there was an effect of competitor frequency on the visually presented probes. High-frequency probes were responded to more quickly than low-frequency probes as would be expected; but low-frequency probes with high-frequency neighbours were responded to more slowly than low-frequency probes with low-frequency neighbours. His explanation was that there is, in both modalities, a competitor frequency effect but this is transitory and of brief duration. It is thus dissipated in the auditory modality by the time the full sensory input is completed; in the visual domain, the input of sense information and the decision process are virtually simultaneous and this allows competitor frequency effects to become manifest.

While visual word recognition is of particular interest to this study, research into neighbourhood effects in word naming has raised some issues which appear to have relevance to the present domain. This is for two main reasons. First, because visual word recognition and pronunciation quite probably share a number of processes in normal reading (Monsell, 1987). Second, because readers have available to them phonological codes which can play a part in the process of understanding written words (see reviews in Carr & Pollatsek, 1985; Coltheart, 1978; Henderson, 1982, 1985; Humphreys & Evett, 1985; Kay, 1985; Papagno, Valentine & Baddeley, 1991; Patterson & Morton, 1985; Seidenberg, 1985b; Service, 1992). To that extent, an understanding of the role of neighbourhood influences in the processes of the conversion from orthography to phonology may well be helpful.

As Patterson and Coltheart (1987) pointed out, for some time the regularity of a word's spelling to sound correspondence was taken to be a sufficient indicator of its ease of naming. Regularity is, of course, closely related as a notion to that of neighbourhood since it is an indication of the sound characteristics of a word relative to its neighbours. A word was taken to be regular if its pronunciation conformed to the normal pronunciation for words sharing the same orthographic body. This notion was challenged by Glushko (1979) who claimed that regularity is not a sufficient indicator of ease of pronunciation; consistency is the relevant factor. A word is consistent only if it has the same pronunciation with *all* other words with which it shares an orthographic body. Thus a word can be regular but inconsistent and in this case its pronunciation will be slower than for a regular consistent word.

More recently, regularity has been to a large extent reinstated as the relevant factor. Seidenberg, Waters, Barnes, and Tanenhaus (1984) showed that consistency only affects low-frequency words; Kay and Bishop (1987) found an inhibitory effect for inconsistent words only when they are low-frequency words which also share their pronunciation with very few of their neighbours. In the formulation of Kay and Bishop (1987), consistency effects are confined to low-frequency words with weak bodies, and to exception words. A body in this context is defined as the medial vowel (or vowel cluster) of a word plus the terminal consonant (or consonant cluster) of that word. A weak body is defined as one whose pronunciation is not the most common pronunciation for that string (e.g., *eat* in *sweat* as opposed to *eat* in *beat*). An exception word is one with a unique pronunciation such as *great.*

It is interesting to compare the notion of body strength to that of competitor frequency in visual recognition. In visual recognition, it is the frequency of a word's neighbours which is important rather than the number of a word's neighbours. Here, it is the regularity of the pronunciation of a word's neighbours (those with which it shares a body) which is the important factor and not the frequency of use of any of the individual items in that set of neighbours. It is not clear however whether this point is sustainable because it begs the question of the possibility of a correlation between type regularity and token frequency. As Andrews (1989) pointed out, regularity effects in the auditory modality and word frequency in the visual modality both affect only low-frequency words and this could be a pointer to some commonality between the two factors.

Seidenberg (1989) went further and identified the effects of repetition at the learning stage as the source of both word frequency and regularity effects. In his connectionist model, patterns which are presented more often at learning, which by definition will be the more frequent and, it is assumed, more regular, have a significant impact on the connection weights which determine the performance of the system. In other words, the system evolves in such a way that word frequency and regularity effects are endemic. However, the case cannot be said to be made until it is established that regular words, like high-frequency words, "occur more often in the language" (Seidenberg, 1989, p. 37).

What is clear is that response to a given word is a function not just of the word itself but of words orthographically related to it. Therefore the processing of an L2 item (by definition low-frequency) may be influenced by the presence of a relatively high-frequency L1 item embedded in it.

The lexicalisation effect

The second reason for anticipating an effect of an embedded word on target recall is to do with the lexicalisation effect (see Connine, 1990; Crowder, 1982; Samuel, 1990). It is something of a tautology to suggest that the language system is designed to seek words but what it entails is that, other things being equal, the system will react more quickly to words than to pseudowords and to pseudowords than to nonwords. If pseudowords are taken to be of ultra-low frequency, the arguments rehearsed in relation to word frequency apply again here. By way of more direct evidence, Baron and Thurston (1973) argued that words and pseudowords share an equal advantage over random strings; however in his review of the issue, Henderson (1982) concluded that although it is clear that people have strong intuitions about characteristic patternings of letters and that these intuitions influence their processing of pseudowords, words nevertheless have an advantage over pseudowords. His tentative explanation for the lexicalisation effect, whether applied to words or pseudowords, was based on the notion of positional or sequential redundancy; to the extent that they resemble words, pseudowords enjoy some of the advantages of redundancy but the redundancy characteristics of words are more robust. Nonwords on the other hand do not have this redundant information at all. Rosson (1983, 1985) showed that the naming latency for a target word is decreased by both words and similar pseudowords and the pronunciation of a pseudoword can be influenced by prior presentation of a visually similar word; she linked these findings to Glushko's (1979, 1981) model of lexical analogy as the basis of nonword pronunciation. The lexicalisation effect therefore implies that the system will attempt to make sense of the input under the influence of lexical constraints; if all other attempts fail, the system will be inclined to match the input string to the nearest best lexical match. This is why readers can cope relatively easily with typographical errors and why accurate proof-reading is a difficult activity.

Some evidence for these claims is to be found in the literature on surface dyslexia. A characteristic of input surface dyslexia is that patients suffer from a deficit in lexical entries which means they have to rely on phonological recoding, of some sort, in order to be able to pronounce words. Despite damage to the lexicon, errors made by dyslexics often appear to be lexically influenced. For example, Kremin (1985) reported the case of the patient H.A.M. who would attempt to guess the meaning of a word which could not be accessed on the basis of its visual similarity to other words. The same patient corrected 70% of word errors made but only 25% of nonword errors; words appear to have a special status in the system. Bub, Cancelliere, and Kertesz (1985) argued that patient M.P., in addition to the purported use of grapheme-phoneme correspondences, also uses "rules" which could only be lexically derived. Thus she pronounces *mild* as *milled* although this pronunciation of the terminal grapheme *ild* is not regular. Saffran (1985) reported the case of the patient L.L. who in reading performance showed a tendency to produce errors resulting in a word rather than a correct pronunciation which would result in a nonword. In a footnote to Saffran's account, Masterson (1985a) pointed out that the lexicalisation tendency, in the sense of the tendency for reading accuracy to be affected by the lexical status of potential alternative responses, may be a property of the purported non-lexical route used by surface dyslexics, but it could equally be a strategy used by subjects when mixed words and nonwords are used as target items. Whether the process is automatic or under strategic control there does seem to be evidence here for a lexical influence on processing. Masterson (1985b) gave the example of lexical effects in a condition where "purely" phonological processing would seem to be more appropriate. In nonword reading, a task well suited to phonological processing, patients E.E. and C.D. appeared to make errors based on visual similarity between the stimulus and the response given, these errors often resulting in (incorrect) real words. Masterson suggested that subjects search for a best match and having encountered an appropriate lexical entry, alter the orthographic entry accordingly to correspond to that best match. Although it is dangerous to base too many conclusions on the often contradictory evidence from brain-damaged patients, it does seem safe to suggest that word representations have some sort of privileged status in the system. Nor is it the case that evidence for the privileged status of words is restricted to studies from brain damaged patients. In the discussion to follow, on word-constituent priming, it is clear that word primes are more effective than other primes under a range of conditions. Monsell (1985) concluded that the evidence points to words having a "psychological reality" (p. 191) mediating between word forms and meanings. In the present domain, what is at issue is not so much a best match between input and stored lexical representations, but the possibility that faced with an unknown string, the French word, the system's attention will be drawn to an automatically processed English word embedded in the French word, even if it has no cognate relationship with that word. This could be because the French word, having no entry in the lexicon, is treated as a nonword by the system, initially at least; or it could be that the English word being more familiar is simply accessed more quickly whether or not the system sees the French word as a nonword.

Word-constituent priming

There is evidence from the phenomenon of word-constituent priming to suggest that subjects do access embedded words (see Caramazza, Laudanna, & Romani, 1988; Henderson, 1985, 1989; Monsell, 1985; Shillcock, 1990). Although the effects of such priming may be transient in the case of single language performance, as will be seen, this transience is to some extent due to speedy identification of the target which effectively inhibits other lexical activity. In the present domain, what is envisaged is the use of such primes when access to the target is slow and difficult or even not forthcoming. As suggested above, the identified embedded word could then be available as a mediating term between the known and the unknown in which case it would be facilitatory; or it could lead to inhibition if subjects wrongly took it as a semantic clue to the target. Word-constituent priming or form priming is described by Forster (1989) as the priming which occurs "when one stimulus facilitates recognition of the other by virtue of similarity of orthographic form" (p. 97). The problem is, of course, that "similarity of orthographic form" is rather vague and could in principle range from two words' sharing one letter, to their sharing the same stem, or to their being identical in form. In effect, many different units of priming have been examined and a case has been made for the full range of possibilities from the individual letter to the whole word; according to Henderson (1985) the "answer to the empirical question about size of priming units is not known" (p. 482). There are two basic lines of explanation for the word-constituent priming phenomenon; one is based on the idea of the decomposition or parsing of complex words into their sub-lexical units in the recognition process; the other is based on the idea of spreading activation. As Seidenberg (1987) pointed out, this division is analogous to that between explanations for phonological recoding based on grapheme-phoneme conversion rules as opposed to those based on analogy.

Explanations based on the idea of access achieved by prior lexical decomposition appear to be on the decline (see Caramazza, Laudanna, & Romani, 1988; Ghatala, Levin, Bell, Truman, & Lodico, 1975). Thus although Taft (1987) continued to argue the case for the Basic Orthographic Syllabic Structure (BOSS), his claims grew progressively weaker as the argument advanced. He defined the BOSS as "the first part of the stem morpheme of a word, up to and including all consonants following the first vowel, but without creating an illegal consonant cluster in its final position" (Taft, 1987, p. 265). It has been shown that nonwords which are BOSSes of real words take longer to reject in a lexical decision task than nonwords which are not BOSSes (Luszcz, Bungey, & Geffen, 1984; Taft & Forster, 1976; Taft, 1979a, 1986), and that words divided after their BOSS (e.g., LANT/ERN) are recognised more quickly than the same words divided after the first syllable (LAN/TERN); however, these effects have not been robust (Lima and Pollatsek, 1983, failed to replicate them). As a result, Taft (1987) acknowledged that there is a problem of definition of the BOSS and he concluded that "the BOSS might best be conceived as being the part of a word that the reader treats as being the stem of the word, even if it is not linguistically a genuine morpheme" (p. 277). It is difficult to see what predictions could be made on the basis of such a specification. In fact, the difficulty of defining both the parsing process and the units of decomposition seems to be insurmountable and attention has progressively moved towards spreading activation models which are not dependent on these concepts.

Seidenberg (1987) argued against the idea of decomposition. He pointed out that experiments designed to establish the size of units have been inconsistent, whether the units are taken to be syllables (Spoehr & Smith, 1973), morphemes, stems (Taft, 1979b), or the Basic Orthographic Syllabic Structure (BOSS) of Taft (1979a). Given this inconsistency, Seidenberg (1987) argued that these units are an emergent property of the process of lexical access rather than a means of achieving it. Orthographic redundancy in a parallel activation model can explain so-called sub-lexical unit effects; there is no need to account for representations of sub-lexical units and no necessity for rules to parse them. In this sense, all that is needed is letters and words. All other units represent coalitions of letters resulting from their co-occurrence in the language. There seems to be no reason, from this account, why embedded words should not emerge in the same way and thus be available for memory purposes.

Marslen-Wilson (1989), also working with a parallel activation model, envisaged "recognisers" for phonological, morphological, syntactic, and semantic information in the input. These recognisers independently feed information into the system and activate a cohort of potential target words in the process; eventually the input will lead to a selection of a best-match. It is a feature of this model that it has built-in "contingency of perceptual choice" (Marslen-Wilson, 1989, p. 7) which means that which word is recognised depends on the overall linguistic context. In the present domain, it is plausible that the embedded English word would have a privileged status given the unfamiliarity of the alternatives.

Even if the parallel activation model is used as an explanation at a general level, more detail is need to cover the range of phenomena observed. As Forster (1989) pointed out, all parallel activation models envisage the output of multiple candidates for selection since activation spreads from letter detectors to all lexical units containing that letter; form priming is endemic in these systems. However, it is not clear why form priming is sometimes found (Evett & Humphreys, 1981; Hillinger, 1980) and sometimes not (Bradley, Savage, & Yelland, 1983; Colombo, 1986); and why only identical form priming is found with four letter words (Forster & Davis, 1984), but form priming for different size units is readily found with eight letter words (Forster, Davis, Schoknecht, & Carter, 1987). Forster concluded that word length as such is not the reason for the difference but neighbourhood density. Short words tend to be more frequent and to have a higher neighbourhood density, which means that they have more competition; any priming effect is thereby dispersed. In support of this argument, Forster, Davis, Schoknecht, & Carter (1987) showed that form priming for short words can be obtained if those words are controlled for low neighbourhood density.

The explanation of Humphreys, Evett, Quinlan, and Besner (1987) was somewhat different. They argued that a distinction needs to be made between effects of masked priming when the prime is not recognised and effects of priming when the prime is recognised. Evett and Humphreys (1981) had shown that target identification could be facilitated when primes contained many of the letters of the target word in the same position in the string; they argued that this was due to a reduction of the threshold of activation of the target by the prime. However, it could also be due to intrusion errors; that is, the correct responses were based on an amalgamation of prime and target information which happened to result in the correct response. Humphreys, Evett, Quinlan, and Besner (1987) found that at short stimulus onset asynchronies, a prime will facilitate recognition of a target with which it shares the same letters, and that this is not an intrusion effect. They argued that the facilitation is due to a minimising of interference which would occur if prime and target differed significantly. However, they showed in a further experiment that with longer stimulus onset asynchronies only identical primes facilitate target recognition. The idea here is that any competition from an orthographically different prime has been dissipated by the time the target is presented therefore more or less similarity (short of identity) will be irrelevant.

A different kind of discussion centres on the issue of how much different kinds of prime are effective in facilitating access to target words. Various accounts have been given. Humphreys, Evett, Quinlan, and Besner (1987) in a masked priming experiment reported 81.77% correct responses in the case of identical primes, 73.33% for graphemically related primes, 57.77% for unrelated primes, and 64.01% for neutral primes. Performance in the identical prime condition was significantly better than the others, $p < 0.05$; performance in the unrelated prime condition was significantly worse than both the related prime condition and the neutral condition, $p < 0.01$. Kirsner, Dunn, and Standen (1987) claimed that the presence of the same morpheme constitutes a sufficient condition for repetition priming whereas this is not true for graphemic information, phonemic information, semantic information, morphologically unrelated spoken and printed forms, synonyms. Monsell (1985) showed that there is evidence for priming at sub-lexical, supra-lexical, and lexical level; however, lexical level priming is both stronger and more persistent than the other two forms. He demonstrated this by considering the case of compound words. Three types of compound words can be distinguished: transparent compounds (e.g., *rope* in *tightrope*); opaque compounds (e.g., *butter* in *butterfly*); and pseudocompounds (e.g., *fur* in *furlong*). Items repeated intact were responded to more quickly than control words and more quickly than constituent-primed words (both $p < 0.01$). Constituent-primed words showed an effect roughly 25% of that of complete word priming. However, this does not seem to reflect a morpheme parsing process because it applied equally to all forms of compound words. Henderson (1985) showed something of the order of effects in his summary of the findings of Glushko (1981) where identical priming is the most effective, followed by words with only an initial segment difference (*fire* - *dire*), a medial segment difference (*dome* - *dime*), and a final segment difference (*daze* - *date*).

What the word-constituent priming data seems to show is that access to a word which shares letters or clusters of letters (right through to identity) with another word can be facilitated by that other word. If this is so, then there would be reason to suppose that the English word *pen* might prime the French word *pencher*, and *vice versa*, although there is no semantic relationship between the two words. Whether or not subjects use the priming effect to mediate between two words in word-pair learning or whether the non-cognate prime interferes with the memory process is the subject of the experiment to follow. Before that, one further matter needs to be discussed and that is the question of whether the difference in language between the embedded word and the prime effectively precludes any word-constituent priming effect. In other words, if lexical representations for L1 and L2 are completely separate, then all of the above may be invalidated.

Between-language priming effects

The fundamental reason for supposing that an English word embedded in a French word will be accessed at input is that the lexical identification process is taken to be initially indiscriminate and automatic. As was discussed earlier, within-language priming effects appear to take place at all levels of representation (see Monsell, 1985) and in most models it is envisaged that all possibilities are computed and a selection is made by the system by competition and/or inhibition of one sort or another. The only reason for supposing that this might not happen when the embedded word is in a different language from the word in question would be if it could be shown that L1 and L2 are separate systems which are unconnected at anything other than the conceptual level. In this case, the representation of an L1 word embedded in an L2 word would not be activated even automatically because the L1 system would be inoperative. The question to be discussed is, therefore, whether the system is language-dependent or language-independent. At the outset it is necessary to clear up what is potentially confusing terminology. From the point of view of the *languages* in question, these can be said to be independent if they do not interconnect at other than the conceptual level and inter-dependent if they do interconnect. From the point of view of the *system*, it is language-independent if it does not distinguish between languages at input, and language-dependent if it is in some way reliant on the language of input. The present discussion will take place from the point of view of the system. The system will be called language-independent if its processing of L1 and L2 allows a large degree of interaction between the two languages; it will be called language-dependent if the opposite is the case.

It is outside the scope of this study to conduct a full review of the question since there is a very large body of relevant literature from both second language learning practitioners and from psycholinguistics dealing with such issues as compound and co-ordinate bilingualism (e.g., Ervin & Osgood, 1954; Weinrich, 1953); the so-called language-switch (e.g., Macnamara, 1967, 1970; Macnamara & Kushnir, 1971); and cross-language Stroop effects (e.g., Gerard & Scarborough, 1989). Nevertheless it is true to say that the conclusion reached in most surveys (e.g., Beauvillain & Grainger, 1987; Gekoski, 1980; Gerard & Scarborough, 1989; McCormack, 1977; Obler, 1984; Schwanenflugel & Rey, 1986) is that the evidence for language dependence is inconclusive and that there is evidence to show that on occasions cross-language effects are obtained which are inconsistent with a language-dependent system (as defined above).

The idea of separate representations for two languages was put forward by Kolers (e.g., 1963, 1966, 1968), Lambert and Fillenbaum (1959), and Tulving and Colotla (1970) among others. More recently, Kirsner, Brown, Abrol, Chadla, and Sharma (1980) found repetition priming in a lexical decision task only for same language primes and not for translations. Kirsner, Smith, Lockhart, King, and Jain (1984) found consistent within-language facilitation but no trace of between-language facilitation. Scarborough, Gerard, and Cortese (1984) found that bilinguals can function like monolinguals and reject nonwords and L2 words with equal facility in a lexical decision task; subjects appear to have the ability to make the L2 system inoperative. Potter, So, von Eckardt, and Feldman (1984) found evidence for connections between languages only at the conceptual level.

On the other hand, there is evidence for between-language priming effects in semantic priming experiments where there is no significant lag between the presentation of the prime and the target. Meyer and Ruddy (1974) showed similarities in between-language performance in word-pair lexical decision tasks. Guttentag, Haith, Goodman, and Hauch (1984) showed between-language effects in a "flanker task"; in this the performance of subjects who were instructed to pay attention only to items in the target language was nevertheless influenced by the presence of flanker items in the second language. Kirsner, Smith, Lockhart, King, and Jain (1984) found between-language effects when subjects were instructed to use translations at the learning stage and found bi-directional priming effects in a mixed-word-pair lexical decision task. Schwanenflugel and Rey (1986) showed no cost of switching languages in a bilingual lexical-decision task, and this regardless of the semantic distance of the target from the prime. Chen and Ng (1989) found between-language facilitation by translations in a lexical-decision task. De Groot and Nas (1991) found that the effect of unmasked cognates is equivalent to that for same-language priming in a lexical-decision task. On a slightly different tack, the Stroop effect between languages has often been demonstrated (e.g., Chen & Ho, 1986; Dyer, 1971; Hamers & Lambert, 1972, 1974; Preston & Lambert, 1969). Beauvillain and Grainger (1987) showed that the system prefers the high-frequency "version" of a bilingual homograph irrespective of the language of input.

In experiments involving repetition effects at relatively longer lags, the evidence is less consistent than it is for semantic priming; however it is clear that between-language effects can be obtained under certain circumstances. Glanzer and Duarte (1971) showed that when same-language items are blocked in a mixed-language learning list, between-language repetitions are more effective than same-language repetitions. Cristoffanini, Kirsner, and Milech (1986) showed that all forms of cognates (that is: identical cognates, regularly derived cognates, and irregularly derived cognates) behave in much the same way as same-language inflections and derivations in terms of repetition effects in a lexical decision task. Gerard and Scarborough (1989) found that cognates and homographic non-cognates both showed a repetition effect over long lags.

The conclusion to be drawn from this evidence is that the system itself is neither language-independent nor language-dependent; it is rather the case that it responds in different ways to different task demands (De Groot & Nas, 1991; Durgunoglu & Roediger, 1987; Gerard & Scarborough, 1989; Snodgrass, 1984). There is a range of variables involved, all of which can affect experimental performance. Semantic priming and repetition priming show different effects. In general terms, early processing of the input appears to be language-independent and between-language effects can be found in both semantic priming and repetition paradigms. At longer lags, other sources of information appear to come into play so that language-dependent effects are more likely to be found (Chen & Ng, 1989; Kirsner, Smith, Lockhart, King, & Jain, 1984; Schwanenflugel & Rey, 1986). Repetition effects directly affect the state of the logogen and are relatively longer lasting whereas semantic priming effects are less directed, more dispersed, and therefore of shorter duration. A distinction is made between data-driven tasks and conceptually-driven tasks (Durgunoglu & Roediger, 1987; Ransdell & Fischler, 1987). A task such as fragment completion, based on a perceptual record, is data-driven and language-dependent; a task such as free recall is conceptually-driven and more likely to be language-independent. In the same vein, Snodgrass (1984) talked about semantic memory (on-line, measured by reaction times) and episodic memory (delayed, measured by errors made). De Groot and Nas (1991) used the concepts of episodic tasks which involve reconstruction of an episode and lexical tasks which relate to automatic priming. Finally, as Chen and Leung (1989) showed, the age of subjects, their proficiency in L2, and the age at which they acquired L2, can all influence experimental performance.

The most parsimonious account of cross-language priming is probably that of Albert and Obler (1978), Hamers and Lambert (1972), Obler (1984), Obler and Albert (1978), Paradis (1977), and Treisman (1969). The language system processes all input in a language-independent manner and a change in the language of input is no more demanding on the system than a change in register within a language. However, for the purposes of this study, all that was necessary to show was that the strong version of the language-dependence hypothesis cannot be upheld. It is clear from this review that between-language effects can be obtained under certain circumstances. Whether or not this obtains in the present domain, and what use, if any, subjects make of it is the subject of the experiment which follows.

EXPERIMENT 10

The principle behind the notion of verbal mediation is that it is easier to associate two items in memory which are perceived as similar than to associate two items perceived as wholly dissimilar (Paivio, 1971). If a French word "contains" a high-frequency English word embedded within it, then the embedded word could affect subject performance in a number of ways provided the system accesses the embedded word, and provided subjects make use of the word when accessed. The neighbourhood effect, the lexicalisation effect, word-constituent priming, and cross-language priming all suggest the possibility of this access occurring. At learning, subjects could then establish an association between the English component of the word-pair and the embedded English word which could be exploited at testing. When the task is generation, the presence of the English cue and the relevant association might allow easy generation of at least part of the French target word. When the task is comprehension, the embedded word might serve as an additional cue for retrieval of the appropriate English target. On the other hand, particularly where comprehension is the task, the embedded English word might function as a false cognate and lead subjects to use it as a semantic clue rather than an associative one and guess incorrectly the meaning of the French word on this basis.

An additional factor examined in the experiment was the influence of word category on performance. On the evidence of the previous experiment, concrete nouns are more easily remembered than verbs. Here the interest was in finding out whether the category effect would be replicated and whether it interacted with embedding. It might be that embedding effects have more influence in the case of the difficult items, verbs, since the embedded word provides a ready source of mediation to overcome the difficulty, than in the processing of items which are relatively easy to learn and recall.

Method

Design

The design of the experiment was a 2 x 2 x 2 factorial design. The between-subjects factor was the test condition. It had two levels, generation and comprehension. As in previous experiments, the effect of embedding could well be different for these two tasks.

The two within-subjects factors were embedding and word category. Embedding had two levels. Ten French items had a high-frequency English word embedded in them; the other 10 did not. Word category had two levels: whether the English word in the stimulus position was a concrete noun (more readily learned and recalled) or a verb (less readily learned and recalled).

Materials

Word-pair lists of 20 items were prepared. Constraints on the choice of items included: that subjects should not have encountered the French items previously; that an equal number of concrete nouns and verbs be available; that the French equivalents for the English items should have the requisite characteristics of either having a high-frequency English word embedded in them or not. Accordingly, L1 cue items were chosen from the low-frequency range of 1-12 per million, with a mean of 6 occurrences per million (Kucera & Francis, 1967). As in the previous experiment, the imageability of the concrete nouns chosen was confirmed by independent judges and with reference to imageability ratings contained in Quinlan (1992).

There were two categories of L2 items. Ten items had a high-frequency L1 word embedded in them (greater than 100 occurrences per million, within the range 102-3001, and with a mean of 576.6 occurrences per million). For example, the French word *borner, to restrict*, has embedded in it the high-frequency English word *born* (113 occurrences per million). Where more than one word was embedded, the higher-frequency word was used as a principle of selection. Thus the French word *ébranler, to rattle*, has embedded in it both the high-frequency English word *ran* (134 occurrences per million), and the low-frequency word *bran* (1 occurrence per million). The word was selected on the basis of the embedded high-frequency word.

Of the other 10 items, three had a low-frequency English word embedded in them and the rest had no English word embedded in them. For example, the French word *picoter, to peck*, has embedded the low-frequency English word *cot* (1 occurrence per million); the French word *fléchir, to sag*, has no English word embedded in it. It was not possible to control the category of embedded words given the other constraints applying.

All subjects learned the same items in a word-pair list; all subjects were tested in a simple sentence context with the necessary adjustments being made for the generation and comprehension tasks.

Subjects

A pool of 51 subjects from School C10 took part in the experiment. Subjects, aged 11-13, were from the same school as used in the previous experiment but none of the subjects had taken part in that experiment. The experiment took place in the Spring term of the school year when subjects had experienced six months of French teaching.

Procedure

Subjects were randomly assigned to two groups. Group 1, tested for generation, was made up of 25 subjects (14 boys, 11 girls). Group 2, tested for comprehension, was made up of 26 subjects (14 boys, 12 girls). The procedure was in other respects as in previous experiments.

Results and discussion

An analysis of variance was performed. The presence of an embedded high-frequency English word had a clear effect on performance, $F\,(1, 49) = 43.19, p < 0.01$ (see Table 8.1). Performance with items containing an embedded word averaged 52.63% as compared with 35.37% for items not containing an embedded word. The results also confirmed the relative learnability of concrete nouns compared with verbs. Subjects recalled 50.66% concrete nouns and 37.34% verbs; this difference was significant, $F\,(1, 49) = 24.48, p < 0.01$ (see Table 8. 2).

Mode of testing had a significant effect on performance $F\,(1, 49) = 4.16, p < 0.05$; (see Table 8.1). Subjects tested for generation averaged 38.20% items recalled and subjects tested for comprehension 49.80%.

Table 8.1. Experiment 10.

Mean percentage scores for items recalled.

	With embedding		**Without embedding**		
Task	**Nouns**	**Verbs**	**Nouns**	**Verbs**	**Overall**
Generation	53.60	40.00	35.20	24.00	38.20
Comprehension	65.38	51.53	48.46	33.84	49.80
Overall	**52.63**		**35.37**		44.00

There was no interaction between embedding and word category, $F\ (1, 49) = 0.02, p > 0.88$. The presence of an embedded English word in the French item appears to have led to better recall independently of the learnability of the item itself. Conversely, concrete nouns were more easily recalled than verbs irrespective of the presence of an embedded English word.

Embedding did not interact with mode of testing, $F\ (1, 49) = 0.00, p > 0.98$. It was equally effective in both the generation and the comprehension task. Word category was also consistent across tasks. There was no interaction between the two factors, $F\ (1, 49) = 0.11, p > 0.73$.

Although the effect of embedding (as opposed to the effect of cognate similarity) is rarely, if ever, mentioned in the second language learning literature, it is very clear from this experiment that the performance of subjects was enhanced by such embedding across difficult and easy categories of vocabulary items, and across the generation and comprehension tasks. The experiment is not able to distinguish between automatic and

Table 8.2. Experiment 10.

Mean percentage scores for items recalled by category.

	Generation		Comprehension		
	With embedding	Without embedding	With embedding	Without embedding	Overall
Nouns	53.60	35.20	65.38	48.46	50.66
Verbs	40.00	24.00	51.53	33.84	37.34

conscious processes involved but it does at least raise questions to be answered by those who argue for a strict separation of language systems. The simplest explanation would seem to be that subjects do access the embedded English word, although they are processing a French word, and that they are able to use the embedded word to assist in their memory processes. A further implication of these results is that the number of factors affecting word learnability is probably much larger than had been anticipated (Rubin, 1980). As psycholinguistics comes to terms with the complexity of language processes, there is every possibility that other factors, such as embedding, will have to be taken into account by those who wish to understand how and why subjects learn some vocabulary items and not others.

CHAPTER 9

Conclusion

It was stated at the outset of this thesis that there was something of a dissociation between the work done in psycholinguistics and theories of second language learning. Research into second language learning, and vocabulary learning in particular, has been described unflatteringly as unsystematic and lacking in scientific rigour (see Broadbent, 1967; Carroll, 1963; Meara, 1980, 1983, 1987; Stern, 1983). On the other hand, psycholinguistic research was rarely directed at second language learning *per se* (Carroll, 1963; Nation, 1982). When second language learning practitioners attempted to make use of it, they often failed to appreciate the complexity of the processes involved and were inclined to simplification (Stern, 1983). The purpose of the study was, therefore, to review some areas of psychological research which appeared to be relevant to second language vocabulary learning and to assess, through a series of experiments, the extent to which that research was applicable in the second language vocabulary learning domain, and in particular to second language vocabulary list learning.

The overall impression gained from the thesis is that psychology has a contribution to make to the domain of second language vocabulary learning, but that this contribution is going to be indirect rather than direct. Psychology cannot tell practitioners what to do, or indeed predict the effectiveness of particular learning techniques, but it can begin to explain why particular procedures may work, and under what conditions, and why those same procedures may not work under other conditions. The reasons for this are threefold. Each aspect of the psychology of second language vocabulary list learning is complex in itself and sensitive to even slight changes in subjects, materials, and tasks. Thus in Experiments 1-4, subjects who were at the same level of learning French, in ostensibly similar schools and similar circumstances, produced significantly different results. Where materials are concerned, Experiments 7 and 8 showed that the frequency of English items alone was not a reliable predictor of performance; this conclusion was reinforced by data from Experiments 9 and 10 which showed the influence of word category and embedded English words on the learning process. While Experiment 1 showed that in general terms generation is a more difficult task than comprehension, it was higher-ability learners who were more disadvantaged in the generation task than lower-ability learners when the test was recall in a simple sentence context (Experiment 2), and lower-ability learners when the task was comprehension (Experiment 3). The activity of second language vocabulary list learning involves several aspects simultaneously, and their interaction is difficult to predict. Finally, a great deal of the most important work in psycholinguistics is done at a low level of description involving reaction times to verbal stimuli as for example in the lexical decision task and the naming task. The activity of second language vocabulary list learning very probably involves a combination of low level automatic processes and processes involving attention, if not conscious strategy, on the part of learners. The combination of factors makes direct application difficult.

The way forward would seem to be that psycholinguistics should pay more direct attention to second language learning in general and to vocabulary learning and list learning in particular. On the other hand, practitioners in the field of second language learning will need to be aware of the complexity of the issues involved and to ask the appropriate questions of psychologists. The appropriate questions in this case are going to be about what outcomes might be expected given these subjects, with this background, under these conditions, and with these tasks in mind; or, alternatively, given these subjects, with this background, under these conditions, and with these tasks, why a particular expected outcome did not materialise. In other words the job of psychology in this domain would seem to be not so much to do with teaching and learning methods and "approaches", in which the second language learning literature abounds, but with the careful analysis of the psychological processes upon which a sound teaching methodology could be based (Hamers & Blanc, 1989).

This thesis was intended as a contribution to that process in the sense that it set out to analyse aspects of the psychology of a particular area of second language learning (second language vocabulary list learning), with a particular set of learners (beginners learning French at school), with two basic tasks (generation and comprehension of the written word), with controlled materials. In one sense, then, the scope of the study was limited. On the other hand, a painstaking bottom-up approach is probably going to be more satisfactory, in the long run, than an over-generalising top-down approach. The effectiveness of the "big picture" is finally dependent on the effectiveness of its constituent parts. There were two main issues addressed in the study: forms of presentation, and word difficulty. These will be treated in turn.

In Chapter 2, some basic issues about word-pair presentation and vocabulary testing were addressed. Given that learners do learn word-pairs, and given that text books and dictionaries use word-pair presentation as an economical way of giving "the meaning" of words, then the question of the order of presentation is important. There has been a considerable amount of research on paired-associate learning, particularly in the Behaviorist tradition and its immediate aftermath, but a great deal of it used either nonsense syllables, or a combination of words or nonsense syllables, or pairs of words which had no relationship to each other outside the particular experimental paradigm. The problem in applying this research is that it is not clear what is the status of the L2 target item for the learner. On the one hand, L2 items may be learned simply as items to be learned in a memory test; on the other hand, subjects may be well aware of the fact that the item to be learned is an alternative way (to the L1 item) of accessing a rich store of conceptual information (though it is doubtful whether any school learner would formulate the issue in quite this way). Since a great deal of the discussion concerns the relative effectiveness of the presentation of the more meaningful component of the word-pair in the stimulus or response position, the status of the L2 item assumes some importance. In the event, it was clear that for both higher-ability and lower-ability learners, the word-pair association was bi-directional but asymmetrical as argued by Johnston (1967), Lockhart (1969), Wolford (1971). There was no basic difference in effectiveness between the L1-L2 or L2-L1 form of presentation measured across the generation and comprehension tasks. However, the forward-association was stronger than the backward-association and the generation task was more difficult that the comprehension task. This meant that L2-L1 learners had more difficulty with the more difficult generation task, which for them meant using the backward-association, than L1-L2 learners had with the comprehension task, which for them meant using the backward-association. The difficulty was more marked for lower-ability learners than it was for higher-ability learners. For this reason, the more conservative approach was adopted in subsequent experiments and items were presented in the order L1-L2.

Many theorists and practitioners express reservations about word-pair list learning. Some of these reservations would appear to be incontrovertible, as, for example, the argument that there is more to learning the meaning of a word than learning its denotation (Balhouq, 1976; Meara, 1980; Nation, 1987; J. C. Richards, 1976). However, less convincing are arguments to the effect that word-pair list learning is of little use, even as part of the process of learning the meaning of a word (Hill, 1965; Hughes, 1968; Judd, 1978; Turner, 1983). They are less convincing because there is a good deal of evidence over many years to suggest that simultaneous presentation of the L1 items and their L2 equivalents is effective (see reviews in Lado, Baldwin, and Lobo, 1967; Mishima, 1966; Nation, 1987).

The psychological discussion underlying this issue is concerned with the notions of encoding specificity (Thomson & Tulving, 1970; Tulving, 1974; Tulving & Osler, 1968; Tulving & Psotka, 1971; Tulving & Thomson, 1973); with form-based repetition priming effects (Feustel, Shiffrin, & Salasoo, 1983; Jacoby, 1983a, 1983b; Kirsner, Dunn, & Standen, 1987); and with logogen-based explanations for repetition priming effects which derive mainly from Morton (1969, 1970). What seems to emerge is that the effectiveness of a particular item as a cue is not to be thought of in absolute terms but as being dependent on task demands. Where an episodic recall is required, a good deal of overlap between the learning and testing conditions is effective; where semantic tasks and indirect recall are concerned, less overlap is necessary.

In Chapter 3, therefore, two experiments were carried out to assess the extent to which word-pair list learning was effective for other than testing in word-pair completion tasks. In other words, the experiment set out to determine whether subjects would be able to generate L2 items and comprehend L2 items when tested in a sentence context or whether their word-pair list-learning would transfer badly to this more "natural" kind of task. The results were both complex and interesting. Higher-ability learners who were tested in a context were less successful than subjects tested in a word-pair completion task when the task was generation but not when the task was comprehension. The opposite was the case for lower-ability learners. An explanation was offered in terms of task-dependent behaviour. Higher-ability learners used a strategy based on an anticipated word-pair list completion test. When this did not materialise, in the generation task, the strategy was not appropriate and performance was inhibited. This disadvantage was offset in the comprehension task by their use of the context provided. Lower-ability learners, on the other hand, did not adopt any particular strategy. This meant, paradoxically, that they were not disadvantaged in the context test where the task was generation. However, in the comprehension test, the combination of the use of the backward-association, the context test, and their inability to use the context provided seems to have resulted in poorer recall. Given this explanation, it seemed important to try to ascertain whether performance would be improved by the provision of a context at learning. This could enable higher-ability learners to adopt an appropriate strategy for testing in context and should offset some of the unfamiliarity of the context test for lower-ability learners.

In Chapter 4, then, the use of context at learning was examined. The issue was of particular interest because of the strong assumption in the second language learning literature, again held over many years, that learning vocabulary in "context" is more "normal" than learning in word-pair lists. This assumption is summed up in the claim of Judd (1978) that "most people agree that vocabulary should be taught in context" (p. 135). The relevant psychological literature appeared to be that involving discussion of elaborated learning, encoding distinctiveness, and transfer-appropriate processing. It could be that provision of a context at testing would enhance recall not only because of the overlap between the learning and testing condition but also because subjects would be encouraged to engage in a kind of processing (more elaborated, more distinctive) which would transfer well to testing in context.

Again, there were important differences between higher-ability and lower-ability learners and between the two tasks. Higher-ability learners who learned in a context recalled more items than learners who learned in a list when the task was generation; where the task was comprehension, the effect was less marked. It could be, then, that higher-ability learners did use an appropriate strategy as a result of the provision of a context at learning and that the advantage of this strategy was less marked in the comprehension task than in the generation task because of the use which list learners could make of the context at testing in the comprehension task. Lower-ability learners, on the other hand, seem to have benefitted little from the provision of a context at learning and indeed they appear to be disadvantaged by departures from familiar ways of learning and testing (see Bialystock, 1985). On the basis of this evidence, therefore, there is a limited case to be made against list learning in the sense that learning in a list may disadvantage higher-ability learners in the generation task. In the comprehension task, and where lower-ability learners are concerned, this is not the case (see Mishima, 1966, for a similar conclusion).

In Chapter 5, list position effects and various forms of word-pair presentation were examined. One possible argument against the use of lists would be that list learning may encourage list dependency (Mandler & Dean, 1969; Murdock, 1962); that is, learners may develop a strategy based, in part at least, on list position and serial order with the result that recall of items when the list order no longer obtains is made more difficult. It might be expected that list dependency would be manifested by the presence of a recency effect, a primacy effect, or a serial order effect, all of which are well documented in the literature on memory.

The evidence against the list learning of word-pairs in Chapter 4 was somewhat mixed. In addition to list dependency, of interest in Chapter 5 was the possibility that one of the reasons why the provision of a context at learning does not necessarily affect performance, particularly where lower-ability learners in concerned, is that learners perceive the learning task as a word-pair learning task, whatever the form of presentation, or that they provide their own context, or both (Gershman, 1970; Lado, Baldwin, & Lobo, 1967; Mishima, 1966).

Experiment 5 showed a clear primacy effect in the generation task, as well as a serial order effect. It appeared that subjects, in this case lower-ability subjects, adopted a strategy of starting at the beginning of the list and working their way through since items at the beginning of the list and in the middle of the list were better recalled than items at the end of the list. In Experiment 6, with higher-ability subjects, and in the comprehension task, there was again a primacy effect, but performance otherwise was spread over the rest of the list. In other words, higher-ability learners appeared to be less dependent on serial order as such, although items at the start of the list seem to have been given special attention.

Item presentation as such, that is in a word-pair list, in a simple sentence context, and in a disrupted format, did not have a significant effect in either experiment. These results are compatible with the results from Experiment 4 in the sense that lower-ability learners in the generation task (Experiment 5) probably treated all forms of presentation in the same way, and for higher-ability learners in the comprehension task (Experiment 6), any word-pair presentation effects would be offset by the availability of a context at testing. Either way, simple word-pair presentation appears to have been as effective as either the simple context, or the disrupted list.

The second main issue to be addressed in the study was word difficulty and Chapter 6 considered the role of word frequency in item learning. While the effect of word frequency is well established in the psychological literature (Besner & McCann, 1987; Whaley 1978), and is found across a range of experiments (see review in Monsell, 1991), the locus of word frequency is a subject of much debate. In particular, there is disagreement about whether word frequency affects the state of the logogen, in which case it would be relevant in the present domain, or whether word frequency effects arise at a decision stage or production stage in the language process (for example, Balota & Chumbley, 1984, 1985; Besner & McCann, 1987), in which case their influence would be relatively unimportant in this domain. In the second language learning literature, there is unease about the reliability of written-word frequency counts where language learning is concerned since there seems to be a discrepancy between objective word frequency counts and subjective word frequency, or familiarity, counts. Finally, whichever way word frequency is established, it is unclear how it will affect the learning and recall of vocabulary items since learning and recall involve processes which are much more extended over time than the reaction-time experiments by which word frequency has usually been studied.

The opportunity was also taken to study further list position effects. In previous experiments, conclusions about list position had been based on lists of words where word frequency had not been a major concern. The interest here was in seeing whether word frequency interacted with list position, and if so, whether word frequency or list position was a more important indicator of word learnability. In the event, in Experiment 7, with learners of intermediate ability and with a generation task, word frequency had a significant effect on performance with high-frequency words being significantly better recalled than low-frequency words. The effect of word frequency was particularly marked in the primacy and recency positions. Where list position was concerned, there was a significant advantage for items at the end of the list over items in both other positions. In other words, in relation to list position, these results were at odds with those from Experiments 5 and 6. It appeared that "natural" options for organising learning were to concentrate on items at either the beginning or end of the list but that neither was obligatory.

In Experiment 8, with higher-ability learners and a comprehension task, word frequency had the opposite effect to the one anticipated with low-frequency words significantly better recalled than high-frequency words. It was suggested that this could have been due to subjects' adopting a strategy of concentrating on what they perceived to be the more difficult word-pairs (that is word-pairs where the English component was of low-frequency) particularly at the end of the list. Items in the middle of the list were relatively unaffected by word frequency in comparison with items at the beginning and end of the list.

What emerges from Experiments 5-8 is, then, that it is possible on occasions to show list position effects, serial effects, positive effects of word frequency, and negative effects of word frequency. However, each of these factors can interact with subjects, tasks, and possibly materials. This means that list position, serial position, and word frequency as such are not reliable indicators of word learnability.

In Chapter 7, word category was examined as an indicator of word learnability. There seems to be general agreement in the second language literature that certain categories of words, notably concrete words, are more easily learned than abstract or function words. There is a similar agreement in the psycholinguistic literature. What is more difficult to establish is why certain categories of words are more easily learned; this is a point strongly made by Rubin (1980). The dual-coding approach of Paivio (passim) is questioned by many (for example, Anderson & Bower, 1973; Pylyshyn, 1973; Richardson, 1975a; Winograd, Cohen, & Barresi, 1976). Attempts to distinguish between imageability, concreteness, operativity, and word frequency have mainly been confounded. It was suggested that the relative ease of learning of concrete words compared with abstract words is to do with context availability (Schwanenflugel & Shoben, 1983), with different kinds of representation in the system (De Groot, 1989), with ease of predication (Jones, 1985), and with hemisphere differences (Coltheart, 1987c; Marcel & Patterson, 1978; Patterson & Marcel, 1977; Shallice & Warrington, 1975; Warrington, 1981).

Whatever the reason for the learnability of certain categories of words, it is clear from Experiment 9 that word category has a significant effect on learning and recall, with word frequency held constant, and independently of whether the task is generation or comprehension. This appears to be a promising area for further investigation. In particular, it is likely that useful discriminations can be made among what are here classed as "other words" and what are referred to in the literature as "function words" (see also Patterson, 1981). The word category effect could also, perhaps, shed some light on the unexpected performance of subjects with high-frequency words in Experiment 8 since an analysis of recall arranged by word category found that of the seven high-frequency words of below average recall in this experiment, four were abstract nouns and two were verbs; these were two of the the more difficult categories of words in terms of learnability.

Word category was returned to in Chapter 8 along with another possible aspect of word learnability which was suggested by psycholinguistic theories on verbal mediation, spreading activation, and multiple lexical access. There are reasons for believing that the presence of an English word embedded in a French word might affect learning and recall of that French word even if the English word and French word are not semantically related. These reasons include the neighbourhood effect (Andrews, 1989); pronounceability (Henderson, 1982, 1985; Service, 1992); lexicalisation (Connine, 1990; Crowder, 1982; Samuel, 1990); word-constituent priming (Henderson, 1985, 1989; Monsell, 1985; Shillcock, 1990); and cross-language effects (De Groot & Nas, 1991). In Experiment 10, French words with a high-frequency English word embedded in them were better recalled than French words without this embedding and the effect was consistent across tasks and word category. Again, concrete nouns were better recalled than verbs, independently of the embedding effect, and across tasks. The embedding effect, like the word category effect, would appear to be a promising area for further investigation into word learnability.

All of these experiments have been conducted at a relatively high level of description. That is, the tasks concerned were complex tasks in which time constraints were not tight. As a result there was a good deal of scope for subjects to engage in learning strategies and for factors to interact. Nevertheless, it was possible to call in question some of the widely-held assumptions in the second language learning domain about the presentation of vocabulary items and in particular about the relative merits of list learning and learning in context. From the point of view of psychology, although word frequency effects are well established in the psycholinguistics literature, particularly with respect to simple tasks such as lexical decision and naming, they were inconsistent in the present domain. On the other hand, word category effects and embedding effects were observed. The way forward would be to conduct further experiments into word frequency, word category effects, and embedding with a wider range of subjects and at a lower level of description where more control could be exercised and a clearer picture obtained of what effects are automatically induced and which the result of strategic decisions by learners.

Another area deserving investigation, but one outside the scope of this thesis, concerns the difference between what have been called here higher-ability and lower-ability learners. It was stated at the outset that this terminology "refers merely to subjects' performance in the particular task under consideration; higher-ability subjects are those who performed more successfully; lower-ability subjects are those who performed less successfully" (p. 20). The interesting question is why subjects perform differently when they are so similar in age and background, and apparently comparable in general academic ability. Throughout this thesis, an explanation is given in terms of subject strategy (as for example on p. 107 at the conclusion of Chapter 4) but this explanation can only be partial since it simply leads to the further question of why subjects similar in the ways mentioned should adopt different strategies. There are a number of possibilities. It is worth noting that with the exception of Experiment 1, subjects from the all-girls school (School B) out-performed those from the mixed-sex school (School A) and that there was no significant difference between subjects from the two mixed-sex schools in Experiment 5 (School A and School C). It would be worth pursuing this line of enquiry by performing similar experiments with gender as a factor. At a more general level, it should be remembered that "streaming" within schools is not necessarily a reliable guide to comparability between schools. In other words, it could be that general academic ability is a factor in performance and that this does indeed manifest itself in students' ability to adopt more or less suitable and effective strategies. This could only be established by more elaborate pre-testing of general academic ability. At a more general level again, there are what might be called "milieu effects" which are very difficult to specify. Subjects were taken from different cohorts and it is clear that from year to year different cohorts within nominally the "same" stream may well differ in attitude and motivation. Schools too differ in their teaching methodologies and academic ethos and these could contribute to the kind of learning undertaken by subjects. The complexity of these issues reinforces the notion that generalisations will be hard to come by in this domain and must be based on experiments which take into account the wide range of potential factors involved.

ASPECTS OF THE PSYCHOLOGY OF SECOND LANGUAGE VOCABULARY LIST LEARNING

BIBLIOGRAPHY AND APPENDICES

TABLE OF CONTENTS

BIBLIOGRAPHY

The presentation of the thesis is in accordance with the general instructions contained in *Requirements for the Presentation of Research Theses*, University of Warwick, 1989.

Referencing, punctuation, and presentation of data follow the guidelines of Sternberg (1988), based on *Publication Manual of the American Psychological Association*, (3rd. edition, 1983).

Adams, J. A. (1967). *Human Memory*. New York: McGraw Hill.

Albert, M. L., & Obler, L. K. (1978). *The Bilingual Brain*. New York: Academic Press.

Allen, V. F. (1983). *Techniques in teaching vocabulary*. Oxford: Oxford University Press.

Anderson, J. P., & Jordan, A. M. (1928). Learning and retention of Latin words and phrases. *Journal of Educational Psychology*, **19**, 485-496.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ.: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA.: Harvard University Press.

Anderson, J. R. (1985). *Cognitive Psychology and its Implications*. New York: Freeman.

Anderson, J. R. (1990). *Cognitive psychology and its Implications*. New York: Freeman.

Anderson, J. R., & Bower, G. H. (1973). *Human Associative Memory*. Washington, DC.: Winston.

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews*. Newark, DE.: International Reading Association.

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles: A problem of polysemy. *Cognitive Psychology*, **7**, 167-180.

Andreewsky, E., Deloche, G., & Kossanyi, P. (1987). Analogies between speed-reading and deep dyslexia: Towards a procedural understanding of reading. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia*. (Second Edition). London: Routledge & Kegan Paul.

Andrews, S. (1989). Frequency and neighbourhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 802-814.

Anglin, J. M. (1970). *The growth of word meaning.* Cambridge, MA.: MIT.

Appel, R., & Muysken, P. (1987). *Language contact and bilingualism.* London: Arnold.

Asch, S. (1968). The doctrinal tyranny of associationism: Or what is wrong with rote learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, NJ.: Prentice-Hall.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society,* **106**, 135-163.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist,* **30**, 821-837.

Baddeley, A. D. (1976). *The psychology of memory.* New York: Basic Books.

Baddeley, A. D., Ellis, N. C., Miles, T. R., & Lewis, V. J. (1982). Developmental and acquired dyslexia: A comparison. *Cognition,* **11**, 185-199.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory.* Volume 8. New York: Academic.

Baddeley, A. D., & Hitch, G. J. (1977). Recency reexamined. In S. Dornic (Ed.), *Attention and Performance VI.* Hillsdale, NJ.: Erlbaum.

Bahrick, H. P. (1969). Measurement of memory by prompted recall. *Journal of Experimental Psychology,* **79**, 213-219.

Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review,* **77**, 215-222.

Bahrick, H. P. (1979). Broader methods and narrower theories for memory research: Comments on the papers by Eysenck and Cermak. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Bahrick, H. P. & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **13**, 344-349.

Balhouq, S. A. (1976). *The place of lexis in foreign language acquisition.* Unpublished M.A. thesis. University of Sheffield.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance,* **10**, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language,* **24**, 89-106.

Balota, D. A., & Chumbley, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, & Haggard (1989). *Journal of Experimental Psychology: General,* **119**, 231-237.

Banks, W. P., and White, H. (1982). Single ordering as a processing limitation. *Journal of Verbal Learning and Verbal Behavior,* **21**, 39-54.

Banta, F. G. (1981). Teaching German vocabulary: The use of English cognates and common loan words. *Modern Language Journal,* **65**, 129-136.

Baron, J., & Thurston, I. (1973). An analysis of the word superiority effect. *Cognitive Psychology,* **4**, 207-228.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition,* **10**, 82-93.

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

Battig, W. F. (1966). Evidence for coding processes in 'rote' paired-associate learning. *Journal of Verbal Learning and Verbal Behavior,* **5**, 177-181.

Battig, W. F. (1968). *Paired-associate learning.* In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, NJ.: Prentice-Hall.

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, **26**, 658-672.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **2**, 556-566.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 252-259.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 389-401.

Begg, I., & Paivio, A. (1969). Concreteness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior*, **8**, 821-827.

Bellezza, F. S. (1983). Mnemonic-device instruction with adults. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Psychological foundations*. New York: Springer-Verlag.

Bellezza, F. S., Cheesman, F. L., & Reddy, B. G. (1977). Organization and semantic elaboration in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 539-550.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading*. Hillsdale, NJ.: Erlbaum.

Bialystock, E. (1985). The compatibility of teaching and learning strategies. *Applied Linguistics*, **6**, 255-262.

Biemiller, A. (1977-78). Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly,* **13**, 223-253.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long term free recall. *Cognitive Psychology,* **6**, 173-189.

Blum, S. & Levenston, E. A. (1978). Universals of lexical simplification. *Language Learning,* **28**, 399-415.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology,* **49**, 229-240.

Bousfield, W. A., Cohen, H. H., & Whitmarsh, G. A. (1958) *Psychological Reports,* **4**, 39-44.

Bower, G. H. (1967). A multicomponent theory of the memory trace. In K. W. Spence, & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory.* Volume 1. New York: Academic Press.

Bower, G. H. (1972). Mental imagery and associative learning. In L. W. Gregg (Ed.), *Cognition in learning and memory.* New York: Wiley.

Bower, G. H., Clark, M. C., Lesgold, A. M. & Winenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior,* **8**, 323-343.

Bradley, D. C., Savage, G. R., & Yelland, G. W. (1983). Form-priming or not? Paper delivered at Fourth Language and Speech Conference, Monash University. [Cited in K. I. Forster, 1989, Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process.* Cambridge, MA.: MIT Press.]

Bradley, M. M., & Glenberg, A. M. (1983). Strengthening associations: Duration, attention, or relations. *Journal of Verbal Learning and Verbal Behavior,* **22**, 650-666.

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cermak & F. I. M. Craik, (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Bransford, J. D., & McCarrell, N. S. (1974). A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. Hillsdale, NJ.: Erlbaum.

Briones, I. T. (1937). An experimental comparison of two forms of linguistic meaning. *Psychological Record,* **1**, 205-214.

Broadbent, D. E. (1967). Notes on current knowledge concerning the psychology of learning modern languages. *Manpower and Applied Psychology*, **1**, 148-157. [Cited in P. Meara, (1985), Some priorities for research in second language learning. In C. Brumfit, H. Lunt, and J. Trim (Eds.), *Second language learning: Research problems and perspectives*. London: CILT.]

Broadbent, D. E., & Broadbent, M. H. P. (1980). Priming and the passive/active model of word recognition. In R. S. Nickerson (Ed.), *Attention and Performance VIII*. Hillsdale, NJ.: Erlbaum.

Brown, A. L. (1979). Theories of memory and the problems of development: Activity, growth, and knowledge. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology,* **55**, 1-5.

Bub, D., Cancelliere, A., & Kertesz, A. (1985). Whole-word and analytic translation of spelling to sound in a non-semantic reader. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

Bugelski, B. R. (1962). Presentation time, total time, and mediation in paired-associate learning. *Journal of Experimental Psychology,* **63**, 409-412.

Bugelski, B. R., Kidd, E., & Segmen, J. (1968). Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology,* **76**, 69-73.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition,* **28**, 297-332.

Carr, T. H., & Pollatsek, A. (1985). Recognising printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. McKinnon (Eds.), *Reading research.* Vol. 5. Orlando: Academic Press.

Carroll, J. B. (1963). Research on teaching foreign languages. In N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally.

Carroll, J. B. (1966). Language teaching: Broader contexts. *North East Conference on the teaching of foreign languages; reports of the working committees*. Cited in S. J. Gershman (1970), *Foreign language vocabulary learning under seven conditions*. Unpublished Ph. D. thesis. Columbia University.

Carroll, J. B. (1968). Word frequency studies and the lognormal distribution. In E. M. Zale (Ed.), *Proceedings of the conference on language and language behavior.* New York: Appleton-Century-Croft.

Carroll, J. B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior,* **10**, 722-729.

Carroll, J. B. (1972). A new word frequency book. *Elementary English,* **49**, 1070-1074.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book.* New York: American Heritage.

Carter, R., & McCarthy, M. (1988). Word lists and learning words: Some foundations. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching*. London: Longman.

Cason, H. (1933). Association between the familiar and unfamiliar. *Journal of Experimental Psychology*, **16**, 295-305.

Chen, H-C. (1990). Lexical processing in a non-native language: Effects of language proficiency and learning strategy. *Memory and Cognition,* **18**, 279-288.

Chen, H-C., & Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **12**, 397-401.

Chen, H-C., & Leung, Y-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 316-325.

Chen, H-C., & Ng, M-L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory and Cognition,* **17**, 454-462.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory and Cognition,* **12**, 590-606.

Cieutat, V. J., Stockwell, F. E., & Noble, C. E. (1958) The interaction of ability and amount of practice with stimulus and response meaningfulness (*m, m'*) in paired-associate learning. *Journal of Experimental Psychology,* **56**, 193-202.

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

Clarke, R., & Morton, J. (1983) Cross modality *facilitation in tachistoscopic word* recognition. *Quarterly Journal of Experimental Psychology,* **35A**, 79-96.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology,* **71**, 858-866.

Cohen, B. H. (1966). Some-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior,* **5**, 182-187.

Cohen, A. D., & Aphek, E. (1979). *Easifying second language learning.* Hebrew University. [Reprinted in *Studies in Second Language Acquisition,* **3**, 221-236. 1981.]

Cohen, A. D., & Aphek, E. (1980). Retention of second language vocabulary over time: Investigating the role of mnemonic association. *System,* **8**, 221-235.

Cole, R. A., & Jakimik, J. A. (1978). Understanding speech: How words are heard. In G. Underwood (Ed.), *Strategies of information processing.* London: Academic.

Cole, R. A., & Jakimik, J. A. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech.* Hillsdale, NJ.: Erlbaum.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Pyschological Review,* **82**, 407-428.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance,* **12,** 226-234.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing*. London: Academic Press.

Coltheart, M. (1980). Deep dyslexia: A right-hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul.

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M. I. Posner and O. S. M. Marin (Eds.), *Mechanisms of attention: Attention and Performance XI.* Hillsdale, NJ.: Erlbaum.

Coltheart, M. (1987a). Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Coltheart, M. (1987b). The semantic error: Types and theories. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Coltheart, M. (1987c). Deep dyslexia: A right-hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI.* Hillsdale, NJ.: Erlbaum.

Coltheart, M., & Funnell, E. (1987). Reading and writing: One lexicon or two? In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language Perception and Production.* London: Academic Press.

Coltheart, M., Patterson, K., & Marshall, J. C. (Eds.). (1987). *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Connine, C. (1990). Effects of sentence context and lexical knowledge in speech processing. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, MA.: MIT Press.

Corder, S. P. (1973). *Introducing Applied Linguistics.* Harmondsworth: Penguin.

Cornu, A-M. (1979). The first step in vocabulary teaching. *Modern Language Journal,* **63**, 262-272.

Craik, F. I. M. (1968). Two components in free recall. *Journal of Verbal Learning and Verbal Behavior,* **7**, 996-1004.

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior,* **9**, 143-148.

Craik, F. I. M. (1979). Levels of processing: Overview and closing comments. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Craik, F. I. M. & Levy, B. A. (1970). Semantic and acoustic information in primary memory. *Journal of Experimental Psychology,* **86**, 77-82.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* **11**, 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General,* **104**, 268-294.

Cramer, P. (1968). *Word association.* New York: Academic Press.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology,* **38A**, 367-393.

Crothers, E., & Suppes, P. (1967). *Experiments in second-language learning.* New York: Academic.

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica,* **50**, 291-323.

Crowder, R. G. (1982). *The psychology of reading*. Oxford: Oxford Univesity Press.

Curran, C. (1976). *Counselling-learning in second languages*. Apple River, Illinois: Apple River Press.

Curtis, M. E. (1987). Vocabulary testing and vocabulary instruction. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. Hillsdale, NJ.: Erlbaum.

Davis, F. B. (1944). The interpretation of frequency ratings obtained from "The Teacher's Word Book". *Journal of Educational Psychology*, **35**, 169-174.

Deese, J. (1957). Serial organization in the recall of disconnected items. *Psychological Reports,* **3**, 577-582.

Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, **5**, 305-312.

Deese, J. (1962). Form class and the determinants of association. *Journal of Verbal Learning and Verbal Behavior,* **1**, 79-84.

Deese, J. (1968). Association and memory. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory.* Englewood Cliffs, NJ.: Prentice-Hall.

De Groot, A. M. B. (1989), Representational aspects of word imageability and word frequency as assessed through word association. *Journal Of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 824-845.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language,* **30**, 90-123.

Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, **19**, 608-623.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition*, **8**, 195-209.

Doehring, D. G. (1976). Acquisition of rapid reading responses. *Monographs of the Society for Research in Child Development*, **41**. [Cited in K. T. Spoehr, & R. E. Schuberth, 1981, Processing words in context. In O. J. L. Tzeng, & H. Singer, (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.]

Drewnowski, A., & Healy, A. F. (1977) Detection errors on *the* and *and*: Evidence for reading units larger than the word. *Memory and Cognition*, **5**, 636-647.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. Hillsdale, NJ.: Erlbaum.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language Two*. Oxford: Oxford University Press.

Durgunoglu, A. Y., & Roediger, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language*, **26**, 377-391.

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior*, **10**, 297-302.

Ebbinghaus, H. (1885). *Memory*. New York: Teachers College Press.

Eich, J. E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory and Cognition*, *8*, 157-173.

Ellis, H. D., & Shepherd, J. W. (1974). Recognition of abstract and concrete words presented in left and right visual fields. *Journal of Experimental Psychology*, **103**, 1035-1036.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford Univesity Press.

Elman, J. L., & McClelland, J. L. (1986). Exploiting the lawful variability in the speech wave. In J. S. Perkell, and D. H. Klatt (Eds.), *Invariance and variability of speech processes*. Hillsdale, NJ.: Erlbaum.

Elshout-Mohr, M., & Van Daalen-Kapteijns, M. M. (1987). Cognitive processes in learning word meanings. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition.* Hillsdale, NJ.: Erlbaum.

Ervin, S., & Osgood, C. E. (1954). Second language learning and bilingualism. Supplement to the *Journal of Abnormal and Social Psychology,* **49**, 139-146.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology,* **33A**, 325-350.

Eysenck, M. W. (1978). Levels of processing: A critique. *British Journal of Psychology,* **69**, 157-169.

Eysenck, M. W. (1979). Depth, elaboration, and distinctiveness. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Eysenck, M. W. (1984). *A handbook of cognitive psychology.* Hillsdale, NJ.: Erlbaum.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General,* **112**, 309-346.

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior,* **18**, 1-20.

Fischler, I., & Goodman, G. O. (1978). Latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception and Performance,* **4**, 455-470.

Fodor, J.A. (1975). *The language of thought.* New York: Cromwell.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales, & E. Walker, (Eds.), *New approaches to language mechanisms.* Amsterdam: New Holland.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper, & E. C. T. Walker, (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ.: Erlbaum.

Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading*. Hillsdale, NJ.: Erlbaum.

Forster, K. I. (1989). Basic issues in lexical processing. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA.: MIT Press.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior,* **12**, 627-635.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **10**, 680-698.

Forster, K. I., Davis, C., Schocknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology,* **39A**, 211-251.

Forster, K. I., & Dickinson, R.G. (1976). More on the language-as-fixed-effect fallacy. *Journal of Verbal Learning and Verbal Behavior,* **15**, 135-142.

Foss, D. J. (1982). A discourse on semantic priming. *Cognitive Psychology,* **14**, 590-607.

Frederiksen, J. R. (1978). Word recognition in the presence of semantically constraining context. Paper presented at the annual meeting of the Psychonomic Society, San Antonio, Texas, November, 1978. [Cited in K. T. Spoehr, and R. E. Schuberth, 1981, Processing words in context. In O. J. L. Tzeng, and H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.]

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance,* **2**, 361-379.

Fry, E. B. (1960). A study of teaching machine response modes. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning*. Washington: National Education Association.

Galbraith, R. C., & Underwood, B. J. (1973). Perceived frequency of concrete and abstract words. *Memory and Cognition,* **1**, 56-60.

Gannon, D. R., & Noble, C. E. (1961). Familiarization (*n*) as a stimulus factor in paired-associate verbal learning. *Journal of Experimental Psychology,* **62**, 14-23.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory and Cognition,* **15**, 24-28.

Garnham, A. (1985). *Psycholinguistics: Central topics.* London: Methuen.

Gekoski, W. L. (1980). Language acquisition context and language organization in bilinguals. *Journal of Psycholinguistic Research,* **9**, 429-449.

George, H. V. (1972). *Common errors in language learning.* Rowley, MA.: Newbury House.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 305-315.

Gernsbacher, M. A. (1983). *The experiential familiarity norms and their psychological reality.* Paper presented at the 29th Annual Meeting of the Southwestern Psychological Association, San Antonio, TX. [Reported in Gernsbacher, 1984.]

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General,* **113**, 256-281.

Gershman, S. J. (1970), *Foreign language vocabulary learning under seven conditions.* Unpublished Ph. D. thesis. Columbia University.

Ghatala, E. S., & Levin, J. R. (1976). Phenomenal background frequency and the concreteness/imagery effect in verbal discrimination learning. *Memory and Cognition,* **4**, 302-306.

Ghatala, E. S., Levin, J. R., Bell, J. A., Truman, D. L. & Lodico, M. G. (1975). The effect of semantic and nonsemantic factors on the integration of verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 647-655.

Glanzer, M. (1962). Grammatical category: A rote learning and word association analysis. *Journal of Verbal Learning and Verbal Behavior,* **1**, 31-41.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and theory.* Volume 5. New York: Academic Press.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory,* **2**, 21-31.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior,* **5**, 351-360

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning and Verbal Behavior,* **10**, 625-630.

Glanzer, M., & Meinzer, A. (1967). The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior,* **6**, 928-935.

Glushko, R. J. (1979) The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance,* **5**, 674-691.

Glushko, R. J. (1981). Principles for pronouncing print: The psychology of phonography. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ.: Erlbaum.

Godden, D. R., & Baddeley, A. D. (1975). Context dependent memory in two natural environments: On land and underwater. *British Journal of Psychology,* **66**, 325-331.

Goodglass, H., & Geschwind, N. (1976). Language disorders (aphasia). In E. Carterette & M. Friedman (Eds.), *Handbook of Perception*, Volume 7. New York: Academic Press.

Goodglass, H., Hyde, M. R., & Blumstein, J. (1969) Frequency, picturability and availability of nouns in aphasia. *Cortex*, **5**, 104-119.

Goodman, K. (1976). Reading: A psycholinguistic guessing game. In H. Singer, and R. Ruddell (Eds.), *Theoretical models and processes of reading*. (Second edition). Newark, Del.: International Reading Association.

Goodman, K. S., & Bird, L. B. (1984). On the wording of texts: A study of intra-text word frequency. *Research in the Teaching of English*, **18**, 119-145.

Goodwin, D. W., Powell, B., Breme, D., Hoine, H., & Stern, J. (1969). Alcohol and recall: State dependent effects in man. *Science*, ***163***, *1358*.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language*, **24**, 631-645.

Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, **61**, 23-29.

Goss, A. E., & Nodine, C. F. (1965). *Paired-associate learning: The role of meaningfulness, similarity, and familiarization*. New York: Academic.

Gough, P. B., Alford, J. A., & Holley-Wilcox, P. (1981). Words and contexts. In O. J. L. Tzeng, & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.

Gouin, F. (1880). *L'art d'enseigner et d'étudier les langues*. Paris. [Cited in H. H. Stern, 1983, *Fundamental Concepts of Language Teaching*. Oxford: Oxford University Press.]

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, **29**, 228-244.

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology*, **39A**, 295-319.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighbourhood frequency effect. *Perception and Psychophysics,* **45**, 189-195.

Graves, M. F., Boettcher, J. V., Ryder, R. J., & Peacock, J. L. (1980). Word frequency as a predictor of students' reading vocabularies. *Journal of Reading Behavior,* **12**, 117-127.

Graves, M. F., Ryder, R. J., & Slater, W. F. (1983). Family frequency as a predictor of word difficulty. In J. A. Niles & L. A. Harris (Eds.), *Searches for meaning in reading/language processing*. Rochester, NY.: National Reading Conference.

Graves, M. F., Ryder, R. J., Slater, W. H., & Calfee, R. C. (1987). The relationship between word frequency and reading vocabulary using six metrics of frequency. *Journal of Educational Research,* **81**, 81-90.

Green, J. F. (1970). The use of the mother tongue and the teaching of translation. *English Language Teaching,* **24**, 217-223.

Guilford, F. P., & Zimmerman, W. S. (1948). The Guilford-Zimmerman aptitude survey. *Journal of Applied Psychology,* **32**, 24-34.

Guttentag, R. E., Haith, M. M., Goodman, G. S., & Hauch, J. (1984). Semantic processing of unattended words by bilinguals: A test of the input switch mechanism. *Journal of Verbal Learning and Verbal Behavior,* **23**, 178-188.

Hall, J. F. (1971). *Verbal learning and retention.* Philadelphia, PA: Lippincot.

Hamers, J. F., & Blanc, M. H. A. (1989). *Bilinguality and bilingualism.* Cambridge: Cambridge University Press.

Hamers, J. F., & Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior,* **11**, 303-310.

Hamers, J. F., & Lambert, W. E. (1974). Bilingual reactions to cross-language semantic ambiguity. In S. T. Carey (Ed.), *Bilingualism, Biculturalism and Education.* Edmonton: University of Alberta Press.

Hatch, E. M. (1983). *Psycholinguistics: A second language perspective.* Rowley, MA.: Newbury House.

Heim, A. W., Watts, K. P., Bower, I. B., & Hawton, K. E. (1966). Learning and retention of word-pairs with varying degrees of association. *Quarterly Journal of Experimental Psychology,* **18**, 193-205.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Henderson, L. (1985). Issues in the modelling of pronunciation assembly in normal reading. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia.* London: Erlbaum.

Henderson, L. (1989). On mental representation of morphology and its diagnosis by measures of visual access speed. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA.: MIT Press.

Henning, G. H. (1973). Remembering foreign language vocabulary: Acoustic and semantic parameters. *Language Learning,* **23**, 185-196.

Higa, M. (1965). The psycholinguistic concept of 'difficulty' and the teaching of foreign vocabulary. *Language Learning,* **15**, 167-169.

Hill, L. A. (1965). *A picture vocabulary.* London: Oxford University Press.

Hillinger, M. L. (1980). Priming effects with phonemically similar words: The encoding-bias hypothesis reconsidered. *Memory and Cognition,* **8**, 115-123.

Hines, D. (1976). Recognition of verbs, abstract nouns, and concrete nouns from the left and right visual half fields. *Neuropsychologia,* **14**, 211-216.

Hines, D. (1977). Differences in tachistoscopic recognition between abstract and concrete words as a function of visual half-field and frequency. *Cortex,* **13**, 66-73.

Holley, F. M. (1973). A study of vocabulary learning in context: The effect of new-word density in German reading materials. *Foreign Language Annals,* **6**, 339-347.

Holley, F. M., & King, J. K. (1971). Vocabulary glosses in foreign language reading materials. *Language Learning,* **21**, 213-219.

Holmes, V. M., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behavior,* **15**, 559-566.

Horowitz, L. M., & Gordon, A. M. (1972). Associative symmetry and second language learning. *Journal of Educational Psychology*, **63**, 287-294.

Horowtiz, L. M., & Newman, W. (1969). An interrupted stimulus can facilitate PA learning. *Journal of Verbal Learning and Verbal Behavior*, **8**, 219-224.

Howes, D. (1954). On the interpretation of word frequency as a variable affecting speed of recognition. *Journal of Experimental Psychology,* **48**, 106-112.

Hughes, J. P. (1968). *Linguistics and language teaching.* New York: Random House.

Humphreys, G. W., Besner, D., & Quinlan, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General,* **117**, 51-67.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioral and Brain Sciences,* **8**, 689-740.

Humphreys, G. W., Evett, L. J., Quinlan, P. T., & Besner, D. (1987). Orthographic priming: Qualitative differences between priming from identified and unidentifed primes. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* Hillsdale, NJ.: Erlbaum.

Hyde, T. S., & Jenkins, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology,* **82**, 472-481.

Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior,* **12**, 471-480.

Jacoby, L. L. (1983a). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning Memory and Cognition,* **9**, 21-38.

Jacoby, L. L. (1983b). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior,* **22**, 485-508.

Jacoby, L. L., & Bartz, W. H. (1972). Rehearsal and transfer to LTM. *Journal of Verbal Learning and Verbal Behavior*, **11**, 561-565.

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General,* **110**, 306-340.

Jacoby, L. L., & Goolkasian, P. (1973). Semantic versus acoustic coding: Retention and conditions of organization. *Journal of Verbal Learning and Verbal Behavior,* **12**, 324-333.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance,* **1**, 130-136.

Jenkins, J. J. (1974). Can we have a theory of meaningful memory? In R. L. Solso (Ed.), *Theories in Cognitive Psychology*. Potomac, MD.: Erlbaum.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Jenkins, J. J., Mink, W. D., & Russell, W. A. (1958). Associative clustering as a function of verbal association strength. *Psychological Reports,* **4**, 127-136.

Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology,* **47**, 818-821.

Jenkins, J. R., Stein, N. L., Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal,* **21**, 767-787.

Jensen, A. R. (1962). Temporal and spatial effects of serial position, *American Journal of Psychology,* **75**, 390-400.

Johnston, W. A. (1967). S-R, R-S independence and the interference potency of latent R-S associations. *Journal of Experimental Psychology,* **74**, 511-516.

Jones, G. V. (1976). A fragmentation hypothesis of memory: Cued recall of pictures and of sequential position. *Journal of Experimental Psychology: General,* **105**, 277-293.

Jones, G. V. (1985). Deep dyslexia, imageability, and ease of predication. *Brain and Language,* **24**, 1-19.

Jones, G. V., & Martin, M. (1980). Recall cued by selectively attended and unattended attributes. *Memory and Cognition,* **8**, 94-98.

Jorm, A. F. (1979). The nature of the reading deficit in developmental dyslexia: A reply to Ellis. *Cognition*, **7**, 421-433.

Judd, E. L. (1978). Vocabulary teaching and TESOL: A need for reevaluation (*sic*) of existing assumptions. *TESOL Quarterly*, **12**, 71-76.

Jung, J., & Skeebo, S. (1967). Multitrial free recall as a function of constant *versus* varied input order and list length. *Canadian Journal of Psychology,* **21**, *329-336.*

Katz, J. J., & Postal, P. M. (1964). *An integrated theory of linguistic descriptions.* Cambridge, MA.: MIT.

Kay, J. (1985). Mechanisms of oral reading: A critical appraisal of cognitive models. In A.W. Ellis (Ed.), *Progress in the psychology of language,* Vol. 2. London: Erlbaum.

Kay, J., & Bishop, D. (1987). Anatomical differences between nose, palm, and foot, or, the body in question: Further dissection of the processes of sub-lexical spelling-sound translation. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale: Erlbaum.

Kellerman, E. (1977). Towards a characterisation of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin,* **2**, 58-145.

Kellerman, E. (1983). Now you see it, now you don't. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning*. Rowley, MA.: Newbury House.

Kempley, S. T., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology,* **73**, 441-454.

Kibby, M. W. (1977). Note on relationship of word difficulty and word frequency. *Psychological Reports,* **41**, 12-14.

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin,* **85**, 532-554.

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ.: Erlbaum.

Kirkpatrick, J. J., & Cureton, E. E. (1949). Vocabulary item difficulty and word frequency. *Journal of Applied Psychology,* **33**, 347-351.

Kirsner, K., Brown, H. L., Abrol, S., Chadha, N. K., & Sharma, N. N. (1980). Bilingualism and lexical representation. *Quarterly Journal of Experimental Psychology,* **32**, 585-594.

Kirsner, K. & Dunn, J. C. (1985). The perceptual record: A common factor in repetition priming and attribute retention. In M. I. Posner, and O. S. M. Marin (Eds.), *Mechanisms of attention: Attention and Performance XI.* Hillsdale, NJ.: Erlbaum.

Kirsner, K., Dunn, J., & Standen, P. (1987). Record-based word recognition. In Coltheart, M. (Ed.), *Attention and Performance XII: The psychology of reading.* Hillsdale, NJ.: Erlbaum.

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M-L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior,* **23**, 519-539.

Klee, H., & Legge, D. (1976). Estimates of concreteness and other indices for 200 transitive verbs. *Journal of Experimental Psychology: Human Learning and Memory,* **2**, 497-507.

Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory and Cognition,* **8**, 336-344.

Klein, W. (1986). *Second Language Acquisition.* Cambridge: Cambridge University Press.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior,* **2**, 291-300.

Kolers, P. A. (1966). Interlingual facilitation of short-term memory. *Journal of Verbal Learning and Verbal Behavior,* **5**, 314-319.

Kolers, P. A. (1968). Bilingualism and information processing. *Scientific American,* **218**, 78-86.

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

Kolers, P. A. (1979). A pattern-analyzing basis of recognition. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Kolers, P. A. (1985). Skill in reading and memory. *Canadian Journal of Psychology,* **39**, 232-239.

Kolers, P. A., & Roediger, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior,* **23**, 425-449.

Kopstein, F. F., and Roshal, S. M. (1955). Method of presenting word pairs as a factor in foreign vocabulary learning. *American Psychologist,* **10**, 354.

Krashen, S. D. (1977). The monitor model of adult second language performance. In M. Burt, H. Dulay, & M. Finocciaro (Eds.), *Viewpoints on English as a Second Language.* New York: Regents.

Krashen, S. D. (1978). The monitor model for second language acquisition. In R. C. Gingras (Ed.), 1978, *Second-language acquisition and foreign language teaching.* Arlington, VA.: Center for Applied Linguistics.

Krashen, S. D. (1981). *Second-language acquisition and second language learning.* Oxford: Pergamon.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* London: Longman.

Kremin, H. (1985). Routes and strategies in surface dyslexia and dysgraphia. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI.: Brown University Press.

Kumar, R., Stolerman, I. P., & Steinberg, H. (1970). Psychopharmacology. *Annual Review of Psychology,* **21**, 595-628.

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology,* **6**, 293-323.

Lachman, R., & Laughery, K. R. (1968). Is a test trial a training trial in free recall learning? *Journal of Experimental Psychology,* **76**, 40-50.

Lado, R. (1957). *Linguistics across cultures.* Ann Arbor: University of Michigan Press.

Lado, R., Baldwin, B., & Lobo, F. (1967). *Massive vocabulary expansion in foreign language beyond the basic course: The effects of stimuli, timing, and order of presentation.* United States Office of Education. Bureau of Research. Project 5, 1095.

Lambert, W., & Fillenbaum, S. (1959). A pilot study of aphasia among bilinguals. *Canadian Journal of Psychology,* **13**, 28-34.

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior,* **12**, 119-131.

Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research.* London: Longman.

Laxon, V. J., Coltheart, V., & Keating, C. (1988). Children find friendly words friendly too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology,* **58**, 103-119.

Lesgold, A. M. & Bower, G. H. (1970). Inefficiency of serial knowledge for associative responding. *Journal of Verbal Learning and Verbal Behavior,* **9**, 456-466.

Levenston, E. A. (1979). Second language vocabulary acquisition: Issues and problems. *Interlanguage Studies Bulletin*, **4**, 147-160.

Levy, B. A. (1981). Interactive processes during reading. In A. M. Lesgold and C. A. Perfetti (Eds.), *Interactive processes in reading*. Hillsdale, NJ.: Erlbaum.

Lima, S.D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The Basic Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior*, **22**, 310-332.

Lockhart, R. (1969). Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, **79**, 12-17.

Lockhart, R. S., & Craik, F. I. M. (1978). Levels of processing. A reply to Eysenck. *British Journal of Psychology*, **69**, 171-175.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. *Research on Speech Perception, technical report no. 6*. Bloomington Ind.: Department of Psychology, Speech Research Laboratory. [Cited in J. Charles-Luce, P. A. Luce, M. S. Cluff, Retroactive influence of syllable neighborhoods. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA.: MIT Press].

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA.: MIT Press.

Luszcz, M. A., Bungey, J., & Geffen, G. (1984). Orthographic-morphemic factors in reading: A developmental study. *Australian Journal of Psychology*, **36**, 355-365.

Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior*, **6**, 729-736.

Macnamara, J. (1970). Bilingualism and thought. In J. Alatis (Ed.), *Bilingualism and language contact*. Washington, DC.: Georgetown University Press.

Macnamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, **10**, 480-487.

Madigan, S. A. (1969). Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior,* **8**, 828-835.

Mägiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior,* **18,** 79-89.

Maki, R. H., & Schuler, J. (1980). Effects of rehearsal duration and level of processing on memory for words. *Journal of Verbal Learning and Verbal Behavior,* **19**, 36-45.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory.* Volume 1. New York: Academic Press.

Mandler, G. (1968). Association and organization: Facts, fancies, and theories. In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, NJ.: Prentice-Hall.

Mandler, G. (1977). Organization and memory. In G. H. Bower, *Human memory: basic processes*. New York: Academic.

Mandler, G., & Dean, P. J. (1969). Seriation: Development of serial order in free recall. *Journal of Experimental Psychology,* **81**, 207-215.

Mandler, G., Worden, P. E, & Graesser, A. C. (1974). Subjective disorganization: Search for the locus of list organization. *Journal of Verbal Learning and Verbal Behavior,* **13,** 220-235.

Marcel, A. J. (1976). Unconscious reading. Paper to the British Association for the Advancement of Science. [Cited in B. A. Levy (1981), Interactive processes during reading. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading*. Hillsdale, NJ.: Erlbaum].

Marcel, A. J., & Patterson, K. E. (1978). Word recognition and production: Reciprocity in clinical and normal studies. In J. Requin (Ed.), *Attention and Performance VII* Hillsdale, NJ.: Erlbaum.

Marcel, T. (1987). Surface dyslexia and beginning reading: a revised hypothesis of the pronunciation of print and its impairments. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Margolin, D. I., Marcel, A. J., & Carlson, N. R. (1985). Common mechanisms in dysnomia and post-semantic surface dyslexia: Processing deficits and selective attention. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia. Neuropsychological and Cognitive Studies of Phonological Reading.* London: Erlbaum.

Marshall, J. C., & Newcombe, F. (1966). Syntactic and semantic errors in paralexia. *Neuropsychologia,* **4**, 169-176.

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research,* **2**, 175-199.

Marshall, J. C., & Newcombe, F. (1987). The conceptual staus of deep dyslexia: An historical perspective. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Marshall, J. C., Newcombe, F., & Holmes, J. M. (1975). Lexical memory: A linguistic approach. In R. A. Kennedy & A. L. Wilkes (Eds.), *Studies in Long Term Memory.* New York: Academic.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition. In H. Bouma, & D. Bouwhuis (Eds.), *Attention and Performance X.* Hillsdale, NJ.: Erlbaum.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition,* **25**, 71-102.

Marslen-Wilson, W. D. (1989). Access and integration: Projecting sound onto meaning. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: MIT Press.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA.: MIT Press.

Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.

Maskarinec, A. S., and Brown, S. C. (1974). Positive and negative recency effects in free recall learning. *Journal of Verbal Learning and Verbal Behavior,* **13**, 328-334.

Massaro, D. W., Jones, R. D., Lipscomb, D., & Scholz, R. (1978). Role of prior knowledge on naming and lexical decision with good and poor stimulus information. *Journal of Experimental Psychology: Human Learning and Memory,* **4**, 498-512.

Masterson, J. (1985a). Appendix III (p. 70) to E. M. Saffran, Lexicalisation and reading performance in surface dyslexia. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

Masterson, J. (1985b). On how we read non-words: Data from different populations. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance,* **13**, 14-24.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance,* **14**, 693-706.

McCarthy, M. J. (1984). A new look at vocabulary in EFL. *Applied Linguistics*, **5**, 12-22.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review,* **86**, 287-330.

McClelland, J. L. (1985). Putting knowledge in its place: A scheme for programming parallel processing structures on the fly. *Cognitive Science,* **9**, 113-146.

McClelland, J. L. (1986). The programmable blackboard model of reading. In J. L. McClelland, D. E. Rumelhart, and the PDP research group (Eds.), *Parallel distributed processing. Explorations in the microstructure of cognition: Volume II: Psychological and Biological Models.* Cambridge, MA.: MIT Press.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* Hillsdale, NJ.: Erlbaum.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology,* **18**, 1-86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Pyschological Review,* **88**, 375-407.

McClelland, J. L., Rumelhart, D. E., & the PDP research Group. (1986). *Parallel distributed processing. Explorations in the microstructure of cognition: Volume II: Psychological and Biological Models.* Cambridge, MA.: MIT Press.

McCormack, P. D. (1977). Bilingual linguistic memory: The independence / interdependence issue revisited. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social, and educational implications.* New York: Academic.

McDaniel, M. A., & Masson, M. E. (1977). Long term retention: When incidental semantic processing fails. *Journal of Experimental Psychology: Human Learning and Memory,* **3**, 270-281.

McDonough, S. H. (1986). *Psychology in foreign language teaching.* Second Edition. London: Allen & Unwin.

McKeown, M. C., & Curtis, M. E. (Eds.) (1987). *The Nature of Vocabulary Acquisition.* Hillsdale, NJ.: Erlbaum.

McLaughlin, B. (1987). *Theories of Second-Language Learning*. London: Arnold.

McRae, K., Jared, D., & Seidenberg, M. S. (1990). Roles of frequency and lexical access in word naming. *Journal of Memory and Language*, **29**, 43-65.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts*, **13**, 221-246.

Meara, P. (1982). Word associations in a foreign language. *Nottingham Linguistics Circular*, **11**, 29-38.

Meara, P. (1983). *Vocabulary in a second language*. London: CILT.

Meara, P. (1985). Some priorities for research in second language learning. In C. Brumfit, H. Lunt, and J. Trim (Eds.), *Second language learning: Research problems and perspectives*. London: CILT.

Meara, P. (1987). *Vocabulary in a second language*. Volume 2. London: CILT.

Merry, R. (1980). The keyword method and children's vocabulary learning in the classroom. *British Journal of Educational Psychology*, **50**, 123-136.

Meyer, D. E., & Ruddy, M. G. (1974). Bilingual word recognition: Organization and retrieval of alternative lexical codes. Paper presented to April meeting of the Easter Psychological Association, Philadelphia. [Cited in A. M. B. De Groot & G. L. J. Nas, 1991, Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, **30**, 90-123.]

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, **90**, 227-234.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition*, **2**, 309-321.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects in visual word recognition. In P. M. A. Rabbit and S. Dornic (Eds.), *Attention and Performance V*. New York: Academic Press.

Michéa, R. (1964). Basic vocabularies. *New Research and Techniques for the Benefit of Modern Language Teaching*. Strasbourg: Council for Cultural Cooperation.

Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.

Mishima, T. (1966). *An experiment comparing five modalities of conveying meaning for the teaching of foreign language vocabulary.* Unpublished Ph. D. thesis. Georgetown University.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, **30**, 609-636.

Moeser, S. D. (1974). Memory for meaning and wording in concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior*, **13**, 682-697.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the Psychology of Language*, Volume 2. London: Erlbaum.

Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language Perception and Production*. London: Academic press.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in Reading: Visual Word Recognition*. London: Erlbaum.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition: Where are they? *Journal of Experimental Psychology: General*, **118**, 43-71.

Morgan, C. L., & Bailey, W. L. (1943). The effect of context on learning a vocabulary. *Journal of Educational Psychology*, **34**, 561-565.

Morgan, C. L., & Foltz, M. C. (1944). The effect of context on learning a French vocabulary. *Journal of Educational Research*, **38**, 213-216.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior,* **16**, 519-533.

Morton, J. (1964). The effects of context on the visual duration threshold for words. *British Journal of Psychology,* **55**, 165-180.

Morton, J. (1969). 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. (1987). Two auditory parallels to deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Morton, J., & Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior,* **15**, 43-51.

Morton, J., & Patterson, K. (1980). 'Little words - No!' In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (First Edition). London: Routledge & Kegan Paul.

Morton, J., & Patterson, K. (1987). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge &Kegan Paul.

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior,* **15**, 447-458.

Murdock, B. B. (1962). The serial position effect in free recall. *Journal of Experimental Psychology* , **64**, 482-488.

Murdock, B. B. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General,* **105**, 191-216.

Nagy, W. E, & Herman, P. A. (1987). Breadth and depth of vocabulary acquisition: Implications for acquisition and instruction. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. Hillsdale, NJ.: Erlbaum.

Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, **13**, 14-36.

Nation, I. S. P. (1983). *Teaching and Learning Vocabulary*. Wellington: Victoria University.

Nation, I. S. P. (1987). *Teaching and Learning Vocabulary*. (Revised Edition). Occasional Publication No. 7. English Language Institute. Wellington: Victoria University.

Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.

Nation, P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching*. London: Longman.

Nattinger, J. (1988). Some current trends in vocabulary teaching. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching*. London: Longman.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226-254.

Nelson, D. L. (1979). Remembering words and pictures: Appearance, significance, and name. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior*, **16**, 151-171.

Nilsen, D. L. F. (1976). Contrastive semantics in vocabulary instruction. *TESOL Quarterly*, **10**, 99-103.

Noble, C. E. (1952). An analysis of meaning. *Psychological Review*, **59**, 421-430.

Noble, C. E. (1961). Measurements of association value (*a*), rated associations (*a'*), and scaled meaningfulness (*m'*) for the 2100 CVC combinations of the English alphabet. *Psychological Reports*, **8**, 487-521.

Noble, C. E., & McNeely, D. A. (1957). The role of meaningfulness (*m*) in paired-associate verbal learning. *Journal of Experimental Psychology*, **53**, 16-22.

Norris, D. (1984). The effects of frequency, repetition and stimulus quality in visual word recognition. *Quarterly Journal of Experimental Psychology*, **36A**, 507-518.

Novinski, L. S. (1969). Part-whole and whole-part free recall learning. *Journal of Verbal Learning and Verbal Behavior*, **8**, 152-154.

Obler, L. K. (1984). The neuropsychology of bilingualism. In D. Caplan, A. R. Lecours, and A. Smith (Eds.), *Biological perspectives on language*. London: MIT Press.

Obler, L. K. & Albert, M. L. (1978). A monitor system for bilingual processing. In M. Paradis (Ed.), *Aspects of bilingualism*. Columbia, SC.: Hornbeam Press.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review*, **77**, 257-273.

Oskarsson, M. (1975). On the role of the mother tongue in learning foreign language vocabulary: An empirical investigation. *I.T.L.*, **27**, 19-32.

Ostyn, P., & and Godin, P. (1985). RALEX: An alternative approach to language teaching. *Modern Language Journal*, **69**, 346-354.

Paivio, A. (1969) Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241-263.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rhinehart, Winston.

Paivio, A. (1978). The relationship between verbal and perceptual codes. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception*. New York: Aceadmic Press.

Paivio, A. (1983). The empirical case for dual coding. In J. C. Yuille (Ed.), *Imagery, memory, and cognition*. Hillsdale, NJ.: Erlbaum.

Paivio, A. (1986). *Mental representations: A dual-coding approach.* New York: Oxford Unviersity Press.

Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology,* **80**, 279-285.

Paivio, A., & Desrochers, A. (1979). Effects of an imagery mnemonic on second language recall and comprehension. *Canadian Journal of Psychology,* **33**, 17-28.

Paivio, A., & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology,* **73**, 780-795.

Paivio, A., & Foth, D. (1970). Imaginal and verbal mediators and noun concreteness in paired-associate learning: The elusive interaction. *Journal of Verbal Learning and Verbal Behavior,* **9**, 384-390.

Paivio, A., & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior,* **20**, 532-539.

Paivio, A., & Rowe, E. J. (1970). Noun imagery, frequency, and meaningfulness in verbal discrimination. *Journal of Experimental Psychology,* **85**, 264-269.

Paivio, A., & Yuille, J. C. (1969). Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. *Journal of Experimental Psychology,* **79**, 458-463.

Paivio, A., Yuille, J., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement,* **76**, 1-25.

Paivio, A., Yuille, J. C., & Smythe, P. C. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology,* **20**, 362-377.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language,* **30**, 331-347.

Paradis, M. (1977). Bilingualism and aphasia. In H. Whitaker & H. Whitaker (Eds.), *Studies in neurolinguistics.* Volume 3. New York: Academic Press.

Patterson, K. (1979). What is right with 'deep' dyslexic patients? *Brain and Language*, **8**, 111-129.

Patterson, K. (1981). Neuropsychological approaches to the study of reading. *British Journal of Psychology*, **72**,151-174.

Patterson, K. (1987). Derivational errors. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia*. (Second Edition). London: Routledge & Kegan Paul.

Patterson, K., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale: Erlbaum.

Patterson, K., & Marcel, A. J. (1977). Aphasia, dyslexia, and the phonological coding of written words. *Quarterly Journal of Experimental Psychology*, **29**, 307-318.

Patterson, K., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

Perfetti, C. A. (1979). Level of language and levels of process. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ.: Erlbaum.

Perfetti, C. A., Bell, L. C., Hogoboam, T. W., & Goldman, S. R. (1977). *Verbal processing speed and reading skill*. Paper presented at the meeting of the Psychonomic Society, November, 1977. [Cited in K. T. Spoehr & R. E. Schuberth, 1981, Processing words in context. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.]

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In P. Carpenter & M. Just (Eds.), *Cognitive Processes in Comprehension*. Hillsdale, NJ.: Erlbaum.

Perfetti, C. A., & Lesgold, A. M. (1979). Coding and comprehension in skilled reading and implications for reading instruction. In L. B. Resnick & P. Weaver (Eds.), *Theory and Practice of Early Reading*. Volume 1. Hillsdale, NJ.: Erlbaum.

Perfetti, C. A., & Roth, S. (1981). Some of the interactive processs in reading and their role in reading skill. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading*. Hillsdale, NJ.: Erlbaum.

Peterson, L.R., & Peterson, M. (1959). Short-term retention of individual items. *Journal of Experimental Psychology*, **58**, 193-198.

Pezdek, K., & Royer, J. M. (1974). The role of comprehension in learning concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior,* **13**, 551-558.

Pickering, M. (1982). Context-free and context-dependent vocabulary learning: An experiment. *System,* **10**, 79-83.

Pimsleur, P. (1967). A memory schedule. *Modern Language Journal*, **51**, 73-75.

Posner, M.I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), *Information processing and cognition.* Hillsdale, NJ.: Erlbaum.

Postman, L. (1962). The effects of language habits on the acquisition and retention of verbal associations. *Journal of Experimental Psychology,* **64**, 7-19.

Postman, L. (1971). Organization and interference. *Psychological Review,* **78**, 290-302.

Postman, L. (1972). A pragmatic view of organization theory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory.* New York: Academic Press.

Postman, L., & Knecht, K. (1983). Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior,* **22**, 133-152.

Postman, L., and Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology,* **17**, 132-138.

Postman, L., & Schwartz, M. (1964). Studies of learning to learn: I. Transfer as a function of method of practice and class of verbal materials. *Journal of Verbal Learning and Verbal Behavior*, **3**, 37-49.

Postman, L., Burns, S., & Hasher, L. (1970). Studies of learning to learn: X. Nonspecific transfer effects in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, **9**, 707-715.

Potter, M. C., So, K-F., Von Eckhardt, B., and Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior,* **23**, 23-38.

Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research,* **52**, 61-92.

Pressley, M., Levin, J. R., & McDaniel, M.A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition.* Hillsdale, NJ.: Erlbaum.

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology: Human Learning and Memory,* **5**, 22-29.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior,* **8**, 295-301.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin,* **80**, 1-24.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review,* **88**, 16-45.

Quinlan, P. T. (1992). *Oxford psycholinguistic database.* Oxford: Oxford University Press.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language*, **26**, 392-405.

Ratcliff, R. A., & McKoon, G. (1981). Does activation really spread? *Psychological Review,* **88**, 454-462.

Raugh, M. R. & Atkinson, R. C. (1975). A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, **67**, 1-16.

Raymond, B. (1969). Short-term storage and long-term storage in free recall. *Journal of Verbal Learning and Verbal Behavior,* **8**, 567-574.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition,* **14**, 191-201.

Reder, L. M., Anderson, J. R., & Bjork, R. A. (1974). A semantic interpretation of encoding specificity. *Journal of Experimental Psychology,* **102**, 648-656.

Richards, J. C. (1970). A psycholinguistic measure of vocabulary selection. *IRAL*, **8**, 87-102.

Richards, J. C. (1974). Word lists: Problems and prospects. *RELC Journal*, **5**. 69-84.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, **10**, 77-89.

Richards, L. G. (1976). Concreteness as a variable in word recognition. *American Journal of Psychology,* **89**, 707-718.

Richardson, J. T. E. (1975a). Imagery, concreteness, and lexical complexity. *Quarterly Journal of Experimental Psychology,* **27**, 211-223.

Richardson, J. T. E. (1975b). The effect of word imageability in acquired dyslexia. *Neuropsychologia,* **13**, 281-288.

Rivers, W. M. (1981). *Teaching foreign-language skills.* Second Edition. Chicago: University of Chicago Press.

Roberts, W. A. (1969). The priority of recall of new items in transfer from part-list learning to whole-list learning. *Journal of Verbal Learning and Verbal Behavior,* **8**, 645-652.

Rodgers. T. S. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *IRAL*, **7**, 327-343.

Romaine, S. (1989). *Bilingualism.* Oxford: Blackwell.

Rosson, M. B. (1983). From SOFA to LOUCH: Lexical contributions to pseudoword pronunciation. *Memory and Cognition*, **11**, 152-160.

Rosson, M. B. (1985). The interaction of pronunciation rules and lexical representations in reading aloud. *Memory and Cognition,* **13**, 90-99.

Rothkopf, E. Z., & Coke, E. U. (1961). The prediction of free recall from word association measures. *Journal of Experimental Psychology,* **62**, 433-438.

Rowe, E. J., & Paivio, A. (1971). Discrimination learning of words and pictures. *Psychonomic Science,* **22**. 87-88.

Rowe, E. J., & Paivio, A. (1972). Effects of noun imagery, pronunciation, method of presentation, and intrapair order of items in verbal discrimination. *Journal of Experimental Psychology,* **93**, 427-429.

Rubenstein, H., & Aborn, J. (1958). Learning, prediction, and readability. *Journal of Applied Psychology,* **42**, 28-32.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior,* **9**, 487-494.

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior,* **19**, 736-755.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Pyschological Review,* **89**, 60-94.

Rundus, D. (1977). Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior,* **16**, 665-681.

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior,* **9**, 99-105.

Saffran, E. M. (1985). Lexicalisation and reading performance in surface dyslexia. In K. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia: Neuropsychological and cognitive studies of phonological reading.* London: Erlbaum.

Saffran, E. M., Bogyo, L. C., Schwartz, M. F., & Marin, O. S. M. (1980). Does deep dyslexia reflect right hemisphere reading? In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul.

Saffran, E. M., Bogyo, L. C., Schwartz, M. F., & Marin, O. S. M. (1987). Does deep dyslexia reflect right hemisphere reading? In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Samuel, A. G. (1990). Using perceptual-restoration effects to explore the architecture of perception. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, MA.: MIT Press.

Samuels, S. J., Begy, G., & Chen, C. C. (1975-76). Comparison of word recognition speed and strategies of less skilled and more highly skilled readers. *Reading Research Quarterly,* **11**, 72-88.

Sanford, A. J., Garrod, S., & Boyle, J. M. (1977). An independence of mechanism in the origins of reading and classification-related semantic distance effects. *Memory and Cognition,* **5**, 214-220.

Saville, M., & Troike, R. (1982). *A handbook of bilingual education.* Washington, DC.: TESOL.

Savin, H. (1963). Word frequency effects and errors in the perception of speech. *Journal of the Acoustical Society of America,* **35**, 200-206.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance,* **3**, 1-17.

Scarborough, D. L., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. *Memory and Cognition,* **7**, 3-12.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior,* **23**, 84-99.

Scheerer, E. (1987). Visual word recognition in German. In A. Allport, D. G. Mackay, W. Prinz, & E. Scheerer (Eds.), *Language Perception and Production*. London: Academic press.

Schneider, W., & Fisk, A. D. (1982). Concurrent automatic and controlled visual search: Can processing occur without resource cost? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **8**, 261-278.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, **84**, 1-66.

Schnorr, J. A., & Atkinson, R. C. (1970). Study position and item difficulty in the short- and long-term retention of paired associates learned by imagery. *Journal of Verbal Learning and Verbal Behavior*, **9**, 614-622.

Schuberth, R. E., & Eimas, P. D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 27-36.

Schvaneveldt, R., Ackerman, B. P. & Semlear, T. (1977). The effect of semantic context on children's word recognition. *Child Development*, **48**, 612-616.

Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language*, **25**, 605-618.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **9**, 82-102.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1977). An analysis of agrammatic reading in aphasia. Presented at International Neuropsychological Society meeting, Santa Fe. [Cited in M. Coltheart, 1987, Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia*. (Second Edition). London: Routledge & Kegan Paul.]

Segal, M. A. and Mandler, G. (1967). Directionality and organizational processes in paired-associate learning. *Journal of Experimental Psychology*, **74**, 305-312.

Seibert, L. C. (1930). An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. *Journal of Educational Psychology,* **21**, 297-314.

Seidenberg, M. S. (1985a). The time course of phonological code activation in two writing systems. *Cognition,* **19**, 1-30.

Seidenberg, M. S. (1985b). The time course of information activation and utilisation in visual word recognition. In D. Besner, T. G. Waller, & G. E. McKinnon (Eds.), *Reading research.* Vol. 5. Orlando: Academic Press.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy? In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* Hillsdale, NJ.: Erlbaum.

Seidenberg, M. S. (1989). Visual word recognition and pronunciation: A computational model and its implications. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: MIT Press.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review,* **96**, 523-568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior,* **23**, 383-404.

Sejnowski, T. J., & Rosenberg, C. R. (1986). *NETtalk; A parallel network that learns to read aloud.* Report # 13 of the Cognitive Neuropsychology Laboratory, The Johns Hopkins University. [Reprinted in J. A. Anderson & E. Rosenfield (Eds.), (1988), *Neurocomputing.* Cambridge, MA.: MIT Press.]

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics,* **10**, 209-231.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology,* **45A**, 21-50.

Shallice, T. (1978). Cited as a personal communication in A. J. Marcel, & K. Patterson, 1978, (p. 218), Word recognition and production: Reciprocity in clinical and normal studies. In J. Requin (Ed.), *Attention and Performance VII.* Hillsdale, NJ.: Erlbaum.

Shallice, T., & McCarthy, R. (1985). Phonological reading: From patterns of impairments to possible procedures. In K. Patterson, J. C. Marshall, and M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. London: Erlbaum.

Shallice, T., & Warrington, E. K. (1975). Word recognition in a phonemic dyslexic patient. *Quarterly Journal of Experimental Psychology*, **27**, 187-199.

Shallice, T., & Warrington, E. K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul.

Shallice, T., & Warrington, E. K. (1987). Single and multiple component central dyslexic syndromes. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* (Second Edition). London: Routledge & Kegan Paul.

Shannon, C.E. (1951). Prediction and entropy in printed English. *Bell System Technical Journal,* 30, 50-64. [Cited in P. B. Gough, J. A. Alford, & P. Holley-Wilcox (1981), Words and contexts. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology.* Hillsdale, NJ.: Erlbaum.]

Shapiro, B. J. (1969). The subjective estimation of relative word frequency. *Journal of Verbal Learning and Verbal Behavior,* **8**, 248-251.

Shiffrin, R. (1970). Memory search. In D. Norman (Ed.), *Models of human memory.* New York: Academic Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review,* **84**, 127-190.

Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA.: MIT Press.

Simons, W. H. G., Vonk, W., & Noordman, L. G. M. (1989). *The representation of expert and novice knowledge structures in economics.* [Unpublished manuscript cited in A. M. B. De Groot (1989), Representational aspects of word imageability and word frequency as assessed through word association. *Journal Of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 824-845.]

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology,* **76**, 504-513.

Slamecka, N. J., Moore, T., & Carey, S. (1972). Part-to-whole transfer and its relation to organization theory. *Journal of Verbal Learning and Verbal Behavior,* **11**, 73-82.

Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition,* **14**, 230-237.

Smith, F. (1971). *Understanding reading.* New York: Holt, Rinehart, and Winston.

Smith, P. (1969). Teaching vocabulary. *Modern Language Journal,* **53**, 531-537.

Smith, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory and Cognition,* **10**, 405-412.

Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior,* **23**, 3-22.

Solso, R. L. (1974). Memory and the efficacy of cues. In R. L. Solso (Ed.), *Theories in Cognitive Psychology*. Potomac, MD.: Erlbaum.

Spoehr, K. T. & Schuberth, R. E. (1981). Processing words in context. In O. J. L. Tzeng, & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.

Spoehr, K. T., & Smith, E. E. (1973). The role of syllables in perceptual processing. *Cognitive Psychology,* **5**, 71-89.

Spyropoulos, T., & Ceraso, J. (1977). Categorized and uncategorized attributes as recall cues: The phenomenon of limited access. *Cognitive Psychology,* **9**, 384-402.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior,* **14**, 259-264.

Stanovich, K. E. (1981). Attentional and automatic context effects in reading. In M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ.: Erlbaum.

Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory and Cognition,* **7**, 77-85.

Stern, H. H. (1983). *Fundamental Concepts of Language Teaching.* Oxford: Oxford University Press.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. C. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition.* Hillsdale, NJ.: Erlbaum.

Sternberg, R. J. (1988). *The Psychologist's Companion.* Cambridge: Cambridge University Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. In W. G. Koster (Ed.), *Attention and Performance II.* Amsterdam: New Holland.

Stoddard, G. D. (1929). An experiment in verbal learning. *Journal of Educational Psychology,* **20**, 452-457.

Strick, G. J. (1980). A hypothesis for semantic development in a second language. *Language Learning,* **30**, 155-176.

Summers, D. (1988). The role of dictionaries in language learning. In R. Carter & M. McCarthy (Eds.), *Vocabulary and Language Teaching.* London: Longman.

Taft, M. (1979a). Lexical access via an orthographic code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior,* **18**, 21-39.

Taft, M. (1979b). Recognition of affixed words and the word frequency effect. *Memory and Cognition,* **7**, 263-272.

Taft, M. (1986). Lexical access codes in visual and auditory word recognition. *Language and Cognitive Processes,* **1**, 49-60.

Taft, M. (1987). Morphographic processing: The BOSS re-emerges. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* Hillsdale, NJ.: Erlbaum.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior,* **15**, 607-620.

Taft, M., & Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition,* **22**, 259-282.

Taylor, C. V. (1983). Vocabulary for education in English. *World Language English,* **2**, 100-104.

Theios, J., & Muise, J. G. (1977). The word identification process in reading. In N. J. Castellan, D. B. Pisoni, and G. R. Potts (Eds.), *Cognitive Theory.* Volume 2. Hillsdale, NJ.: Erlbaum.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology,* **86**, 255-262.

Thorndike, E., & Lorge, I. (1944). *The teacher's word book of 30,000 words.* New York: Columbia University, Teachers College Press.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review,* **76**, 282-299.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review,* **69**, 344-354.

Tulving, E. (1966). Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior,* **5**, 193-197.

Tulving, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory.* Englewood Cliffs, NJ.: Prentice-Hall.

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

Tulving, E. (1974). Recall and recognition of semantically encoded words. *Journal of Experimental Psychology,* **102**, 778-787.

Tulving, E. (1979). Relation between encoding specificity and levels of processing. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ.: Erlbaum.

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology,* **1**, 86-98.

Tulving, E., & Donaldson, W. (Eds.) (1972). *Organization of Memory.* New York: Academic Press.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology,* **77**, 593-601.

Tulving, E., & Patkau, J. E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology,* **16**, 83-95.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior,* **5**, 381-391.

Tulving, E., & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology,* **87**, 1-8.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, **80**, 352-373.

Turner, G. (1983). Teaching French vocabulary: A training study. *Educational Review*, **35**, 81-88.

Tzeng, O. J. L. (1973). Positive recency effect in a delayed free recall. *Journal of Verbal Learning and Verbal Behavior*, **12**, 436-439.

Tzeng, O. J. L. & Singer, H. (Eds.). (1981). *Perception of print: Reading research in experimental psychology*. Hillsdale, NJ.: Erlbaum.

Underwood, B. J. (1966). *Experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts.

Underwood, B. J. (1982). Paired associate learning: Data on pair difficulty and variables that influence difficulty. *Memory and Cognition*, **10**, 610-617.

Underwood, B. J., & Schulz, R. W. (1960). *Meaningfulness and Verbal Learning*. Chicago: Lippincott.

Vaid, J. (1988). Bilingual memory representation: A further test of dual coding theory. *Canadian Journal of Psychology*, **42**, 84-90.

Voss, J. F. (1972). On the relationship of associative and organizational processes. In E. Tulving, & W. Donaldson (Eds.) *Organization of Memory*. New York: Academic.

Voss, J. F. (1979). Organization, structure, and memory: Three perspectives. In C.R. Puff (Ed.), *Memory organization and structure*. New York: Academic.

Wallace, M. J. (1982). *Teaching Vocabulary*. London: Heinemann.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, **27**, 635-657.

Warrington, E. K. (1981). Concrete word dyslexia. *British Journal of Psychology*, **72**, 175-196.

Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain*, **102**, 43-63.

Warrington, E. K., & Shallice, T. (1980). Word form dyslexia. *Brain*, **103**, 99-112.

Weber, R. (1970). First graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books.

Weinrich, U. (1953). *Languages in contact: Findings and problems.* New York: Linguistic Circle of New York.

Weist, R. M. (1970). Optimal versus nonoptimal conditions for retrieval. *Journal of Verbal Learning and Verbal Behavior,* **9**, 311-316.

Weist, R. M. (1972). The role of rehearsal: Recopy or reconstruct. *Journal of Verbal Learning and Verbal Behavior,* **11**, 440-450.

West, M. P. (1953). *A general service list of English words.* London: Longmans.

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development,* **49**, 717-727.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **8**, 385-399.

Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior,* **17**, 143-154.

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review,* **77**, 1-15.

Wind, M., & Davidson, M. (1969). Facilitation of paired-associate learning by language context. *Psychonomic Science,* **15**, 184-185.

Winograd, T., Cohen, C., & Barresi, J. (1976). Memory for concrete and abstract words in bilingual speakers. *Memory and Cognition,* **4**, 323-329.

Wolford, G. (1971). Function of distinct associations for paired-associate performance. *Psychological Review,* **78**, 303-313.

Wood, G. (1969). Whole-part transfer from paired-associate to free recall learning. *Journal of Experimental Psychology,* **82**, 532-537.

Wood, G. (1972). Organizational processes and free recall. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory.* New York: Academic Press.

Wood, G., & Clark, D. (1969). Instructions, ordering and previous practice in free-recall learning. *Psychonomic Science,* **14**, 187-188.

Yuille, J. C. (1968). Concreteness without imagery in PA learning. *Psychonomic Science,* **11**, 55-56.

MATERIALS APPENDIX

The learning and test lists were designed to fit an A4 page; exemplars of each lay-out will be given as they are introduced. The materials themselves will be given in summary form.

INDEX

NOTE

Despite the best endeavours of those who kindly helped check the materials for accuracy, some inaccuracies and infelicities were contained in the experimental materials used. These do not make any material difference to the experimental results, but for the sake of accuracy preferred versions are given in the appendix which follows. The incorrect version is marked with an asterisk, and the preferred version is given in brackets, as for example:

Elle vient .... *
[Elle va venir ....]

INSTRUCTIONS TO TEACHERS

Instructions to teachers took the following form and were, *mutatis mutandis*, similar for all experiments:

TO THE TEACHER - DAY ONE

Thank you very much for helping me conduct this experiment. I would be grateful if you could follow the procedure below as closely as possible - so that the various groups doing the experiment are all given the same task.

1. Explain that this is a special test connected with an experiment which is designed, eventually, to make vocabulary learning easier. Ask the pupils, therefore, to do it on their own and as carefully as possible since this is more important than just getting the answers to the test correct.

2. Explain.
• that they will be given some words to learn (please avoid the word 'list' since the test will have the words in a different order from the original).
• that they will have eight minutes to learn them
• they will then hand back the papers and will receive the test papers to complete

3. <u>Hand out the contents of envelope A 1</u>.

4. Ask the pupils to learn the words silently for eight minutes. Please do not give any more information - whether about pronunciation, best way of learning, or whatever.

5. After eight minutes collect up the papers.

6. Hand out the test papers from envelope A 2. Ask the children to write their names and class number at the top of the page. This is most important. Then ask them to complete the answers to the test by filling in the blanks. The test order is different from the learning order. If you are asked about this, ask the children to deal with the items as they come. If they do not ask - do not comment!

7. After two minutes collect all the papers and return them to envelope A2. Please make sure that all pupils have registered their name and class number.

EXPERIMENT 1

The effect of order of presentation

Learning list exemplar.

NAME.....................................CLASS...........................

PLEASE LEARN THE FOLLOWING WORDS.

a wardrobe	-	une armoire
between	-	entre
later	-	plus tard
the knee	-	le genou
the poster	-	l'affiche
to bet	-	parier
a bus-stop	-	un arrêt
housework	-	le ménage
cold	-	froid
often	-	souvent
really	-	vraiment
to brush	-	brosser
washing-up	-	la vaisselle
to swim	-	nager
to wait for	-	attendre
the soap	-	le savon
to say	-	dire
seated	-	assis
old	-	vieux
love	-	l'amour

Experiment 1 continued.

Test list exemplar.

NAME.....................................CLASS...........................

PLEASE FILL IN THE BLANKS WITH THE FRENCH WORDS.

a wardrobe -
love -
a bus-stop -
later -
to wait for -
between -
really -
the knee -
seated -
to bet -
washing-up -
housework -
often -
to brush -
cold -
the poster -
to swim -
the soap -
to say -
old -

Experiment 1 continued.
Word-pairs. School A1. Group 1. Group 2 used the reverse form of this list.

a wardrobe	-	une armoire
between	-	entre
later	-	plus tard
the knee	-	le genou
the poster	-	l'affiche
to bet	-	parier
a bus-stop	-	un arrêt
housework	-	le ménage
cold	-	froid
often	-	souvent
really	-	vraiment
to brush	-	brosser
washing-up	-	la vaisselle
to swim	-	nager
to wait for	-	attendre
the soap	-	le savon
to say	-	dire
seated	-	assis
old	-	vieux
love	-	l'amour

Word-pairs. School B1. Group 1. Group 2 used the reverse form of this list.

a wardrobe	-	une armoire
between	-	entre
later	-	plus tard
the armchair	-	le fauteuil
the poster	-	l'affiche
to bet	-	parier
a car-jack	-	un cric
housework	-	le ménage
misty	-	brumeux
gently	-	doucement
really	-	vraiment
to brush	-	brosser
washing-up	-	la vaisselle
to swim	-	nager
to understand	-	entendre
the soap	-	le savon
to say	-	dire
standing	-	debout
old	-	vieux
love	-	l'amour

Experiment 1 continued.
Word-pairs. School A1 and School B1. Group 3. Group 4 used the reverse form of this list.

to amaze	-	étonner
among	-	parmi
a brake	-	un frein
cheerful	-	riant
a pushchair	-	une poussette
to complete	-	achever
to dry	-	sécher
an elbow	-	un coude
the cost	-	le frais
to heat	-	chauffer
a century	-	un siècle
the purpose	-	le but
to sew	-	coudre
a trial	-	un procès
sticky	-	collant
grating	-	grinçant
yet	-	pourtant
punishment	-	la peine
soaked	-	trempé
backwards	-	en arrière

EXPERIMENT 2

The transferability of list learning to testing in a simple context.

Generation.

Word-pairs. School A2 and School B2.

the wardrobe	-	l'armoire
between	-	entre
later	-	plus tard
the knee	-	le genou
the poster	-	l'affiche
to bet	-	*parier*
the bus-stop	-	l'arrêt
housework	-	le ménage
cold	-	froid
often	-	souvent
really	-	vraiment
to brush	-	se brosser
washing-up	-	la vaisselle
to swim	-	nager
to wait for	-	attendre
the soap	-	le savon
to say	-	dire
seated	-	assis
old	-	vieux
love	-	l'amour

Experiment 2 continued.

Exemplar for groups tested in context.

NAME..........................CLASS............................DATE............

The wardrobe is in the bedroom. .... est dans la chambre.
The book was about **love.** Le sujet du livre était ....
He gets on the bus at **the bus-stop.** Il monte dans l'autobus à ....
She is coming **later.** Elle vient ....
He has **to wait for** the bus. Il doit .... le car
Paul fell down **between** Marie and Jean. Paul est tombé .... Marie et Jean.
She is **really** beautiful. Elle est ....belle.
His **knee** was hurting him. Son .... lui faisait mal.

⋮

His grandfather is very **old.** Son grand-père est très ....

Experiment 2 continued. Test sentences for School A2 and School B2. Group 2. Day 1.

The wardrobe is in the bedroom.
.... est dans la chambre.

The book was about **love**.
Le sujet du livre était ....

He gets on the bus at **the bus-stop**.
Il monte dans l'autobus à ....

She is coming **later**.
Elle vient .... *
[Elle va venir ....]

He has **to wait for** the bus.
Il doit .... le car.

Paul fell down **between** Marie and Jean.
Paul est tombé .... Marie et Jean.

She is **really** beautiful.
Elle est .... belle.

His **knee** was hurting him.
Son .... lui faisait mal.

He was **seated** under a tree.
Il était .... sous un arbre.

Young people are forbidden **to bet**.
Il est interdit aux jeunes de ....

He is doing **the washing up**.
Il fait ....

She had to do **the housework**.
Elle a dû faire ....

He goes **often** to his aunt's house.
Il va ....chez sa tante.

She likes **to brush** her teeth.
Elle aime ....les dents.

It is **cold** in winter.
Il fait ....en hiver.

She has **the poster** for the concert.
Elle a .... pour le concert.

He likes **to swim** in the sea.
Il aime .... dans la mer.

The soap is in the bathroom.
.... est dans la salle de bain.

What do you want **to say**?
Qu'est-ce que vous voulez ....?

His grandfather is very **old**.
Son grand-père est très ....

Experiment 2 continued.
Test sentences for School A2 and School B2. Group 2. Day 2.

His coat was in **the wardrobe**.
Son manteau était dans ....

He spoke about the **love** of God.
Il parlait de .... de Dieu.

He saw John at **the bus-stop**.
Il a vu Jean à ....

His wife came **later**.
Sa femme est venue ....

He had **to wait for** his friends.
Il a dû .... ses amis.

He played **between** 2 and 2.30.
Il a joué .... 2h. et 2h.30.

He was **really** poor.
Il était .... pauvre.

The boy was holding his **knee**.
Le garçon se tenait ....

Everyone was **seated**.
Tout le monde était ....

He liked **to bet** on the horses.
Il aimait .... sur les chevaux.

He had a machine to do **the washing up**.
Il avait une machine pour faire ....

He did **the housework** before leaving.
Il a fait .... avant de partir.

It rains **often** in England.
Il pleut .... en Angleterre.

After the meal she had **to brush** her teeth.
Après le repas elle a dû .... les dents.

His bed was very **cold.**
Son lit était très ....

He saw **the poster** in the town.
Il a vu ....dans la ville.

He liked **to swim** in the pool.
Il aimait .... dans la piscine.

The soap was in the shower.
.... était dans la douche.

He did not know what **to say**.
Il ne savait pas quoi ....

The village was very **old**.
Le village était très ....

EXPERIMENT 3

The transferability of list learning to testing in a simple context.

Comprehension.

Word-pairs. School A3 and School D3.

the tyre	-	le pneu
outside	-	dehors
drunk	-	ivre
the pram	-	le landau
the forehead	-	le front
to succeed	-	réussir
the menu	-	la carte
a smile	-	un sourire
moody	-	maussade
except	-	sauf
forward	-	en avant
to wet	-	mouiller
the fight	-	la lutte
to heal	-	guérir
to sow	-	semer
the track	-	la trace
to worry	-	inquiéter
squeaky	-	grinçant
smooth	-	lisse
reward	-	la récompense

Experiment 3 continued.
Test sentences for School A3 and School D3. Group 2. Day 1.

Le pneu de son vélo était crevé.
............of his bike was punctured.

Il avait **un sourire** aux lèvres.
He had .... on his lips.

La récompense était de deux cents francs.
.... was two hundred francs.

La voiture allait **en avant.**
The car was going...........

Il savait **inquiéter** ses amis.
He knew how .... his friends.

Il voulait jouer **dehors.**
He wanted to play ....

Le dessus de la table était **lisse.**
The top of the table was ....

Son **front** était couvert de sang.
His .... was covered with blood.

Son père était **ivre** après la fête.
His father was ....after the party.

La lutte contre les drogues continue.
.... against drugs continues.

La carte au restaurant était longue.
.... at the restaurant was long.

Elle était belle **sauf** le nez.
She was beautiful .... for her nose.

Un docteur doit **guérir** les malades.
A doctor has .... the sick.

Son vélo était **grinçant**.
His bike was ....

L'enfant aimait **le landau.**
The baby liked ....

Elle a dû **semer** les graines pour les légumes.
She had .... the seed for the vegetables.

La trace de l'assassin était difficile à suivre.
.... of the murderer was hard to follow.

Il voulait **réussir** dans ses études.
He wanted .... in his studies.

Il a dû lui **mouiller** les cheveux avant de les peigner.
He had .... his hair before combing it.

Il était jeune et toujours **maussade**.
He was young and always ....

Experiment 3 continued.

Test sentences for School A3 and School B3. Group 2. Day 2.

Le pneu de son vélo était dégonflé.
.... of his bike was flat.

Il lui parlait avec **un sourire.**
He spoke to him with ....

La récompense de ses services était grande.
.... for his services was great.

L'économie va toujours **en avant.**
The economy is always going ....

Il ne voulait pas **inquiéter** ses parents.
He did not want .... his parents.

Le chien ne voulait pas rester **dehors.**
The dog did not want to stay ....

La surface du miroir était **lisse.**
The surface of the mirror was ....

Son **front** était plissé.
His .... was wrinkled.

Il était légèrement **ivre** après le repas.
He was slightly......... after the meal.

La lutte antipollution est très importante.
.... against pollution is very important.

La carte au restaurant chinois était compliquée.
.... at the Chinese restaurant was complicated.

Tout le monde a été sauvé **sauf** lui .
Everyone was saved.............. him.

Le sommeil peut **guérir** les maladies.
Sleep is able ....illnesses.

Le gond de la porte était **grinçant.**
The hinge of the door was ....

L'enfant dormait dans **le landau.**
The child was sleeping in ....

Il faut **semer** les graines en hiver.
You have .... seeds in winter.

La trace du renard est toujours claire.
.... of a fox is always clear.

Il faut **réussir** pour être riche.
You have .... to be rich.

La pluie commençait à **mouiller** la foule.
The rain started .... the crowd.

Il avait toujours un air **maussade.**
He always had a .... look about him.

EXPERIMENT 4

Learning in a list versus learning in a context

Exemplar for groups learning in context.

NAME............................CLASS..............................DATE............

There was **a monkey** in the circus.
Il y avait **un singe** dans le cirque.

Perhaps it will be fine tomorrow.
Peut-être *fera-t-il beau demain.*

The old man was very **kind.**
Le vieil homme était très **gentil.**

He liked to take **a walk** in the afternoon.
Il aimait faire **une promenade** l'après-midi.

He had a large **chin.**
Il avait un grand **menton.**

He had **to clean** the house.
Il a dû **nettoyer** la maison.

He was going **fishing.**
Il allait à **la pêche.**

He had **a fortnight** in France.
Il a passé **une quinzaine** en France. *
[Il a passé **une quinzaine** de jours en France].

.
.
.
.
.
.
.
.
.
.

He paid **the rent** for the flat.
Il a payé **le loyer** pour l'appartement. *
[Il a payé **le loyer** de l'appartement].

Experiment 4 continued.
Word-pairs. School A4 and School B4. Group 1.

a monkey	-	un singe
perhaps	-	peut-être
kind	-	gentil
a walk	-	une promenade
the chin	-	le menton
to clean	-	nettoyer
fishing	-	la pêche
a fortnight	-	une quinzaine *
		[une quinzaine de jours]
rushed	-	pressé
sometimes	-	quelquefois
alright	-	d'accord
reply	-	répondre
the cost	-	le coût
to comb	-	se peigner
to joke	-	plaisanter
swimming	-	la natation
to return	-	rentrer
busy	-	affairé
thin	-	maigre
the rent	-	le loyer

Experiment 4 continued.
Word-pairs in sentence context. School A4 and School B4. Group 2.

There was **a monkey** in the circus.
Il y avait **un singe** dans le cirque.

Perhaps it will be fine tomorrow.
Peut-être fera-t-il beau demain.

The old man was very **kind.**
Le vieil homme était très **gentil.**

He liked to take **a walk** in the afternoon.
Il aimait faire **une promenade** l'après-midi.

He had a large **chin.**
Il avait un grand **menton.**

He had **to clean** the house.
Il a dû **nettoyer** la maison.

He was going **fishing.**
Il allait à **la pêche.**

He had **a fortnight** in France.
Il a passé **une quinzaine** en France. *
[Il a passé **une quinzaine** de jours en France].

Everyone is **rushed** these days.
Tout le monde est **pressé** aujourd'hui.

Sometimes he went to his aunt's.
Quelquefois il allait chez sa tante.

It's **alright** for tomorrow.
C'est **d'accord** pour demain.

He had **to reply** to the letter.
Il a dû **répondre** à la lettre.

The cost of living has risen.
Le coût de la vie a augmenté.

She liked **to comb** her hair.
Elle aimait **se peigner** les cheveux.

To joke with him was always easy.
Plaisanter avec lui était toujours facile.

Swimming is a popular sport.
La natation est un sport populaire.

He did not want **to return** after the holiday.
Il ne voulait pas **rentrer** après les vacances.

In the morning he was very **busy**.
Le matin il était très **affairé**.

After his illness he was very **thin**.
Après sa maladie il était très **maigre**.

He paid **the rent** for the flat.
Il a payé **le loyer** pour l'appartement. *
[Il a payé **le loyer** de l'appartement].

Experiment 4 continued.
Test sentences for School A4 and School B4. Groups 1 and 2. Day 1.

He saw **a monkey** in the tree.
Il a vu .... dans l'arbre.

The rent for the house was high.
.... pour la maison était élevé. *
[............ de la maison était élevé].

Fishing was his great pleasure.
.... était son grand plaisir.

His face was **kind.**
Son visage était ....

It is difficult **to joke** when you are sad.
Il est difficile de .... quand on est triste.

Perhaps he was going too fast.
.... allait-il trop vite.

'Agreed', he said, smiling.
.... dit-il, en souriant.

It was too cold to take **a walk.**
Il faisait trop froid pour faire ....

He was too **busy** to take a holiday.
Il était trop........ pour prendre des vacances.

The car was easy **to clean.**
La voiture était facile à ....

The cost of credit is high.
.... du crédit est haut. *
[.............. du crédit est élevé].

He was ill for **a fortnight.**
Il était malade pendant .... *
[Il était malade pendant ....de jours].

Sometimes he couldn't sleep.
.... il ne pouvait pas dormir.

I must **reply** to his question.
Je dois .... à sa question.

He was too **rushed** to eat.
Il était trop .... pour manger.

He cut his **chin.**
Il s'est coupé ....

He did not want **to comb** his hair.
Il ne voulait pas .... les cheveux.

Swimming in the sea is difficult. *
.... dans la mer est difficile. *
[**Swimming** is an Olympic sport].
[.............. est un sport olympique].

It was difficult **to return** to his office.
Il était difficile de .... au bureau.

He was **thin** as a rake.
Il était .... comme un clou.

Experiment 4 continued.
Test sentences for School A4 and School B4. Groups 1 and 2. Day 2.

The monkey was eating a banana.
.... mangeait une banane.

He had no money for **the rent.**
Il n'avait pas d'argent pour ....

The boys loved to go **fishing.**
Les garçons aimaient aller à ....

He always has a **kind** word for everyone.
Il a toujours un mot .... pour chacun.

I'm not in the mood **to joke.**
Je ne suis pas d'humeur à ....

Perhaps he was too young.
.... était-il trop jeune.

'Agreed', he said, closing the book.
.... dit-il, en fermant le livre.

He liked to take **a walk** by the sea.
Il aimait faire .... au bord de la mer.

He always seemed to be **busy.**
Il avait toujours l'air d'être ....

He wanted **to clean** the windows.
Il voulait .... les fenêtres.

The cost of living is still high.
.... de la vie est encore haut. *
[............... de la vie est encore élevé].

He was travelling for **a fortnight.**
Il voyageait pendant .... *
[Il voyageait pendant ....de jours].

Sometimes he was very happy.
.... il était très heureux.

You must **reply** to letters.
On doit .... aux lettres.

I am really **rushed** today.
Je suis vraiment .... aujourd'hui.

His **chin** was covered in spots.
Son .... était couvert de boutons.

After his shower he had **to comb** his hair.
Après la douche il a dû .... les cheveux.

Swimming is good for the health.
.... est bonne pour la santé.

He liked **to return** home.
Il aimait .... chez lui.

He ate a lot but stayed **thin.**
Il mangeait beaucoup mais il restait ....

Experiment 4 continued.

Word-pairs. School A4 and School B4. Group 3.

the platform	-	le quai
despite	-	malgré
nice	-	chouette
a path	-	un sentier
the cheek	-	la joue
to dirty	-	salir
happiness	-	le bonheur
a diversion	-	une déviation
forgetful	-	distrait
gently	-	doucement
rather	-	plutôt
to demand	-	exiger
boredom	-	l'ennui
to murder	-	assassiner
to warn	-	alerter
cooking	-	la cuisine
to lend	-	prêter
free	-	gratuit
ugly	-	laid
the meaning	-	le sens

Experiment 4 continued.
Word-pairs in sentence context. School A4 and School B4. Group 4.

He was waiting for the train on **the platform.**
Il attendait le train sur **le quai.**

He was warm **despite** the wind.
Il avait chaud **malgré** le vent.

His friend was very **nice**.
Son ami était très **chouette**.

He found **a path** in the forest.
Il a trouvé **un sentier** dans la forêt.

He cut his **cheek** with the razor.
Il s'est coupé **la joue** avec le rasoir.

He did not want **to dirty** his shoes.
Il ne voulait pas **salir** ses chaussures.

Happiness is rare.
Le bonheur est rare.

There was **a diversion** after the accident.
Il y avait **une déviation** après l'accident.

He looked **forgetful**.
Il avait un air **distrait**.

He always spoke **gently**.
Il parlait toujours **doucement**.

He is **rather** rich.
Il est **plutôt** riche.

He wanted **to demand** an apology.
Il voulait **exiger** des excuses.

Boredom is a problem for the young.
L'ennui est un problème pour les jeunes.

He wanted **to murder** the president.
Il voulait **assassiner** le président.

He wanted **to warn** the police.
Il voulait **alerter** la police.

French **cooking** is famous.
La cuisine française est renommée.

Would you like **to lend** me a pound?
Voulez-vous me **prêter** une livre?

The ticket was **free**.
Le billet était **gratuit**.

The dog was very **ugly**.
Le chien était très **laid**.

The meaning of the book was clear.
Le sens du livre était clair.

Experiment 4 continued.

Test sentences for School A4 and School B4. Groups 3 and 4. Day 1.

Elle l'attendait sur **le quai.**
She was waiting for him on ....

Il ne voulait pas **salir** sa chemise.
He was afraid .... his shirt.

Il y avait **un sentier** entre les maisons.
There was .... between the houses.

Il a dû prendre **une déviation** avant la ville.
He had to take .... before the town.

L'Angleterre n'est pas renommée pour **la cuisine.**
England is not famous for its ....

La vieille dame parlait **doucement** à l'enfant.
The old lady spoke .... to the child.

Sa nouvelle chemise était **chouette**.
Her new dress was very ....

Il a dû **exiger** ses droits.
He had .... his rights.

Il a crié pour **alerter** son ami.
He shouted .... his friend.

L'ennui signifie la stupidité.
.... is a sign of stupidity.

Tout le monde cherche **le bonheur.**
Everyone is looking for ....

Il a tenté d'**assassiner** sa femme.
He tried .... his wife.

Il est **plutôt** malheureux aujourd'hui.
He is .... unhappy today.

Voulez-vous me **prêter** votre livre?
Would you .... me your book?

Sa **joue** était couverte de sang.
His .... was covered in blood.

Le voyage était **gratuit**.
The journey was ....

Le vieil homme était très **laid**.
The old man was very ....

Il était toujours **distrait**.
He was always ....

Il ne pouvait pas trouver **le sens** du poème.
He could not see .... of the poem.

Il faisait froid **malgré** le soleil.
It was cold .... the sun.

Experiment 4 continued.
Test sentences for School A4 and School B4. Groups 3 and 4. Day 2.

La foule sur **le quai** attendait le train.
The crowd on .... was waiting for the train.

Il ne voulait pas **salir** le tapis.
He did not want .... the carpet.

Il pouvait voir **un sentier** dans la neige.
He could see .... in the snow.

Il y avait **une déviation** autour de la ville.
There was .... round the town.

Il n'aimait pas **la cuisine** anglaise.
He did not like English ....

Il a touché le bras du vieil homme **doucement.**
He touched the old man's arm ....

Elle trouvait sa nouvelle école très **chouette.**
She found her new school really ....

Exiger n'est pas poli.
.... is not polite.

Il a téléphoné pour **alerter** ses parents.
He phoned .... his parents.

L'ennui cause le sommeil.
.... causes sleep.

Il a trouvé **le bonheur** dans l'amour.
He found .... in love.

Assassiner est un crime grave.*
[**Assassiner** est un délit grave].
.... is a serious crime.

Elle est **plutôt** pauvre.
She is ..... poor.

Il voulait me **prêter** sa voiture.
He wanted .... me his car.

Sa **joue** était très rouge.
Her .... was very red.

Le premier exemplaire était **gratuit**.
The first copy was ....

Son visage était très **laid.**
His face was very ....

Quand il était fatigué, il était **distrait.**
When he was tired he was ....

Chacun doit chercher **le sens** de la vie.
Each person must seek .... of life.

Il était triste **malgré** son argent.
He was sad .... his money.

EXPERIMENT 5.

List position and serial order effects and word-pair presentation.

Generation.

Exemplar for groups learning in disrupted list.

Please learn the word indicated in bold type:
The English word is printed on the left in each case.

the tyre	The tyre of his car was flat. **Le pneu** de sa voiture était dégonflé.
despite	Despite his illness he was happy. **Malgré** sa maladie il était content.
smooth	The surface of the mirror was smooth. La surface du miroir était **lisse.**
a walk	It was too cold to take a walk. Il faisait trop froid pour faire **une promenade.**
forehead	His forehead was wrinkled. Son **front** était plissé.
to clean	She liked to clean the car. Elle aimait **nettoyer** la voiture.
fishing	Fishing was his great pleasure. **La pêche** était son grand plaisir.
a fortnight	He was ill for a fortnight. Il était malade pendant **une quinzaine.** * [Il était malade pendant **une quinzaine** de jours].
⋮	
reply	I must reply to his question Je dois **répondre** à sa question.

Experiment 5 continued.
Word-pairs. School A5 and C5. Group 1.

the tyre	-	le pneu
despite	-	malgré
smooth	-	lisse
a walk	-	une promenade
forehead	-	le front
to clean	-	nettoyer
fishing	-	la pêche
a fortnight	-	une quinzaine. *
		[une quinzaine de jours].
rushed	-	pressé
except	-	sauf
agreed	-	d'accord
the meaning	-	le sens
to sow	-	semer
to joke	-	plaisanter
swimming	-	la natation
to return	-	rentrer
nice	-	chouette
thin	-	maigre
the pram	-	le landau
to reply	-	répondre

Experiment 5 continued.
Word-pairs in sentence context. School A5 and School B5. Group 2 and Group 3.

The tyre of his car was flat.
Le pneu de sa voiture était dégonflé.

Despite his illness he was happy.
Malgré sa maladie, il était content.

The surface of the mirror was **smooth.**
La surface du miroir était **lisse.**

It was too cold to take **a walk.**
Il faisait trop froid pour faire **une promenade.**

His **forehead** was wrinkled.
Son **front** était plissé.

She liked **to clean** the car.
Elle aimait **nettoyer** la voiture.

Fishing was his great pleasure.
La pêche était son grand plaisir.

He was ill for **a fortnight.**
Il était malade pendant **une quinzaine.** *
[Il était malade pendant **une quinzaine** de jours].

He was too **rushed** to eat.
Il était trop **pressé** pour manger.

Everyone was ill **except** him.
Tout le monde était malade **sauf** lui.

'Agreed', he said, smiling.
'D'accord', dit-il, en souriant.

The meaning of the word was unclear.
Le sens du mot n'était pas clair.

You have **to sow** seeds in winter.
Il faut **semer** les graines en hiver.

It is difficult **to joke** when you are sad.
Il est difficile de **plaisanter** quand on est triste.

Swimming in the sea is difficult.
La natation dans la mer est difficile. *
[**Swimming** is an Olympic sport].
[**La natation** est un sport olympique].

He had **to return** after the holidays.
Il a dû **rentrer** après les vacances.

His friend was always **nice**.
Son ami était toujours **chouette**.

He was as **thin** as a rake.
Il était **maigre** comme un clou.

The child was sleeping in **the pram.**
L'enfant dormait dans **le landau.**

I have **to reply** to his question.
Je dois **répondre** à sa question.

Experiment 5 continued.
Test sentences for School A5 and School C5. All groups.

He forgot **the tyre** for his bike
Il a oublié .... pour son vélo.

Her husband wanted **to clean** the kitchen.
Son mari voulait .... la cuisine.

She had to take **a walk** every day.
Elle devait faire .... chaque jour.

His holidays lasted **a fortnight.**
Ses vacances duraient ....

To return home is always easy.
.... chez vous est toujours facile. *
[.............. chez soi est toujours facile].

They all left **except** him.
Tout le monde .... lui est parti. *
[Tout le monde est parti .... lui].

The sea was **smooth** yesterday.
La mer était ....hier.

To find **the meaning** of life is difficult.
Trouver .... de la vie est difficile.

Swimming in a pool is expensive.
.... dans une piscine coûte cher. *
[**Swimming** is good for the health].
[.............. est bonne pour la santé].

He went **to sow** the seeds in the garden.
Il est allé ....les graines dans le jardin.

Fishing is a sport for all.
.... est un sport pour chacun.

He loved **to joke** with his friends.
Il aimait .... avec ses copains.

They were **agreed**.
Ils étaient ....

His new house was **nice**.
Sa maison nouvelle était ....

His **forehead** was covered with blood.
Son .... était couvert de sang.

The sick man was very **thin**.
Le malade était très ....

She put the baby in **the pram.**
Elle a mis l'enfant dans ....

Despite his age he could run very quickly.
.... son âge il pouvait courir très vite.

He did not like being **rushed**.
Il n'aimait pas être ....

He wanted **to reply** but he couldn't.
Il voulait .... mais il ne pouvait pas.

EXPERIMENT 6.

List position and serial order effects and word-pair presentation.

Comprehension.

Word-pairs. Group 1.

the deck	-	le pont
rough	-	rêche
to flow	-	couler
the claw	-	la griffe
health	-	la santé
laughing	-	riant
the liver	-	le foie
till	-	jusqu'à
to scare	-	effrayer
the wedding	-	les noces
beyond	-	au-delà
to hide	-	cacher
opposite	-	en face de
treated	-	traité
reduced	-	réduit
the average	-	la moyenne
the ashtray	-	le cendrier
to complete	-	achever
the crew	-	l'équipage
to teach	-	enseigner

Experiment 6 continued.
Word-pairs in sentence context. Group 2 and Group 3.

The deck of the ship was wet.
Le pont du bateau était mouillé.

Your skin becomes **rough** in the sun.
La peau devient **rêche** au soleil.

The river seemed **to flow** to the east.
La rivière semblait **couler** vers l'est.

The cat had a wound on **the claw.**
Le chat avait une blessure à **la griffe.**

Nothing is more important than **health.**
Rien n'est plus important que **la santé.**

To be always **laughing** is very appealing.
Etre toujours **riant** est très attirant.

The liver is damaged by alcohol.
Le foie se détruit par l'alcool. *
[L'alcool détruit **le foie**].

He waited patiently for him **till** the train arrived .
Il l'attendait avec patience **jusqu'à** l'arrivée du train.

She liked **to scare** her brother.
Elle aimait **effrayer** son frère.

The wedding took place in the church.
Les noces ont eu lieu dans l'église.

There was a wide valley **beyond.**
Il y avait une grande vallée **au-delà.**

She was not able **to hide** the money.
Elle ne pouvait pas **cacher** l'argent.

Opposite the station was a square.
En face de la gare il y avait une place.

In hospital he was always well **treated.**
A l'hôpital il était toujours bien **traité.**

The price of cheese was **reduced** in the shop.
Le prix du fromage était **réduit** dans le magasin.

The average of the class was high.
La moyenne de la classe était élevée.

He put his cigarettes in **the ashtray.**
Il a mis ses cigarettes dans **le cendrier.**

He needed **to complete** his work.
Il devait **achever** son travail.

The crew of the ship were very nice.
L'équipage du bateau était chouette.

She wanted **to teach** the children to swim.
Elle voulait **enseigner** les enfants à nager. *
[She wanted **to teach** the children maths].
[Elle voulait **enseigner** les mathématiques aux enfants].

Experiment 6 continued.
Test sentences. All groups.

Il aimait voyager sur **le pont** du bateau.
He liked to travel on .... of the ship.

Une rivière veut toujours **couler** vers la mer.
A river always wants .... to the sea.

La bonne nourriture est nécessaire pour **la santé.**
Good food is necessary for ....

La griffe du tigre est pointue.
.... of a tiger is sharp.

Il voulait **achever** sa tâche.
He wanted .... his task.

Au-delà de la montagne il y avait la mer.
.... the mountain was the sea.

Sa peau était **rêche** à cause du travail.
Her skin was .... because of work.

Son score était plus haut que **la moyenne.** *
[Son score était plus élevé que **la moyenne**].
His score was above ....

L'équipage de l'avion était anglais.
.... of the plane was English.

Effrayer quelqu'un est dangereux.
.... someone is dangerous.

Il était là pour **les noces** de son amie.
He was there for .... of his friend.

Elle tentait de **cacher** le gâteau.
She was trying .... the cake.

Il y avait un bar **en face de** sa maison.
There was a bar .... his house.

Il était bien **traité** par la police.
He was well .... by the police.

Le foie est un organe vital.
.... is a vital organ.

Le prix du livre était **réduit** après Noël.
The price of the book was .... after Christmas.

Le cendrier est tombé par terre.
.... fell to the ground.

Il restait dehors **jusqu'à** son arrivée.
He stayed outside .... she arrived.

Elle voulait voir son visage **riant.**
She wanted to see his .... face.

Il est allé **enseigner** dans une école.
He went .... in a school.

EXPERIMENT 7.

Word frequency and list position. Generation.

Word-pairs. Group 1 and Group 2.

forgetful	-	distrait
to feel	-	sentir
to heal	-	guérir
the cost	-	le coût
seldom	-	rarement
the century	-	le siècle
the monkey	-	le singe
sometimes	-	parfois
happiness	-	le bonheur
free	-	gratuit
the rent	-	le loyer
kind	-	gentil
bored	-	ennuyé
to use	-	employer
the noise	-	le bruit
business	-	les affaires
to comb	-	peigner
the field	-	le champ
the chin	-	le menton
the war	-	la guerre.

Experiment 7 continued.
Test sentences. Group 2.

She **seldom** went home.
Elle allait.............. chez elle.

The cost of the journey was too high.
.... du voyage était trop élevé.

He always had a **forgetful** look.
Il avait toujours un air ....

He saw **the monkey** in the tree.
Il a vu .... dans l'arbre.

He had a cut on **the chin.**
Il s'était coupé ....

Business was good.
.... allaient bien.

There were changes throughout **the century.**
Il y avait des changements pendant ....

The noise of the crowd was terrible.
...............de la foule était terrible.

He had to pay **the rent** for the flat.
Il a dû payer .... pour l'appartement. *
[Il a dû payer .... de l'appartement.]

The war began in 1939.
.... a commencé en 1939.

He was always **bored.**
Il était toujours............

It is difficult to find **happiness.**
Il est difficile de trouver ....

The doctor tries **to heal** the sick.
Le docteur tente de .... les malades.

She liked **to feel** the wind on her face.
Elle aimait .... le vent sur son visage.

She wanted **to use** the car.
Elle voulait .... la voiture.

The ticket for the concert was **free.**
Le billet pour le concert était ....

The horse was waiting in **the field.**
Le cheval attendait dans ....

It was difficult **to comb** his hair.
Il était difficile de lui .... les cheveux.

He was always **kind** with children.
Il était toujours............ avec les enfants.

Sometimes he used to play in the park.
.... il jouait dans le parc.

EXPERIMENT 8.

Word frequency and list position. Comprehension.

Word-pairs. Group 1 and Group 2.

daring	-	hardi
to believe	-	croire
to clog	-	boucher
the land	-	la terre
crazy	-	fou
the rate	-	le taux
the seam	-	la couture
close	-	proche
the bush	-	le buisson
quite	-	assez
the herd	-	le troupeau
light	-	léger
relaxed	-	détendu
to place	-	mettre
the mood	-	l'humeur
thought	-	la pensée
to clip	-	tailler
half	-	la moitié
the sand	-	le sable
the end	-	la fin

Experiment 8 continued.
Test sentences. Group 2.

Boucher un tuyau peut être dangereux.
.... a pipe can be dangerous.

Elle voulait **mettre** le vase sur la table.
She wanted .... the vase on the table.

Il était toujours **détendu** chez lui.
He was always .... at home.

Elle aimait **croire** qu'il viendrait.
She liked .... that he would come.

Il avait toujours un air **hardi**.
He always had a ....look.

La couture de la chemise était défaite.
.... of the shirt was unpicked.

L'amour de **la terre** est normal.
Love of .... is natural.

La fin du séjour s'est bien passée.
.... of the stay went well.

Son nouveau vélo était **léger**.
His new bike was ....

La dignité de l'homme est dans **la pensée.**
Human dignity is based on the power of ....

Elle a trouvé le chat sous **le buisson.**
She found the cat under ....

Il dormait sur **le sable.**
He was asleep on ....

Il a dû laisser **le troupeau** dans le champ.
He had to leave .... in the field.

L'humeur de la foule était terrible.
.... of the crowd was terrible.

Au mois de décembre, l'été n'est pas **proche.**
In December, the summer is not ....

Le taux de la montée des prix est difficile pour tous.
.... of the rise in prices is difficult for everyone.

La moitié du monde a faim.
.... the world is hungry.

Il était difficile de **tailler** la haie.
It was difficult .... the hedge.

Elle pensait que son frère était **fou**.
She thought her brother was ....

Le billet pour le concert était **assez** cher.
The ticket for the concert was .... expensive.

EXPERIMENT 9

The effects of word category on recall

Word-pairs. Group 1 and Group 2.

the flesh	-	la chair
alive	-	vivant
to grow	-	pousser
equally	-	également
the strike	-	la grève
the spot	-	la tache
tall	-	élevé
to dust	-	essuyer
besides	-	outre
the taste	-	le goût
the rifle	-	le fusil
thick	-	épais
to flow	-	couler
beneath	-	sous
spite	-	le dépit
a bear	-	un ours
wet	-	mouillé
to dry	-	se sécher
towards	-	vers
the search	-	la fouille

Experiment 9 continued.
Test sentences. Group 1. Group 2 used the reverse form of these sentences.

The money was divided **equally** among them.
L'argent était divisé .... entre eux. *
[L'argent était .... divisé entre eux].

Spite is not a good motive.
.... n'est pas bon comme raison.

After the rain the roof was **wet**.
Après la pluie le toit était ....

The search of the luggage was finished.
.... des bagages était finie.

To grow when you are young is normal.
.... pendant la jeunesse est normal.

He saw **a bear** in the zoo.
Il a vu .... dans le zoo.

He wanted **to dry** his hair.
Il voulait .... les cheveux.

He was walking **towards** his mother.
Il marchait .... sa mère.

The flesh of a peach is firm.
.... d'une pêche est ferme.

He was **alive** after the accident.
Il était .... après l'accident.

He tried to hide **the spot** on his shirt.
Il tentait de cacher .... sur la chemise.

He liked **the taste** of beer.
Il aimait .... de la bière.

The new building was very **tall**.
Le nouveau bâtiment était très ....

He carried **the rifle** in his car.
Il portait .... dans sa voiture.

His finger was very **thick**.
Son doigt était très ....

She needed **to dust** the table.
Elle devait .... la table.

The water began **to flow.**
L'eau a commencé à ....

He was found **beneath** the bridge.
Il a été trouvé .... le pont.

The strike lasted three months.
.... a duré trois mois.

Besides Jean, the room was empty.
.... Jean, la salle était vide.

EXPERIMENT 10

The effect of embedded words on memory processes

Experiment 10.
Items with embedding.

English item	Frequency	French item	Embedded word(s)	Embedded word frequency
Verbs				
to rattle	5	ébranler	ran	134
			bran	1
to wrap	5	emballer	all	3001
			ball	110
to chatter	7	causer	use	589
			cause	130
			user	4
to restrict	11	borner	born	113
to hover	4	planer	plane	114
			lane	30
Nouns				
the curls	1	la frisure	sure	264
a refill	3	une recharge	charge	122
the roundabout	2	le manège	man	1207
a salesman	12	un démarcheur	march	120
			arch	13
the socket	3	la prise	rise	102

Experiment 10.

Items without embedding or with low-frequency words embedded.

English item	Frequency	French item	Embedded word	Embedded word frequency
Verbs				
to sag	4	fléchir	-	-
to peck	5	picoter	cot	1
to shorten	4	abréger	-	-
to curse	11	maudire	dire	1
to tease	6	taquiner	-	-
Nouns				
a tablet	3	un cachet	ache	4
a boulder	10	un rocher	-	-
a boil	12	un clou	-	-
the froth	1	l'écume	-	-
the sow	3	la truie	-	-

Word-pairs. Group 1 and Group 2.

the curls	-	la frisure
to sag	-	fléchir
a boulder	-	un rocher
to wrap	-	emballer
a refill	-	une recharge
to peck	-	picoter
a boil	-	un clou
to chatter	-	causer
the roundabout	-	le manège
to shorten	-	abréger
the froth	-	l'écume
to restrict	-	borner
a salesman	-	un démarcheur
to tease	-	taquiner
the sow	-	la truie
to hover	-	planer
the socket	-	la prise
to curse	-	maudire
a tablet	-	un cachet
to rattle	-	ébranler

Experiment 10 continued.
Test sentences. Group 1. Group 2 used the reverse form of these sentences.

The wind began **to rattle** the windows.
Le vent a commencé à .... les fenêtres.

The froth on the sand was very white.
.... sur le sable était très blanche.

A salesman came to the door.
.... est arrivé à la porte.

The hawk liked **to hover** above the house.
Le faucon aimait .... au-dessus de la maison.

He had to take **a tablet** every day.
Il devait prendre .... chaque jour.

He wanted **to restrict** her freedom.
Il voulait .... sa liberté.

He wanted **to tease** the cat.
Il voulait .... le chat.

They saw **the sow** in the field.
Ils ont vu .... dans le champ.

The socket was behind the chair.
.... se trouvait derrière le fauteuil.

The witch began **to curse** the cat.
La sorcière commencait à .... le chat.

The roof was beginning **to sag.**
Le toit commencait à ....

He had **a boil** on his nose.
Il avait .... au nez.

There was **a boulder** in the pond.
Il y avait .... dans l'étang.

The chicken tried **to peck** the dog.
Le poulet tentait de .... le chien.

It was forbidden **to chatter** in class.
Il était interdit de .... dans la classe.

She needed **to wrap** the present.
Elle devait.............. le cadeau.

The roundabout at the fair was very small.
.... à la foire était très petit.

He wanted **to shorten** the meeting.
Il voulait .... la réunion.

She loved **the curls** on his head.
Elle aimait .... sur sa tête.

He needed **a refill** for his pen.
Il avait besoin d'.............. pour son stylo.

RESULTS APPENDIX

INDEX

EXPERIMENT 1

The effect of order of presentation

School A1 and School B1. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	F	p
School	1	55837.49	55837.49	45.51	< 0.01
Group	3	136293.70	45431.23	37.03	< 0.01
School x Group	3	3165.14	1055.04	0.86	0.46
Error	181	222027.84	1226.67		
Day	3	19491.75	6497.25	79.87	< 0.01
School x Day	3	1348.60	449.53	5.52	< 0.01
Group x Day	9	1764.06	196.00	2.41	< 0.05
Error	543	44168.24	81.34		

School A1. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	F	p
Association	1	62714.91	62714.91	44.26	< 0.01
Direction of learning	1	839.09	839.09	0.59	0.44
Association x Direction of learning	1	10070.16	10070.16	7.10	< 0.01
Error	89	126091.98	1416.76		
Day	3	5497.55	1832.51	23.60	< 0.01
Direction of learning x Day	3	393.21	131.07	1.68	0.16
Error	267	20728.95	77.63		

School B1. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Association	1	40184.91	40184.91	38.53	< 0.01
Direction of learning	1	92.40	92.40	0.08	0.76
Association x Direction of learning	1	25821.61	25821.61	24.76	< 0.01
Error	92	95935.86	1042.78		
Day	3	15008.82	5002.94	58.91	< 0.01
Direction of learning x Day	3	271.55	90.51	1.06	0.36
Error	276	23439.29	84.92		

EXPERIMENT 2

The transferability of list learning to testing in a simple context.

Generation.

School A2 and School B2. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
School	1	5085.03	5085.03	9.68	< 0.01
Group	1	20269.05	20269.05	38.58	< 0.01
School x Group	1	8219.93	8219.93	15.64	< 0.01
Error	80	42023.12	525.28		
Day	1	15408.73	15408.73	154.93	< 0.01
School x Day	1	8.73	8.73	0.08	0.76
Group x Day	1	118.50	118.50	1.19	0.27
School x Group x Day	1	2.44	2.44	0.02	0.87
Error	80	7956.38	99.45		0.87

School B2. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Group	1	28476.64	28476.64	47.10	< 0.01
Error	36	19011.84	528.10		
Day of testing	1	6726.64	6726.64	93.28	< 0.01
Group x Day of testing	1	39.80	39.80	0.55	0.46
Error	36	2596.09	72.11		

School A2. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Group	1	1471.32	1471.32	2.81	0.11
Error	44	23011.28	522.98		
Day of testing	1	8888.58	8888.58	72.96	< 0.01
Group x Day of testing	1	85.31	85.31	0.70	0.40
Error	44	5360.33	121.82		

EXPERIMENT 3

The transferability of list learning to testing in a simple context.

Comprehension.

School A3 and School B3. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	*F*	*p*
School	1	18316.79	18316.79	43.89	< 0.01
Group	1	2263.29	2263.29	5.49	< 0.05
School x Group	1	5.42	5.42	0.01	0.90
Error	82	34218.21	417.29		
Day	1	6808.53	6808.53	90.98	< 0.01
School x Day	1	751.02	751.02	10.03	< 0.01
Group x Day	1	0.82	0.82	0.01	0.91
School x Group x Day	1	617.78	617.78	8.25	< 0.01
Error	82	6136.50	74.83		

School B3. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	*F*	*p*
Group	1	995.02	995.02	1.82	0.18
Error	39	21220.21	21220.21		
Day of Testing	1	5792.04	5792.04	76.28	< 0.01
Group x Day of testing	1	274.97	274.97	3.62	0.06
Error	39	2961.00	75.92		

School A3. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Group	1	1317.55	1317.55	4.35	< 0.05
Error	43	12998.00	12998.00		
Day of testing	1	1586.72	1586.72	21.48	< 0.01
Group x Day of testing	1	346.72	346.72	4.69	< 0.05
Error	43	3175.50	73.84		

EXPERIMENT 4

Learning in a list versus learning in a context

School A4 and School B4. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	*F*	*p*
School	1	20762.65	20762.65	51.90	< 0.01
Group	3	21964.24	7321.41	18.03	< 0.01
School x Group	3	8698.81	2899.60	7.24	< 0.01
Error	165	65997.35	399.98		
Day	1	19950.66	19950.66	175.85	< 0.01
School x Day	1	919.47	919.47	8.10	< 0.01
Group x Day	3	587.68	195.89	1.72	0.16
School x Group x Day	3	777.78	259.26	2.28	0.08
Error	165	18719.56	113.45		

School B4. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	*F*	*p*
Learning condition	1	12187.10	12187.10	28.29	< 0.01
Test	1	223.01	223.01	0.51	0.47
Learning x Test	1	11532.98	11532.98	26.77	< 0.01
Error	79	34030.47	430.76		
Day	1	14196.57	14196.57	125.82	< 0.01
Learning condition x Day	1	0.91	0.91	0.08	0.92
Test x Day	1	319.27	319.27	2.83	0.09
Learning condition x Test x Day	1	34.18	34.18	0.30	0.58
Error	79	8913.18	112.82		

School A4. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Learning condition	1	162.56	162.56	0.43	0.51
Test	1	895.00	895.00	2.40	0.12
Learning x Test	1	5005.56	5005.56	13.46	< 0.01
Error	86	31966.87	371.70		
Day	1	6386.67	6386.67	56.01	< 0.01
Learning condition x Day	1	35.00	35.00	0.30	0.58
Test x Day	1	203.06	203.06	1.78	0.18
Learning condition x Test x Day	1	798.06	798.06	6.99	< 0.01
Error	86	9806.37	114.02		

EXPERIMENT 5

List position and serial order effects and word-pair presentation. Generation.

School A5 and School C5. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	*F*	*p*
School	1	2535.56	2535.56	2.11	0.14
Group	2	4102.72	2051.36	1.70	0.18
School x Group	2	2238.84	1119.42	0.93	0.39
Error	104	124849.77	1200.47		
List position	2	41297.99	20648.99	92.38	< 0.01
School x List position	2	3749.53	1874.76	8.38	< 0.01
Group x List position	4	1393.26	348.31	1.55	0.18
School x Group x List position	4	2383.09	595.77	2.66	< 0.05
Error	208	46492.47	223.52		

EXPERIMENT 6

List position and serial order effects and word-pair presentation. Comprehension.

Experiment 6. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	F	p
Group	2	4479.73	2239.86	2.45	0.09
Error	56	51154.98	913.48		
List position	2	16043.67	8021.83	21.21	< 0.01
Group x List position	4	2507.55	626.88	1.65	0.16
Error	112	42357.09	378.18		

EXPERIMENT 7

Word frequency and list position. Generation.

Experiment 7. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	F	p
Group	1	41.68	41.68	0.01	0.91
Error	57	189023.75	3316.20		
List position	2	5712.31	2856.15	4.79	< 0.05
Group x List position	2	372.42	186.21	0.31	0.73
Error	114	67890.55	595.53		
Frequency	1	10661.84	10661.84	19.86	< 0.01
Group x Frequency	1	242.55	242.55	0.45	0.50
Error	57	30588.27	536.63		
List position x Frequency	2	9941.61	4970.80	11.68	< 0.01
Group x List position x Frequency	2	1919.91	959.95	2.25	0.10
Error	114	48484.86	425.30		

EXPERIMENT 8

Word frequency and list position. Comprehension.

Experiment 8. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	*F*	*p*
Group	1	1134.38	1134.38	0.64	0.42
Error	56	98726.26	1762.96		
List position	2	4311.76	2155.88	2.74	0.06
Group x List position	2	450.52	225.26	0.28	0.75
Error	112	87860.25	784.46		
Frequency	1	10166.17	10166.17	20.73	< 0.01
Group x Frequency	1	962.53	962.53	1.96	0.16
Error	56	27462.05	490.39		
List position x frequency	2	4474.25	2237.12	3.92	< 0.05
Group x List position x Frequency	2	177.51	88.75	0.15	0.85
Error	112	63798.57	569.63		

EXPERIMENT 9

The effects of word category on recall

Experiment 9. Summary of ANOVA results.

Source of variation	df	sum of squares	mean square	*F*	*p*
Group	1	1248.27	1248.27	0.47	0.49
Error	45	118570.87	2634.90		
Category	4	30018.50	7504.62	17. 68	< 0.01
Group x Category	4	3241.90	810.47	1.91	0.11
Error	80	76380.43	424.33		

EXPERIMENT 10

The effect of embedded words on memory processes

Experiment 10. Summary of ANOVA results.					
Source of variation	**df**	**sum of squares**	**mean square**	***F***	***p***
Test	1	6469.02	6869.02	4.16	< 0.05
Error	49	80872.15	1650.45		
Embedding	1	15176.61	15176.61	43.19	< 0.01
Test x Embedding	1	0.14	0.14	0.00	0.98
Error	49	17215.53	351.33		
Category	1	9038.79	9038.79	24.48	< 0.01
Test x Category	1	42.71	42.71	0.11	0.73
Error	49	18090.61	369.19		
Embedding x Category	1	8.47	8.47	0.02	0.88
Test x Embedding x Category	1	32.00	32.00	0.07	0.78
Error	49	20360.15	415.51		