In Search of a Theory of List Memory: An Inquiry into the Effects of Long-Term Information on Short-Term Recall

Andrew Francis Turvey

Doctor of Philosophy

ASTON UNIVERSITY

May 2003

Thesis Summary

Over recent years, evidence has been accumulating in favour of the importance of long-term information as a variable which can affect the success of short-term recall. Lexicality, word frequency, imagery and meaning have all been shown to augment short term recall performance. Two competing theories as to the causes of this long-term memory influence are outlined and tested in this thesis. The first approach is the order-encoding account, which ascribes the effect to the usage of resources at encoding, hypothesising that word lists which require less effort to process will benefit from increased levels of order encoding, in turn enhancing recall success. The alternative view, trace redintegration theory, suggests that order is automatically encoded phonologically, and that long-term information can only influence the interpretation of the resultant memory trace. The free recall experiments reported here attempted to determine the importance of order encoding as a facilitatory framework and to determine the locus of the effects of long-term information in free recall.

Experiments 1 and 2 examined the effects of word frequency and semantic categorisation over a filled delay, and experiments 3 and 4 did the same for immediate recall. Free recall was improved by both long-term factors tested. Order information was not used over a short filled delay, but was evident in immediate recall. Furthermore, it was found that both long-term factors increased the amount of order information retained. Experiment 5 induced an order encoding effect over a filled delay, leaving a picture of short-term processes which are closely associated with long-term processes, and which fit conceptions of short-term memory being part of language processes rather better than either the encoding or the retrieval-based models.

Experiments 6 and 7 aimed to determine to what extent phonological processes were responsible for the pattern of results observed. Articulatory suppression affected the encoding of order information where speech rate had no direct influence, suggesting that it is ease of lexical access which is the most important factor in the influence of long-term memory on immediate recall tasks.

The evidence presented in this thesis does not offer complete support for either the retrieval-based account or the order encoding account of long-term influence. Instead, the evidence sits best with models that are based upon language-processing. The path urged for future research is to find ways in which this diffuse model can be better specified, and which can take account of the versatility of the human brain.

Acknowledgements

I am very much indebted to my family for moral and financial support throughout the period of this research. I would also like to thank the postgraduates at Aston, especially Miriam, for making my time here an enjoyable one. The input from my supervisors, Fiona Fylan and Andy Bridges, is greatly appreciated, in particular, their encouragement and comments on the final draft of this thesis. Thanks also go to Niteen Mulji and Ian Fawcett for prompt and appropriate technical assistance, and to Lindsay Melling for sharing some of the tedium of data collection. Finally, I am grateful to the British Psychological Society for funding a secondment to the Parliamentary Office of Science and Technology during my final year of study.

Table of Contents

List of Figures and Tables

Chapter One: Verbal Short Term Memory- Structure and Function

It has been suggested that more experiments have been run on list memory than on any other paradigm in psychology (Anderson, Bothell, Lebiere & Matessa, 1998). Consequently, there are innumerable competing models of human memory which attempt to explain the patterns within the vast body of data generated by these experiments. Until recently, much of the focus on short-term memory has been related to the working memory model of Baddeley and his associates (Baddeley, 1986; Baddeley & Hitch, 1974; Salamé & Baddeley, 1982), in which the fundamental aspect of short-term memory has been characterised by a modular articulatory loop, dependent upon a passive store of phonological codes for items and theoretically impenetrable by long-term information (Glanzer, 1972; Baddeley, 1986). Indeed, much research carried out in the last twenty years provided evidence that many factors influencing recall, such as word length and language, have a phonological basis (Salamé & Baddeley, 1982; Baddeley, Thompson & Buchanan, 1975; Hoosain & Salili, 1988; Baddeley, Lewis & Vallar, 1984). Recently, this assumption has been challenged, and evidence is accumulating to the effect that long-term information such as meaning and lexicality can influence the processes of short-term recall (Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart, 1997; Poirier & St-Aubin, 1995, 1996; St-Aubin & Poirier, 1999a, b).

The definition of information as used in this thesis is, simply put, *that which is stored in the brain*. In human memory systems, any data we have access to, whether it be motor programs, knowledge about faces or of language is a form of information. This kind of storage of knowledge is exemplified in verbal memory models by the storage of meanings of words, representations of the frequency with which the word has previously been experienced, measures of the familiarity of that word, and a wealth of other, related information. A number of levels of analysis have been applied to the problem of how this information is stored from neural networks (Humphreys, Bain & Pike, 1989; McClelland, 1998), through cognitive models (Baddeley & Hitch, 1974; Cowan, 1988) to chemical/ electrophysiological modelling (e.g. Hebb, 1949; Rugg, Cox, Doyle & Wells, 1995). The reported effects of many different factors on the processes involved in the wide spectrum of human memory are manifold, but two recent sets of findings are of particular interest in list memory research. Firstly, Delosh & McDaniel (1996) have noted the effects of word frequency on memory for order, and secondly Poirier & St-Aubin (1995) have explored similar effects with the semantic categorisation of word lists. These effects have been discussed as having similar characteristics and as perhaps depending on similar neural functions (Walker & Hulme, 1999; Nairne & Neuman, 1993).

The aim of this thesis is to explore and extend these recent findings. To this end, the present chapter will outline a brief history of memory research and will then explain some of the implications of recent findings on the validity of some of the better-specified models. This will generate a number of research questions which will form the substance of chapters two and three.

Early Short-Term Memory Research

Since the time of William James (1890) and Francis Galton (1883), short term memory has been characterised as a temporary store in which items can be consciously held and manipulated. This was always assumed to be distinct from long-term memory. Indeed, James refers to primary (short-term) memory as the "psychological present" and to secondary (long-term) memory as the "psychological past". Evidence for a distinction between the two types of memory was a common experimental finding. For example, Brown (1958) and Peterson & Peterson (1959) found that memory for items in the short-term was drastically impaired if participants were compelled to count backwards in threes (the *Brown-Peterson* task) or were asked to perform some other phonologically-based task before attempting recall.

Findings such as these were responsible for some of the first testable models of human memory. Broadbent (1958) and Atkinson & Shiffrin (1968) used data from studies such as the Brown-Peterson paradigm to assert that short-term memory was a temporary storage buffer in which information could decay rapidly. If the information was held or repeated for long enough, it could be processed and then sent for storage in long-term memory. Others (for example, Melton, 1963) ascribed forgetting to interference amongst items in memory, and suggested that this principle could account for both short-term and long-term forgetting. Hence, the question of whether short-term and long-term memories were stored by separate processes or by a monistic mechanism began to gain significance.

Melton's (1963) model was based upon the notion that since interference is the common cause of forgetting in short- and long-term memory, the two types of memory should behave in a similar way under a variety of measures. This was demonstrated in an implicit learning study reported by Hebb (1961). Participants were given a task in which pseudo-random sequences of numbers were to be remembered. However, on every third trial, the same number sequence was implicitly presented, and participants showed evidence of an improvement on these trials over the truly random ones. Thus it was demonstrated that long-term learning had an impact on short-term memory tasks.

Miller (1956) reported a related effect in his analysis of 'chunking' – the combining of smaller parts of information into larger wholes, and the resultant increase in the overall informative value of information which can be held in short-term recall studies. When items are arranged into meaningful chunks, rather than random collections of items, it is much easier to recall them: the letters ODG TCA LWO are harder to recall than the grouping DOG CAT OWL.

The counter-argument to the unitary account employed by those researchers favouring the dichotomous view of memory was to make a distinction between memory tasks and memory systems. Waugh & Norman (1965; see also Baddeley, 1997) argued that certain short-term tasks relied both on primary and secondary memory, whereas longer-term tasks were likely to rely mainly upon secondary memory, due to the rapid decay of primary memory information.

This argument embodied the main thrust of research throughout the 1960s and 1970s. As more and more evidence was accumulated, theorists generally began to accept the view that dual-process descriptions of memory certainly had theoretical use. Baddeley (1997) outlines four different approaches which support the dual-process models.

Firstly, he cites a collection of studies which demonstrate a number of different components within a task such as free recall. A classic finding in free recall tasks is that of the recency effect. When orally-presented words are allowed to be recalled in any order, the most recent items are best remembered and are often recalled first (Postman & Phillips, 1965; Glanzer & Cunitz, 1966). After a filled delay[1], this advantage is greatly reduced, despite the persistence of a primacy effect (good recall for words near the start of a list) in both cases.

Primacy and recency are concepts relating to the serial position of stimulus words within a list. Primacy refers to recall for words presented near the start of the list: traditionally these are recalled better than words in the middle. The effect is ascribed to the long-term component of memory (Baddeley & Hitch, 1974). The recency effect refers to the most recently-presented items in a list. These items are purported to be recalled well because the traces created in the participants mind on presentation are still active when it comes to recall those items. A filled delay disrupts this process as it is hypothesised to

---

[1] A filled delay paradigm involves leaving a gap between the presentation of the last stimulus item and the beginning of the recall phase, as with an unfilled delay. However, in the case of a filled delay, the participant is required to perform another task before recall commences. Often, this involves simple, but attentionally demanding tasks such as counting backwards in threes from a high number.

create interference in the conscious part of short-term memory (Peterson & Peterson, 1959).

Further evidence for dual-process models was cited by Glanzer (1972), who found that certain variables, such as presentation rate and word familiarity only affected the recency part of the curve, leaving the primacy effect intact. However, as discussed later on in this chapter, recent findings have suggested that some variables which were thought to affect only long-term memory do indeed influence short-term memory, suggesting that short-term processes are not as functionally independent of long-term memory as was first thought.

Baddeley's (1997) second argument against memory being a monistic structure is that short-term recall appears to have different capacity limitations to long-term recall. Typically, as stated by Miller's (1956) classic paper, the capacity of the short-term store is seven, plus or minus two items. The capacity of long-term memory is all but infinite in comparison. Another major difference between the two types of recall is that short-term recall appears to require attention, whereas long-term recall is sometimes automatic (Cowan, 1997). For example, when a familiar word is presented to us, we can access the word's meaning from memory with very little effort, but attempting to remember a new telephone number for a few seconds requires a great deal more effort.

These distinctions were fortified by analogy with the burgeoning field of computing in the early 1970s. Computers tend also to have two types of memory: RAM, which at the time was both fast-acting and of limited capacity, and a disk-based storage system which was slower to access, but had much greater capacity for storage. These two types of technological hardware have ever since been compared to short and long-term memory structures in the human brain, and the field has been used as a conceptual framework for memory theories ever since (e.g. Atkinson & Shiffrin, 1968). Such an approach has led to a focussing of the debate upon the issue of storage of memories, rather than the processes which occur during encoding and recall.

However, findings similar to those of Waugh (1970), that recent items were recalled faster than initial items on a list, are not conclusive proof that two different memory mechanisms are in play. Tellingly, Cowan (1997) extends the computer analogy further. He notes that it is common practice in the normal functioning of computer systems for the employment of "virtual memory", a process by which part of the long-term storage device can be made to function as if it were fast-acting RAM. This extension of the analogy outlines the possibility that any distinctions between short and long-term components of memory tasks may not be due to differences in fundamental architectures, but more to do with the way that the architecture is behaving.

The basis of Baddeley's third argument for the separation of memory mechanisms into short-and long-term stores is the neuropsychological evidence for a dissociation between short- and long-term processes. Milner (1966) supplied evidence of a patient, HM, who had undergone drastic temporal lobectomy in order to treat epilepsy. He was left with a severely defective long-term memory. HM was able to demonstrate a normal immediate memory span, but was unable to form new memories such as recognising people or remembering what he'd had for breakfast. Baddeley & Warrington (1970) tested the ability of amnesiac patients to recall word lists, and found that they performed very well on an immediate free recall task, and very poorly on a delayed recall task. Primacy effects were reduced and recency effects were normal, if not improved by their condition.

Shallice & Warrington (1970) provided evidence of the opposite problem. They reported a patient, KF, whose short-term memory was extremely limited, but whose long-term abilities generally seemed normal. They found that KF's memory span performance was limited to two or three numbers, whereas for amnesiacs like HM, memory span was normal. This distinction was supported by a free recall experiment in which KF's recall curve showed a strong primacy effect and a disrupted recency effect (for a more complete review of neuropsychological data, see Ellis & Young, 1988; Shallice, 1988). This type of double dissociation is very persuasive in determining the existence of two separate systems (but see Shallice, 1988, for a word of warning on such inferences).

The final argument Baddeley cites against a unitary memory system is that of the different types of encoding apparently used for each of the tasks. Immediate recall tasks seemed to be governed by phonological encoding of items, with phonologically similar items in a list being prone to more errors at recall (Conrad, 1964; Conrad & Hull, 1964). Wicklegren (1965, 1977a) deemed phonological errors in immediate recall to be intimately linked with the processing of the order of the items in memory. He supposed that phonologically unique information allowed accurate associative representation of items, and therefore allowed subjects to use associative cues at recall in addition to pure item information. For example, in the syllabic sequence, NA, FA, TA, if one recalls information pertaining to the 'a' sound, there is no unique association to prompt the next syllable. If one considers the sequence, NA, FO, TI, vowel sounds are uniquely associated with a particular consonant, and order information is inherently available. The theme of order retention and its importance in recall will be a recurring theme of this thesis.

On the other hand, Baddeley (1966a, b) noticed that although phonological coding was the determining factor in immediate retention, long-term recall depended more upon semantic encoding. He found that whilst errors in phonologically similar lists were reduced with delay, semantic errors increased over longer delays. Baddeley (1966a) presented his participants with three sets of word lists. One set was comprised of phonologically related, but semantically unrelated words (e.g. *man, mad, cap, can, map*), one of phonologically unrelated and semantically related words (e.g. *big, huge, tall, long...*) and one set of words which were unrelated either semantically or phonologically. Errors were counted if participants failed to recall an item in the correct place, irrespective of whether or not they recalled it in a different serial position. It was found that under an immediate recall paradigm[2], there were more errors for phonologically related word lists than for either of the other two. Conversely, Baddeley (1966b) found that using similar lists over a filled delay resulted in more errors for the semantically related lists than for either of the other two lists.

---

[2] i.e. where participants were required to recall words immediately after presentation of stimulus items had ended

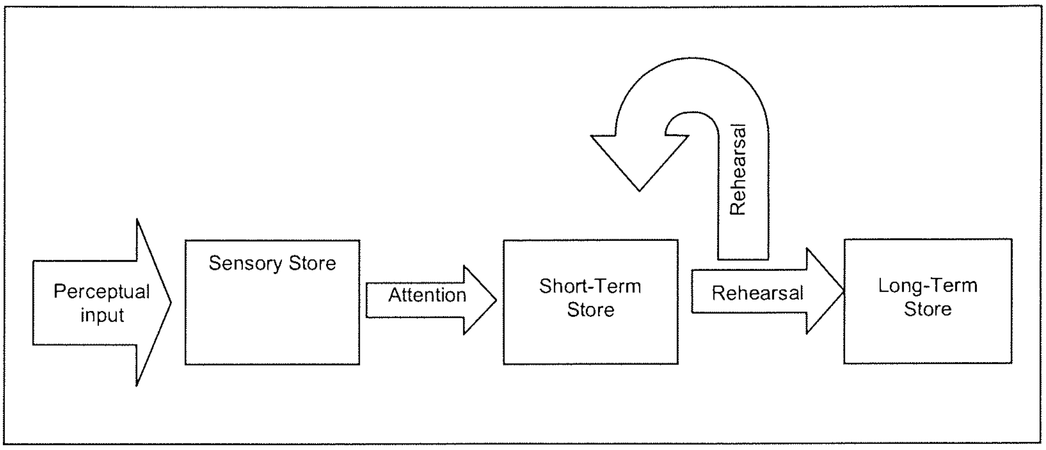
There were some concerns about the nature of these tests, for example that Baddeley's recall criteria depended mainly upon an order-only measure of retention, in which all items were presented at recall and participants were asked merely to reconstruct their order. This measure cannot be used to inform the debate on the effects of semantic factors on item recall. Despite this, results found by other researchers did suggest that for item encoding over filled delays, irrespective of the order of items, semantic representation was the norm (e.g. Sachs, 1967).

Models of Memory

The Modal Model

The first model to incorporate most of the above findings was the modal model (Broadbent, 1958; Atkinson & Shiffrin, 1968). It consisted of three separate stores and centred on a short-term store which functioned as a conscious workspace. Input from the environment was entered through very short-term sensory stores (see Sperling, 1960), was then processed and encoded by the short-term store, and after a degree of rehearsal, information entered the long-term store (see figure 1.1). It was supposed that without rehearsal, information was unlikely to enter the long-term store (Rundus & Atkinson, 1970). Indeed, Atkinson and Shiffrin (1968) specifically predicted that the more times an item was rehearsed in the short-term store, the greater the chance of transfer to long-term memory.

Figure 1.1 The Modal Model (Adapted from Atkinson & Shiffrin, 1968)



However, there were serious problems with the modal model. The first concern was that it was too simplistic. Theorists assumed that each subdivision of the memory system operated in a consistent, uniform manner. This was not necessarily the case. For example, Shepard & Metzler (1971) reported evidence of short-term encoding of mental rotation in addition to linguistic encoding. Warrington & Shallice, (1972, Shallice & Warrington, 1974) discovered that KF's short-term deficit was centred on the auditory modality, and furthermore, that only language sounds were affected, meaning that the problem was not one of echoic memory or of a hearing impairment. Visual stimuli and non-verbal auditory stimuli were recalled normally, suggesting that information from different modalities was processed in different ways by short-term memory.

Furthermore, according to the modal model, short term processing was absolutely necessary for information to enter long-term memory and patients such as KF should not have been able to efficiently form new long-term memories. This was not the case (Basso, Spinnler, Vallar & Zanobio, 1982). HM, the patient with a long-term memory deficit, was also found to be able to store some new memories. Although he had difficulty in recalling episodic information, meaning autobiographical occurrences, he was able to learn new motor tasks, such as tracing a shape in a mirror (Squire & Cohen, 1984). Other studies have shown that some amnesic patients are able to form some non-motor memories also

(Squire & Zola, 1998). The concept of long-term memory as a homogeneous store was apparently an oversimplification too.

More evidence that rehearsal did not directly lead to long-term storage of information came from Tulving (1966), who showed that repeatedly exposing participants to words did not result in improved learning of those words in a later test. The amount of time an item stays in active rehearsal does not predict success in recall (Rundus, 1977; Craik & Watkins, 1973). Nickerson & Adams (1979) demonstrated that people have poor memory for the most common everyday objects. They found that very few Americans could identify the twelve most salient features of a one-cent piece, despite the frequency of their exposure to the object. On reflection, it seems improbable that we necessarily need to rehearse to be able to remember in the long-term. When we read text, we do so fluently and at a fast rate, without having to repeat each word as it occurs. Despite this, we are still able to retain information learnt through reading.

The double dissociation between primacy and recency effects mentioned earlier as evidence for the existence of two separate stores was also found to be too simplistic. Bjork & Whitten (1974) managed to induce a recency effect in long-term memory simply by increasing the length of time between the presentation of successive items in the list. Crowder & Neath (1991) describe this effect with an analogy. It is as if items in a list are similar to telegraph poles standing in a long line. If one stands at the end of the line, one can clearly distinguish between the nearest poles (or the most recent memories), but as they get further and further away, it is harder to separate items. If the viewer himself moves further away from the poles, even the nearest are hard to distinguish, unless the gaps between the poles are widened. Thus, all memories, whether short- or long-term, can be described by the same mechanism, that of distinguishing between items on the basis of temporal distinctiveness.

Levels of Processing Theory

In the light of criticism of the two-store model, some researchers turned their attention to the processes of memory, rather than the architecture. Craik & Lockhart (1972) proposed that the type of processing applied to an item would influence how well it could be remembered. Participants were given lists of words to read, without being aware of a subsequent memory task. They were told to focus on particular aspects of the words which were presented to them. If participants focussed on surface features of words, such as the colour of the ink they were written in, or phonemes which they contained, then recall of those words was poor. If people were asked to process syntactic information about a word, then recall was slightly better than for words processed by surface features. The best levels of recall were obtained by words which had been processed semantically by participants. This finding was replicated in a number of studies of the period (Craik & Lockhart, 1972; Hyde & Jenkins, 1973; Craik & Tulving, 1975).

Many theorists took these findings to indicate that a distinction between short and long-term memory was not necessary (e.g. Crowder, 1982; Postman, 1975). Success of learning could be described in a simple and elegant way which depended upon the type of processing applied to a word at encoding. The depth-of-processing approach is also responsible for marking the need to account for the elaboration of items (i.e. how many different ways a single item is processed) as a factor in recall success.

However, the depth of processing approach is not without its problems. The most difficult problem is that there is no independent measure of 'depth of processing'. For example, some researchers claimed that processing a word's frequency was deep processing, whilst others classified this as a surface feature. Eysenck (1978) warned that the description of depth was in danger of becoming a circular argument, with better recall explained by 'depth-of-processing', and depth defined by better levels of recall. In this light, it can be seen that the theory merely described what was happening in memory, rather than explaining the processes of memory.

In sum, this approach signalled the importance of accounting for the ways in which items are encoded for recall. However, the theory does neglect the various factors affecting the retrieval phase of recall (for example, participants seem to be able to encode information specifically relevant to the type of test they are about to receive; Morris, Bransford & Franks, 1977). Baddeley (1997) notes that the focus of Craik & Lockhart's theory was in reality upon different ways of encoding information for long-term retention, ignoring immediate recall to some extent. A different model of short-term memory surfaced which addressed this issue.

The Working Memory Model

Baddeley & Hitch (1974) emphasised the need for the replacement of the concept of a simple short-term store with that of an active, working memory. They suggested that working memory consisted of a modality-free central executive, responsible for attentional factors, a visuo-spatial sketch pad for visual and spatial coding, and a phonological loop, which was responsible for verbal material. This conception had the benefit not only of being able to explain some of the previous findings (such as emphasis on phonological coding in short-term recall), but also to aid explanation of cognitive tasks such as language acquisition, verbal reasoning and arithmetic (Baddeley, 1997).

Evidence in favour of the working memory model was obtained through an experiment in which rehearsal of a list of digits was carried out simultaneously with verbal reasoning, language comprehension and free recall tasks (Baddeley & Hitch, 1974). In all cases, performance of the task was impaired further when the number of digits retained increased, as would be expected when overloading a system which is responsible for the conscious manipulation of verbal material.

Much of the experimentation in short-term recall since the introduction of the working memory model has concentrated upon the phonological loop. This is largely because it is the most accessible part of the model, and because verbal material is very convenient to use. The central executive is still largely shrouded in mystery, and remains little more than a black box 25 years after its appearance. However the advent of brain-imaging techniques has allowed researchers to begin to cast some light on executive processes. For example, Goldman-Rakic (1995) has used this approach to determine that specific areas of the prefrontal cortex are involved in different aspects of executive function. She suggests that ventral regions handle object-based executive tasks, and that dorsal regions process spatially-based tasks.

In addition, several cognitive models have been developed, not least the Supervisory Activating System (SAS) model of Norman & Shallice (1986). This account offers an explanation for the differences between automatic and willed action. The willed component is very similar to the central executive module of Baddeley's (1986) model, in that it organises the priority of tasks and therefore the order in which they are carried out. This SAS is hypothesised to be something of a conflict-resolver in working memory, and supposedly functions by favourably weighting the importance of one from a number of competing behavioural possibilities. The model has successfully been applied to problem solving in chess (Robbins, Anderson, Barker, Bradley *et al,* 1996) and language comprehension (see Baddeley, 1997). However, most of the focus on working memory has been on the slave systems rather than the central executive itself. In verbal memory research, it has been the phonological loop which has come under the most scrutiny and so for much of the remainder of this thesis, it is to this slave system to which the discussion turns.

A brief discussion of evidence for the phonological loop will help to characterise its properties. Firstly, it appears that the phonological loop is sensitive to the order of items, and that this has a basis in language comprehension. Baddeley & Lewis (1981) presented participants with otherwise meaningful sentences in which syntactic anomalies or semantic anomalies were present. Syntactic errors involved switching the order of two words, whereas semantically abnormal sentences had a word replaced by an inappropriate substitute. When asked to assess whether sentences were meaningful whilst performing an articulatory suppression task (participants were forced to articulate unrelated words concurrently with performance on a memory task), syntactic abnormalities yielded much worse performance than semantic errors. Assuming that articulatory suppression does indeed disrupt phonological encoding of items, this supports the notion that the phonological loop is responsible for encoding the order of items in addition to pure item information.

A related piece of evidence comes from the phonological similarity effect mentioned above. If similar items are stored in a phonological store, those items will have similar codes, and the process of recalling those items will be made more difficult as discrimination amongst those traces will be harder to achieve. Wicklegren (1977a) suggests that this discrimination will affect the order of items recalled much more than the item information alone.

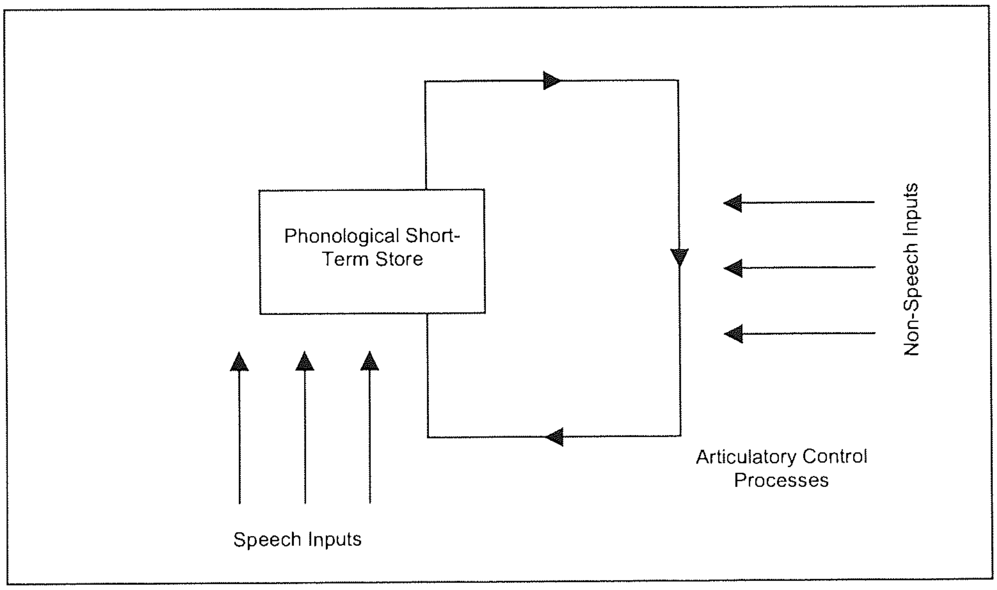
Another source of evidence for the existence of a phonological loop in human memory is from the irrelevant speech effect. When non-verbal background noise is playing whilst participants attempt to recall information, there is no deleterious effect, but when verbal sounds are played, even if the syllables do not make up proper words, recall is disrupted (Salamé & Baddeley, 1982; Macken & Jones, 1995). The reasoning behind this is similar to the articulatory suppression effect, which is that the loop is occupied with sounds other than those generated by interpreting words in a memory task and as a result, performance on that memory task suffers.

A final, convincing piece of evidence for the phonological loop is the word length effect. The longer the words within a list, the less likely that list is to be successfully recalled (Baddeley, Thompson & Buchanan, 1975). However, this effect appears to be dependent on the rate of articulation for the words, rather than their written length. The authors found that an individual's reading rate had an effect on the number of words correctly recalled also. It transpires that the number of words one can recall in immediate memory is equivalent to the number of words one can articulate in approximately two seconds. Such a figure is reminiscent of Miller's (1956) seven-item limit to short-term memory.

A complication of the model is that the word length effect can be disrupted by articulatory suppression for written words, but not for spoken words (Baddeley, Thompson &

Buchanan, 1975). This made it necessary to adapt the model to comprise two components: a passive phonological store, which can hold information for around two seconds, and an articulatory control process, which refreshes information in the store by means of subvocal rehearsal (see Figure 1.2). The articulatory control process is also responsible for the translation of written stimuli into phonological codes, which accounts for the disruption of the written word length effect under articulatory suppression, but allows spoken stimuli direct access to the phonological store.

Figure 1.2. The Phonological Loop as Passive Phonological Store and Active Articulatory Control



As a result of the simplicity of the phonological loop model and the ease with which experiments can be designed to probe its possibilities, Baddeley & Hitch's (1974) working memory model has proved a useful theoretical tool over the past two decades. However, one must raise a caveat at this point. Experiments designed to specifically test the phonological loop model generally involve paradigms in which strict serial ordered recall of words is required, often at the expense of performing experiments with a wider, more naturalistic scope, such as free recall paradigms in which no artificial constraints are imposed upon the participant. It is the aim of this thesis in part to extend some of the findings under discussion here and below to a more liberal paradigm in order to assess whether the findings from serial recall studies are applicable to the kinds of strategies employed naturally by participants in free recall.

Problems with the Phonological Loop Account

Despite all of the evidence presented above, the phonological loop account does not explain all of the available data. Some have questioned the validity of the evidence for the model on a number of wide-ranging levels (e.g. Hulme *et al*, 1997; Crowder, 1989, 1993; Jones & Macken, 1993; Walker, 1999). One criticism is that the number of items which are supposedly held in the phonological loop does not represent the way in which items are fundamentally processed in memory, but rather relies on a 'trick' (Crowder, 1989). Crowder carried out an experiment in 1969 which involved the presentation of very long lists of words to participants, during which the experimenter arbitrarily stopped the participant and asked for immediate backward recall. Only two or three items were usually recalled. Crowder suggests that this result and those of articulatory suppression studies (which yield similar levels of immediate item recall) point to those items simply being those in consciousness, and not in a distinct, practical short-term store.

The distinction between a short-term store and consciousness is a subtle one. Unattended information can still be processed and acted upon even if it is not in a position to be recalled. Cherry's famous (1953) cocktail party experiment demonstrated this point. Participants were asked to listen to a recorded voice in one ear whilst trying to ignore a similar recording in the other. Even though participants were not able to overtly recall information which had been presented in the unattended channel, salient cues, such as the participant's name or emotionally loaded words could cause them to switch attention to the other channel. This principle is most clearly applied to memory research in Cowan's (1988) model of memory and attention. He likens conscious awareness to a spotlight focusing on the various cognitive activities being carried out in the brain. Function, such as recall, is represented by activity and attention marked by the intensity of activity in this model. Once the activity in the relevant cell assemblies reach a certain threshold, the participant becomes consciously aware of them. Hence, Crowder's (1989) "trick", mentioned above, refers to the possibility that recently presented items may be more active in other areas, such as in a lexical store, rather than necessarily being representative of phonological loop function.

Other criticisms of the phonological loop model include the problem of order encoding. If people reactivate items through rehearsal during presentation, then it is not clear how order can be effectively encoded. This is true from an activation-based account of order encoding, in which the most active item equates to the most recently-presented item (see Cowan, 1997), and it is also true from an associational point of view (e.g. Wicklegren, 1977a). Old items would be interpolated with new items, and unique associations between items in the correct order would not occur without filtering out unintended associations between currently rehearsed items and novel items (Tzeng & Cotton, 1980).

Long-term Memory Influences on Short-Term Recall

One extremely important finding which contradicts the phonological loop hypothesis is that information from long-term memory can affect the way in which information is recalled in the short-term. If information were simply entered into the phonological loop and then emptied out of it at recall, one would expect only the phonological features of those words to have any impact upon their storage and recall (this is the hypothesised basis of the word length, phonological similarity and speech rate effects). Since it is only the phonological stream which matters in Baddeley & Hitch's (1974) model, long-term factors such as lexicality or the meaning of the word should neither be accessed nor relevant to the processing of the memory (see Baddeley, 1966a).

However, a number of studies have recently shown such an effect on both item and order information. The paradigm of choice for studying phonological loop activity is that of immediate serial recall, in which it is supposed that participants use only phonological codes to encode and retrieve information. Monists, such as Crowder (1982) would argue that other mechanisms come into play during this process, a theory which is discussed below.

Firstly, in serial recall studies, Hulme, Maughan & Brown (1991) and Hulme, Roodenrys, Brown & Mercer (1995) have demonstrated an effect of lexicality on short-term recall (see also Besner & Davelaar, 1982). Words are recalled better than non-words and familiar words better than unfamiliar, irrespective of the speech rate of the items used in either condition. This effect on item recall extends to words of differing frequency, with high-frequency words being recalled better than low (Hulme *et al*, 1997) and in an order more closely approximating the order of presentation (Delosh & McDaniel, 1996).

Semantic information has also been found to influence immediate serial recall. Baddeley (1966a) found a detrimental effect of semantic categorisation on memory span, but as mentioned previously, his method of scoring for serial recall was unorthodox. Huttenlocher & Newcombe (1976) found a facilitatory effect for semantic categorisation on a memory span task, and Tehan & Humphreys (1988) found that function words (such as *but, if, and*) were not recalled as well as content words. Bourassa & Besner (1994) replicated this finding, but also found that when the words were matched for imagability (a measure of how easy it is for participants to form an image of the word), the effect disappeared. From this, they inferred that the basis of the function/ content effect was in the semantic coding of the word, rather than in a syntactic code. Walker & Hulme (1999) found a similar advantage for concrete words (those with a specific definition, like *tree* or *house*) over abstract words (like *justice* or *mood*). Poirier & St-Aubin (1995) have also reported a semantic effect on item information in serial recall. Semantically categorised lists were better remembered than semantically unrelated lists, but this effect appeared to be independent of phonological loop activity: categorised lists retained their advantage even when articulatory suppression was required. The Poirier & St-Aubin (1995) experiment suggests not only that the serial recall paradigm does not load solely on the phonological loop component of memory, but also adds weight to the hypothesis that memory processing depends on an interactive network of specialised processes, all or some of which may be used, depending on task requirements (see Cowan, 1997).

The precise nature of semantic effects such as this is hard to determine. It could be that long-term semantic information is influencing either the encoding or recall of items, or indeed that the effect is somehow occurring in Baddeley's other mooted memory subsystem, the visuo-spatial sketchpad. Either way, it is certain that care should be taken when designing stimuli for use in such experiments, as the confluence of a number of different factors in word perception, recall and speech can muddy the waters significantly when it comes to explaining the causes of behaviour.

Cowan's Model of Attention and Memory

Cowan's (1988) model of memory and attention was developed as an alternative to the working memory account. He argued that previous definitions of short-term memory were vague in general, with STM defined as either activated elements from long-term memory or elements that are currently receiving attention. However, he noted the existence of implicit memory effects which seemed to be an activation of memory outside awareness. For example, Tulving, Schacter & Stark (1982) demonstrated that participants who had been exposed to words in one task were then more likely to choose those target words in a word fragment completion task, even though other, perhaps more obvious possibilities for completion existed. Even more interesting is their further finding that there was no relationship between this kind of performance on the fragment completion task, and a later test of recognition memory for the target words. This experiment showed that conscious memory of an episodic event such as the presentation of a word is not always necessary for that event to affect memory performance.

Cowan (1988; 1997) used evidence such as this to specify a new way of conceptualising the human memory system. He saw the focus of attention as a subset of short-term memory, and STM as an activated subset of LTM. To some extent, this conception is very similar to Tulving's (1985) classification of memory into anoetic, noetic and autonoetic components. However, whilst Tulving used this classification to sub-divide long-term memory into non-conscious procedural memories, conscious semantic memories and self-conscious episodic memories, Cowan's approach sought to disambiguate short and long-term processes. As evidence for STM being the activated portion of long-term memory, he cites the simple fact that people can remember more words than non-words (e.g. Hulme, Maughan & Brown, 1991), suggesting an obvious effect of long-term memory on short-term recall behaviour.

This model accounts adequately for phenomena such as the word length effect, in which longer words are recalled less well. Baddeley, Thompson & Buchanan (1975) ascribed this effect to the possibility that fewer long words could be maintained in the two seconds of phonological loop available than short words, and therefore that short words were at an advantage. However, Cowan gives evidence that this phenomenon can be explained in terms of output speed for individual items. Cowan, Day, Saults *et al* (1992) reasoned that items in the phonological loop decay over time, and that therefore longer words, which took longer to pronounce during recall, may cause items still remaining in the phonological loop awaiting output to decay more than if a short word was being pronounced. In this way, words output first should have the most impact on recall performance, as the items recalled last would not be delaying any remaining words. In order to assess this possibility, Cowan *et al* (1992) presented participants with five-word lists for which the lengths of the first two and the last two items were manipulated. In one experiment, recall was signalled with a post-list cue to be either forwards or backwards. Crucially, whereas in forwards recall, words in the first half of the list demonstrated the word-length effect but words at the end did not, in backwards recall, the opposite was true. Cowan *et al* had shown that it was words which were to be recalled first which caused the word length effect, and that therefore this effect was output-based, rather than retention based.

However, despite Cowan's model of memory offering a convincing explanation of some of the phenomena of short-term and long-term memory research, it remains a poorly specified model in terms of the specific predictions it makes. Additionally, "activation" is not accurately defined within the model, and therefore it remains difficult to assess empirically. Despite these criticisms, Cowan's (1988) model has provided a plausible alternative to the working memory model, and has provided a basis for similar and better-specified alternatives.

The Feature Model

The feature model (Nairne, 1990b; Neath & Nairne, 1995) accounts for short-term memory traces in terms of vectors of features. Each vector corresponds to a feature of a word, such as a syllable or part of a meaning, and each feature can either represent information about presented items, or information about stored words in LTM. Nairne's model has each item presented in a list overwriting some of the salient features of the previous item. If two items are very similar, unique features distinguishing the two will be more likely to be lost then if two items are very different. Thus, forgetting of information in this model is characterised by interference between items, rather than decay of items over time. If all the items in a list share similar characteristics, there is more chance that information will be lost, and that recall performance will worsen.

In this way, the feature model can account for the effects of phonological similarity mentioned previously. It also accounts for the recency effect, as the last few items in a list will have fewer of their features overwritten, where those near the start and in the middle will suffer from greater interference. The suffix effect can also be explained by this model. The suffix effect was studied by Crowder & Morton (1969), who found that presenting an irrelevant spoken stimulus after a target list interfered with recall, whereas the presentation of a tone did not, even when participants were told to ignore it. The Feature model explains this phenomenon through the automatic overwriting of the feature buffer by the extra syllable.

However, there are some effects that the feature model cannot explain. For example, it does not address recent developments in irrelevant speech research. The irrelevant speech effect is based on the finding that recall for a list of items is disrupted when a stream of irrelevant verbal material is concurrently presented to the participant at the time of list presentation (e.g. Jones & Macken, 1993). Non-verbal material was initially not found to have any similar effect, and Nairne's (1990) model could explain this pattern of results in a similar way to that described above for the suffix effect, namely that phonological features in the irrelevant stream could cause interference with items with similar features in the short term store. However, Macken & Jones (1995) later reported evidence that the irrelevant speech effect can still be induced by using stimulus items and irrelevant items which do not share phonological features. This is problematic for the feature model.

The feature model also has some difficulty in explaining other recent findings in the field of short-term memory research. The word frequency effect (e.g. Hulme et al, 1997) and semantic effects (e.g. St-Aubin & Poirier 1999*b*). Two models that may be able to explain effects such as this are discussed below.

The Trace Redintegration Model

The spiritual successor of working memory is the trace redintegration model proposed by Schweickert (1993; Hulme *et al*, 1997; see also Horowitz & Prytulak, 1969). It attempts to account for the effects of long-term memory on immediate recall by examining the processes occurring at retrieval (it is sometimes called the *retrieval-based hypothesis*). The idea is that list presentation creates a number of memory traces which are phonetically coded, as in the original phonological loop model. As the trial goes on, the traces gradually degrade, either through decay or through interference. When it comes to recalling those items, participants must interpret the degraded traces available to them. The original phonological loop model did not specify how this might happen, but Hulme *et al* (1997) suppose that phonological traces are compared to traces which exist in the mental lexicon. If a unique match is found, then that word can be selected for recall, but if no word matches, participants can only guess at best. Hulme *et al* (1995) provided evidence for this account.

They presented lists of nonwords to participants which were recalled poorly. However, after participants practised saying the words repeatedly, recall improved.

The reason that high frequency words demonstrate improved recall is explained by Hulme *et al* (1997). They propose that high frequency words have higher resting levels of activation in the lexicon (cf. Plaut, McClelland, Seidenberg & Patterson, 1996; Patterson & Hodges, 1992), and are therefore more easily accessed for comparison with the phonological trace. The explanation for the advantage of imagable words (Bourassa & Besner, 1994) and of concrete words (Walker & Hulme, 1999) is similar, assuming that such words also have higher resting levels of activation. The explanation for semantically categorised lists demonstrating an advantage over non-categorised lists (Poirier & St-Aubin, 1995; St-Aubin & Poirier, 1999b) presumably relies on a different process, whereby the activation of a word in the lexicon increases the activation of semantically related words through lateral connections (cf. Hinton & Shallice, 1991; Plaut & Shallice, 1993), increasing the chances that other semantically related words on the list will be matched through trace comparisons.

The trace redintegration account is a useful tool in examining LTM effects on short-term recall, but it remains incomplete. A number of specific criticisms can be levelled at the retrieval-based hypothesis. Firstly, it is based solely on the serial recall paradigm, with no evidence that this strategy is more generally used in more naturalistic free recall designs. If it is to be useful in effectively describing the way humans generally behave, the same processes of trace redintegration must also be shown to appear in immediate free recall. Also, it is entirely possible that the effects of frequency and semantic information have their basis in a more general, unified memory architecture, such as that hypothesised by Crowder (1993). If this is the case, then one would expect to be able to replicate these effects over longer delays, just as the recency effect was shown to extend for far longer if the conditions were right (Crowder & Neath, 1991). Some evidence to this effect will be discussed in the next section.

Another problem for the retrieval-based hypothesis is the way in which it accounts for order information. Some serial recall studies (e.g. Hulme *et al*, 1997) have found an effect of word frequency on the order of items recalled. Namely, high-frequency words appear to be better recalled in the middle serial positions of a list compared to low-frequency words. Similar studies on the role of semantic information have found no evidence of semantic processing directly affecting the order of items recalled (St-Aubin & Poirier, 1999a; Poirier & St-Aubin, 1995). The trace redintegration hypothesis does not suggest why this difference might be. However, one approach which can offer explanation for this phenomenon is that of analysing encoding processes.

The Order-Encoding Hypothesis

The order-encoding hypothesis (Nairne, Riegler & Serra, 1991; Serra & Nairne, 1993; Delosh & McDaniel, 1996) is based upon the assumption that order information acts as an extra cue to the recall of words on a list. The retention of order information from the presentation of a list of items has been a prime candidate as recall cue for a very long time (e.g. Ebbinghaus, 1885/1913). It appears to occur automatically (Deese, 1957; Toglia & Kimble, 1976), suggesting that latent order information will always have a significant influence on recall. The problem is that it has always been difficult to tease apart the separate influences of item recall and order information in a word list.

Item and order information must be distinct to at least a limited degree (see Crowder, 1979). It is possible to forget which order some past events occurred in, and equally possible to remember the order of temporarily forgotten events once reminded of them. Some theories (such as levels of processing theories) imply that item and order information are separately stored and accessed. Individual items are either processed deeply and well remembered, or not processed so deeply, and are less well remembered (Craik & Tulving, 1975). Order encoding is sometimes seen as an entirely separate process, which is secondary to item processing (see Greene, Thapar & Westerman, 1998).

Another approach to the item-order debate is linked to theories which focus upon how items are organised within memory. Item and order information can be seen as being of a similar type. As Crowder (1979) points out, in attempting to remember an item from a list, we are not testing our memory of what that item represents, but rather whether or not that item occurred on a particular list. Hence, item memory is merely a more coarse-grained version of order memory: judging position between lists, rather than between items. This approach fits well with emphasis on organisation, as it allows a distributed memory approach, in the manner of some of the monistic theories mentioned above. For example, TODAM (*Theory Of Distributed Associative Memory*) is able to code item and order information in the same convolved code, and the recall of one is a cue to the other (Lewandowsky & Murdock, 1989; Murdock, 1997). Whichever of these assumptions is correct, most theorists assume that retention of order information can enhance the recall of individual items in a list. The order encoding hypothesis specifies how this enhancement may occur.

It states that words presented to a reader or listener require a certain amount of resources to process. Some words require more cognitive resources than others to process- this can be seen as analogous to the lexical access latencies being different for (for example) high and low frequency words. Words which are accessed more easily promote the encoding of order information. Words which are uncommon or unusual are hypothesised to undergo less relational encoding. Delosh & McDaniel (1996) state that the reason for this is that...

> *"learner's resources are lured to processing and interpreting the individual and idiosyncratic features of the unusual items, leaving fewer resources (or reducing attention) for encoding order information"* (p1137, original parentheses).

As a result of impoverished order information for words which require more individual elaboration, overall recall for those lists is hypothesised to be reduced, since they do not benefit from the extra recall cues available to them from order information. Studies have supported this theory, with low frequency words suffering a disadvantage (Delosh & McDaniel, 1996), and generated words (partial words which must be completed by the participant at presentation) also showing less item and less order retention (Nairne *et al*, 1991). This effect of word frequency occurs over filled delays of 30 seconds, presumably removing the redintegration of phonological codes as a source of the effect, as the phonological element should have disappeared after such a delay. This is further evidence that the trace redintegration account is not a complete one.

One interesting prediction of this approach is that these effects should only occur for pure lists, in which every word is at the same level of advantage. In mixed lists of low and high frequency words, the order encoding hypothesis makes the prediction that low frequency words should gain an advantage in both item and order recall (Delosh & McDaniel, 1996). Item recall should be enhanced because the low-frequency words are elaborated more, and hence are processed more deeply, and order recall should be enhanced relative to pure lists, as the high frequency words in the list will encode order information relative to their low-frequency neighbours. As a result, low frequency words will gain a certain amount of relational information by being associated with high-frequency neighbours and high-frequency words will lose order encoding through their association with items which are not individually encoded for order.

Delosh & McDaniel (1996) ran a series of experiments which aimed to illustrate this effect. In their first experiment, eight words were presented for two seconds each, with an inter-stimulus interval of 500 msecs. Immediately after the presentation of the last word in the list, the participants were asked to solve maths problems for 30 seconds, after which they had to write down as many of the words they could remember from the list. Following this free-recall phase, there was an order-reconstruction phase in which all words were re-displayed on the screen in a random order. Participants were asked to reconstruct the original order of these items with unlimited time. Success on free recall and order reconstruction differed depending on the types of word list presented. Pure lists of high frequency words were recalled better than pure low frequency lists. Mixed lists of high and low frequency words were showed the hypothesised frequency effect, with low-frequency items enjoying an advantage over the high frequency items. When it came to analysing the order reconstruction task, recall of order information for pure high-frequency lists was best, and pure low-frequency lists showed the worst recall for order information. However, order reconstruction for the mixed lists was placed at an intermediate level. A similar pattern was observed when the amount of latent order information evident within the free recall output protocols was analysed using an index of correspondence between input and output orders.

This effect was repeated across a number of contexts and was shown to be independent of word imagability and concreteness. Delosh & McDaniel (1996) believed that high frequency words act as better cues than low-frequency words and that a greater proportion of high frequency words in a list will give more resources to process contextual order information for items within a list.

This pattern of results is shown to occur by Delosh & McDaniel (1996) and is entirely unpredicted by the trace redintegration model as it stands (Walker & Hulme, 1999). The frequency of a word should, according to the retrieval-based hypothesis, only affect the baseline activation of words in the lexicon. Since recall for low frequency words in one list is enhanced and for the same words in a different list is reduced, this casts further doubt on the applicability of trace redintegration theory to free recall paradigms.

Organisation of Memory

As mentioned before, levels of processing theory helped to shed light on memory theory by switching emphasis from an architectural view of memory to a process-based one. Recent research has not only focussed on the phonological loop and its behaviour, but on the organisation and elaboration of the memory trace as a predictor of successful recall (Belezza, Cheesman & Reddy, 1977; LeCompte & Watkins, 1993; Delosh & McDaniel, 1996; Greene, Thapar & Westerman, 1998). It is clear that any coherent model of short-term memory will necessarily have to incorporate item elaboration (e.g. depth of processing) and organisation (e.g. order encoding) effects in addition to determining the structure and behaviour of the phonological loop. One way in which convergent data may be sought is through some of the long-term memory effects described above. The processes of memory may depend upon mechanisms of linguistic processing (Crowder, 1993; Walker, 1999), and studying their effects on recall tasks may help to point to a way of interpreting the vast and often conflicting data on verbal short-term memory.

A good place to start this process would appear to be in assessing how long-term memory affects the encoding of order information across memory tasks. This approach has borne much fruit in the area of serial recall of lists of words, with lexicality (Besner & Davelaar, 1982; Brown & Hulme, 1992, 1995; Hulme, Maughan & Brown, 1991; Schweickert, Hayt, Hersberger & Guentert, 1996), word frequency (Engle, Nations & Cantor, 1990; Poirier & St-Aubin, 1996; Roodenrys, Hulme, Alban, Ellis & Brown, 1994; Watkins & Watkins, 1977) and imageability (Bourassa & Besner, 1994) all showing measurable influence on the encoding of information in serial recall, a paradigm which is heavily dependent on the retention of order information.

It was therefore deemed appropriate that this thesis should adopt a similar approach, and continue the process of gathering information relevant to the themes outlined briefly above. Of prime interest is the role of long-term factors, such as word meaning and word frequency, in the use of order information as a recall strategy. A brief review of previous findings for each of these factors will help to outline the aims of the research.

The Nature of Order Information Encoding

Serial position has an effect upon a participant's memory for items in a list. In a classic paper, Murdock (1962) describes the characteristic shape of the serial position effect with spoken stimuli. It is a bow-shaped curve, with a primacy effect at the start of the list, indicating better recall for list-initial items, and a recency effect, indicating good recall for latter items. To account for these findings, there are two main ways in which order could be encoded. One is to encode exact and absolute positions of items within lists, and the other is to associate each item to the next in a chain of items. Needless to say, there exist variations of these theories, too: for example, one variation of the chaining theory is that items are not only associated to immediate neighbours, but to many other elements in the list simultaneously (cf. Estes, 1972; Phillips & Singer, 1997). Evidence exists for associational encoding (e.g. Smith & Mynatt, 1982; Tzeng & Cotton, 1980; Wicklegren, 1977a) and for positional coding (Ebenholtz, 1963).

The distinction between these two forms of encoding is pertinent to the debate between short- and long-term storage. Associative evidence supports notions of human memory as a network of interactive components, each having an effect upon the other. Evidence of associational encoding would therefore harmonise with the accounts of Murdock (1997), Estes (1972), Crowder (1989) and Cowan (1997), mentioned above. This is not to say that evidence for positional coding undermines support for these theories, but it is perhaps easier to imagine item and order information as being separate in a positional encoding framework, with individual items being placed into 'slots' in memory, rather than being fundamentally associated with each other.

Initially, the weight of research focussed upon positional encoding theories. Ebenholtz (1963; see also Asch & Ebenholtz, 1962b) found that with non-words, temporal position was encoded, but information from associations between items was not used to recall lists. To demonstrate this, participants were given a serial anticipation task, in which they were presented with a series of nonsense syllables, one every two seconds, and were asked at each presentation to guess which came next. The same sequence of 12 syllables was repeated over and over again until the participant guessed all 12 correctly. In the experimental condition, the sequence started at a different position in each of 10 trials, but in the control condition, exactly the same sequence was presented each time. Ebenholtz found that participants took less time to learn the sequence in the control condition than they did for the experimental condition.

This experiment was to some extent biased, however, in that repeated presentations in the positional condition were unchanging, whereas for the associative condition, a different letter began each list, thus creating a new context for each new list, which may have been enough to produce the effects obtained. A second experiment to remedy this did produce similar results, but was subject to an entirely different bias: that only second-order associations were tested in the associative condition (i.e. A and C, C and E, and so on), as intervening words were randomly replaced during each trial. In the positional condition, every other position was randomly replaced. One would expect that first-order associations would be the strongest in associative chaining (Estes, 1972; Murdock, 1982, 1997), and that therefore participants in the positional condition were given more chance to succeed than those in the associative condition.

Effects of absolute position were also obtained by Winzenz & Johnson (1980). Participants were presented with word lists which either adhered to schema, or which were randomly organised. The predictable lists were organised such that individual words of a certain semantic category were always in the same position in the list, but the actual words from list to list were altered. In such lists, recall was better than for randomly organised lists. This effect was not present for lists in which the sequence of categorisation rotated, leading the authors to believe that positional encoding was taking place.

Support for associative chaining can be found in the literature also. Sheull & Keppel (1967) demonstrated that by learning pairs of words previous to learning lists, participants were better able to learn lists of words in which those pairings had occurred. Learning partial information about lists also aids recall of the list as a whole, and this learning is better still if associations between list items are concrete, rather than abstract (Smith & Mynatt, 1982).

Wicklegren (1967, 1977b) gives evidence that participants in immediate recall tasks tend not to code information positionally, but to group lists into sub-lists of start, middle and end groups. Within these groups, it appears that further classifications of initial, middle and final items are made. If short-term memory were structured in such a way as to allow chunks of information to be allocated to non-associative slots, “it is not at all obvious... why a subject would resort to such elaborate grouping schemes” (Wicklegren, 1977a; p223).

One variation of the associative coding theory is that of study-phase retrieval (Tzeng & Cotton, 1980; Nairne & Neumann, 1993). It is suggested that associations are not only made at encoding, but at rehearsal of items. Once an item is presented, it may remind the participant of an earlier item, and thus, whilst the latter item is being rehearsed, relative order is encoded too, as by definition, the item which the individual has been reminded of must have occurred before the new item. It follows from this that if two items are related in a long-term store of information prior to experimental presentation, then they will have a greater chance of cueing each other in a distributed semantic network (Anderson, 1983; Kruschke, 1992, 1996). Connectionist models are based upon out assumptions about the way the brain represents information. The basic principles of connectionism are that events in the real world are represented by patterns of neural activity. Inherent in each pattern is a set of instructions which are transmitted to the rest of the neural network and these instructions in turn produce a further set of patterns with their own inherent set of instructions (McClelland, 1998). In this way, it is easy to conceive that the representation of a single item may automatically generate representations of other, related items (e.g. Humphreys, Bain & Pike, 1989; Metcalfe, 1990; Murdock, 1997).

With respect to the actual processes of order encoding, Murdock and associates have outlined a model, called TODAM (*Theory Of Distributed Associative Memory*), in which item information and order information are both stored in the same memory trace (Murdock, 1982, 1997; Lewandowsky & Murdock, 1989; Lewandowsky & Li, 1994). Item information is encoded as a vector which is then added to a dynamically updated matrix representing the memory trace. In addition to each item being added, a representation of the associativity between subsequent items is also encoded (the outer product of auto-correlated item vectors). Recall processes are simulated by correlating a probe item with the memory vector, which yields an approximation of the item originally associated with the probe item. This approximation is redintegrated in a manner similar to that of the retrieval-based hypothesis, and the new item is recalled and used as the next probe. This simulation is plausible in both neurological (Phillips & Singer, 1997) and psychological (Murdock, 1997;

Murdock & Lewandowsky, 1989) terms. TODAM demonstrates not only that item and order information can be stored within the same architecture, but that each can actively aid the recall of the other.

Similar models to this have been suggested by Metcalfe (1990), Humphreys, Bain & Pike (1989) and in a wider cognitive context by Halford, Bain, Mayberry & Andrews (1998) and Phillips & Singer (1997). Although some of these models have been criticised as being somewhat isolated from the rest of human cognitive function, some have made efforts to integrate other cognitive phenomena into their model. For example, Dennis & Kruschke (1998) have incorporated attentional effects into their model, showing how attention can influence storage and retrieval of items.

Categorisation of Word Lists

Not only does the type of order encoding play a role in memory, but it is necessary to see how these processes interact with other organisational factors. In general, if a word list is categorised, that is to say that the list contains words which share common parameters, it is better remembered than one which has no internal structure (e.g. Gregg, 1976; Tzeng & Cotton, 1980; Cowan, 1997; Crowder, 1979). This applies to many different types of category coding, including semantic categorisation, grouping by word frequency, and word generation paradigms (Crowder, 1979; Nairne & Neumann, 1993). Analysis of category effects is important in current research because many list manipulations are performed through the use of list categorisation. Support for the order-encoding hypothesis relies heavily on such procedures (Delosh & McDaniel, 1996). If categorisation affects list recall independently of the effects of item elaboration, then data supporting the hypothesis will be compromised.

Frequency Effects

The most comprehensive review of frequency effects in list memory to date was carried out by Gregg (1976). Findings were reported from a number of studies that when words of high and low frequency are presented in lists, their recall is characterised by clustering (e.g. Bousfield & Cohen, 1955). That high and low frequency words tend to clump together during recall suggests that words of similar type are good cues for each other.

In addition to this finding, it is commonly found that high frequency words are better recalled than low frequency words, in long-term (e.g. Sumby, 1963; Watkins, 1977; Whiteman, Nairne & Serra, 1994) and medium to short-term memory tasks (e.g. Delosh & McDaniel, 1996; Poirier & St-Aubin, 1996; Hulme *et al*, 1997). The most common explanation of this effect is that high frequency items have greater inter-item associative strength, and this allows words to cue each other during recall (Gregg, 1976). Indeed, Deese (1960) attempted to compare effects of frequency when inter-item associative strength was controlled for. He found that under these conditions, word frequency effects were greatly attenuated. Although some evidence exists that this effect could be due to approximation of stimulus items to English (Underwood & Postman, 1960), the fact that this effect remains in lists of purely nouns or adjectives (Gregg, 1976) suggests that some non-linear process is occurring. These effects are highly compatible with theories that suggest that similar variables affect short and long term memory processes, and especially those which are based upon interactive, distributed representations of cognitive function.

Data presented by Delosh & McDaniel is used to argue against the associative recall model of Gillund & Shiffrin (1984). Their third experiment tested word lists composed of uneven ratios of low and high frequency words. In mixed lists, high frequency words suffered a free recall disadvantage both when high frequency words were dominant, and when low frequency words were dominant in the mixed lists, compared to pure lists. However, order encoding overall was unchanged in mainly high frequency lists, but impaired in mainly low frequency lists, supporting the order encoding account. Also, no difference was found between the mixed lists with different concentrations of high frequency words. Gillund & Shiffrin's model would have predicted recall values to tail off as the number of high frequency words within a list was reduced, as would a model based upon associative encoding. However, the actual figures reported in the article do suggest a trend towards this phenomenon, with mainly low frequency mixed lists obtaining 12% less items correct than mainly high frequency lists. This is an area worthy of further investigation.

One paper which finds support for theories other than order encoding is that of Poirier & St-Aubin (1996), who subscribe to the redintegration model. In presentation of short lists of words for immediate serial recall, they found the standard effects for frequency on positional serial recall (attempting to recall items in correct locations, rather than emphasising associative recall) and for an item recall measure. However, when they came to analyse errors incurred in positional recall, they found no effects of frequency at all. They interpret the frequency effect in recall as due to traces of high frequency words being easier to interpret and more accessible once decay has commenced.

Whiteman, Nairne & Serra (1994) also found no effect of frequency on positional recall *per se*, but when coarser measures of order were used in which participants were asked to identify whether an item had appeared in a particular list, a frequency effect emerged. This suggests that we do encode order associatively, rather than positionally, as discussed in an earlier section.

Evidence for a frequency effect on immediate recall also comes from the shape of serial recall curves. In a serial recall study, Hulme *et al* (1997) found that although word frequency appeared to have no direct effect on order encoding errors, the shape of the serial recall curve was affected, with items in the middle of the curve gaining a high-frequency advantage. It would appear that some measures of order encoding are more sensitive than others.

In sum, frequency effects occur in free recall, but are a little more elusive in serial recall. A free recall paradigm has been shown to elicit the strongest evidence for a frequency effect on order encoding (Delosh & McDaniel, 1996), and serial recall paradigms appear to show weaker effects (Hulme *et al*, 1997; Poirier & St-Aubin, 1996). The most important difference between these experiments is the use of the measure of order. One explanation for this is that serial recall may primarily be sensitive to a positional encoding strategy, at the expense of detecting the use of an associative strategy. Once one order mistake is made, all subsequent items, even if recalled correctly with reference to the erroneous item, are scored as positionally incorrect. Thus, the small effect of frequency on recall success proposed by the order encoding hypothesis may be smothered by the inherent confound between item and order coding in the serial recall paradigm.

If order encoding is based on an associative strategy, serial recall experiments may not be the best way to measure effects on the order of recalled items. The free recall paradigm does not allow the possibility of measuring order retention in a positional way, and so the more flexible measure of associative order encoding was taken in the Delosh & McDaniel study. It could be that this subtler method of detecting inter-item associativity might offer a more informative window upon the frequency-order encoding effect.

Semantic Relatedness

A vast amount of data has been collected on the effects of semantic relatedness upon free recall of words. Generally, the results for free recall follow a similar pattern to those for frequency: it has been reported that if the words in a list are all from the same category, then an item recall advantage is obtained over randomly constructed lists (e.g. Crowder, 1979). This is a common finding in free recall, but not so common in serial recall (see Poirier & St-Aubin, 1995, 1999). The question of interest to this thesis is that of the effect of semantic relatedness upon order encoding *per se* in free recall.

Semantic coding does appear to play a part in the associative recall of words. If a list of words from various semantic categories is given to a participant, then they will tend to recall the items in clusters which represent those categories (Deese, 1959; Jenkins & Russell, 1952; Bousfield, 1953). Semantic codes can therefore influence the order of the output of items. However, evidence of semantic information influencing the recall of the original retention of order is inconsistent.

Underwood & Postman (1960) found that approximation to English was a predictor of serial recall success. Their research showed that if a sentence makes sense grammatically, then it is likely that will be recalled better than one which is less orthodox. They supposed that if items in the lexicon were linked in such a way as to facilitate each other in an associative manner, then semantic factors might also influence these associations. Another reason for supposing that semantic information has a role to play in the encoding of order information is based on the theory of Tzeng & Cotton (1980). If order is encoded through study-phase remindings, then it is presumably the case that similar items will provide associative cues to each other, whereas dissimilar items may not. Thus, one would expect more correct order judgements in semantically categorised lists than in uncategorised lists.

However, Baddeley (1966a) reported that on a short- term order-reconstruction task, semantic categorisation actually disrupted order retention. Similar results have been reported by Nairne (1990*a*) and Lewandowsky & Murdock (1989) in short-term serial recall tasks. Nairne & Neumann (1993, experiment two) examined the effects of semantic categorisation on long-term memory for order. They had participants incidentally learn lists of words which were either categorised, uncategorised or mixed. Instructions were to judge which of two list words was most recent for all lists after all of the lists had been presented, with results showing that categorised lists gave best relative recency scores, then mixed, then uncategorised. However, Nairne & Neumann did not find support for semantic factors influencing order encoding *per se*, but rather that semantic information acted as a distinct and unique cue for the participant to access the correct list record (for each list, a unique category label could be applied).

Analysis of error information made during recall can help to shed light on order encoding. Poirier & St-Aubin (1995) carried out a series of tests involving short-term serial recall which also examined the types of errors made. Both a serial positional measure and an item-based measure of sub-span lists was significantly better for categorised than non-categorised word lists. This effect was independent of the articulation rate of individual items. However, an analysis of errors showed no significant differences for order encoding across conditions, but a definite advantage for item recall. Poirier & St-Aubin conclude that semantic information somehow influences order encoding by enhancing item recall, rather than better order encoding, induced by richer semantic information affecting item recall.

These findings are by no means consistent, however, and most of the above studies have been criticised by St-Aubin & Poirier (1999b) for not taking into account the speech rate of the items used. More importantly, they point out that the measures of order retention used in serial positional recall paradigms often lack validity. Often, order retention is assessed by computing the number of order errors which occur (Healy, 1974). This is not appropriate when analysing factors such as semantic similarity, which have been shown to improve item recall, because if more items are recalled, the chances of order errors within those items increases disproportionately. St-Aubin & Poirier (1999b) advocate the use of methodological or statistical methods of controlling for this discrepancy. In their serial recall study, when order error data were statistically corrected, no overall effect of semantic similarity on recall of order information was found.

However, semantic similarity did show a significant interaction with certain serial positions within their experiments. The effect was small and inconsistent, but earlier serial positions sometimes showed a similarity advantage, whereas later ones often did not. The trace redintegration hypothesis predicts that semantic information will not affect the encoding of order information, as semantic information can only be effective at trace reconstruction, not at the extraction of the next phonological trace in the output protocol.

The order-encoding hypothesis may have different predictions about semantic-order effects. On the one hand, inter-item associations between semantically related words may encourage associative recall. On the other, semantic codes may be automatically elaborated as part of word recognition, as in the Plaut *et al* (1997) network model, which Delosh & McDaniel (1996) hypothesise to interfere with organisational encoding. However, if this activation is automatic, and does not require extra cognitive effort, then the extra elaboration should not take up the resources hypothesised to be necessary for the processing of order information.

Aims of the Research

The experiments described in the next two chapters aim to explore the issues described above in greater depth. Firstly, it is interesting to assess the importance of order encoding on free recall success, and to this end, the experiments described here will probe the relationship between order retention and the overall number of items recalled. Of further interest is the relationship of long-term lexical information to the process of order encoding. Since much of the data described above tends to use a strict serial position score to assess the degree of order information retained in a manipulation with only scant evidence in favour of long-term effects, re-examination of the question with a more sensitive measure of order encoding was sought. Finally, an examination of the importance of phonological coding in short term memory in general and in order encoding in particular was deemed to be of interest in the general discussion as to whether short-term and long-term memory operate within separate and distinct modules or whether there is a degree of overlap between the two.

Chapter Two: The Order Encoding Hypothesis in Depth

The aim of this chapter is primarily to attempt a clarification of order-encoding processes which occur during free recall. As discussed in the previous chapter, many conflicting data exist in determining whether order is encoded naturally as a mnemonic in free recall, on whether the high frequency advantage in pure lists and the low frequency advantage in mixed lists are founded upon order encoding, and on what part long-term semantic information plays in order retention. The initial objective in attempting such a clarification depends in part upon successfully replicating Delosh & McDaniels' (1996) results in which the pure and mixed list frequency effects were found. In addition to this simple replication, it was hypothesised that semantic categorisation would play its own role in both straightforward recall and in order retention (Bridges & May, 1997). A semantic categorisation element was therefore added to the design.

Experiment One

One of the fundamental assumptions of Nairne & Neumann's (1993) paper was that in addition to the existence of intra-list effects of word type, *inter*-list effects could also be induced. This is especially true when accounting for semantic factors. In some of Delosh & McDaniel's (1996) experiments, participants had the opportunity to compare different list types within experiments, and to overtly code them into categories. If they subsequently discovered one type of list to require less effort to recall (for example high frequency, semantically categorised lists), then any effects of list type may have been confounded by motivational factors. Hence in the present study, each participant was assigned to a different condition.

Another improvement made to the original design was to attempt to control for other word-based variables. It has long been known that it is very difficult to separate out the effects of factors such as frequency, imageability and concreteness (see Gregg, 1976), and therefore the present study chose word lists which were matched for age of acquisition, concreteness and imageability, in order to isolate the effects of semantic categorisation alone.

An independent measure of the degree of semantic categorisation of the words lists was also taken. A number of balanced word lists were created by the experimenter, and were then judged for semantic relatedness by an independent group of participants. This validation of the semantic categorisation variable was deemed necessary as the list categories were somewhat diffuse (see Appendix One).

In free recall studies, it is impossible to measure the extent of order information retained by scoring items by correct serial position, as an omission early on in the recalled list would necessarily ensure that all other items recalled would be in the wrong location. A somewhat more flexible measure is called for in which the relative positions of words are assessed. The Asch-Ebenholtz index (Asch & Ebenholtz, 1962a) was used by Delosh & McDaniel (1996) to provide a measure suitable for use with free recall designs. It involves taking a ratio of the number of adjacent pairs recalled in the correct order to the total number of pairs recalled. The index is a number between 0 and 1 with 0.50 representing an output protocol which is random in terms of the order the items. Higher values represent outputs in which more of the original order of the items is retained and a value of one indicates perfect serial recall. A score of zero represents perfect backward recall.[3]

The benefits of such a measure are twofold. Firstly, it enables a more flexible assessment of the amount of order information retained in free recall output protocols, and

---

[3] In their experiments, Delosh and McDaniel found that order was encoded in free recall: they observed Asch-Ebenholtz scores of around 0.70 on average.

secondly, it may be a more useful measure in detecting order effects in semantic categorisation. For, if one assumes that the semantic effect on order may have its basis in inter-item association, a method of order analysis which depends upon associations with specific temporal or spatial 'slots', such as Poirier & St Aubin's (1999) method in serial recall would not be able to detect these subtle interactions. This is because, as before, if one error is made early on and the participant is not aware of the error, all later output will be classified as erroneous. Semantic associations may be stronger in some parts of the list than others, and a relative measure of order encoding will be much more sensitive to these within-list patterns.

Delosh & McDaniel's use of an order reconstruction task could also measure order indirectly, but would cause a certain amount of bias, as after the first trial, participants were aware that order was important, and would have the opportunity to change their recall strategy accordingly. The primary aim of this chapter is to assess how much frequency and semantic categorisation naturally influence order encoding in free recall, and an order reconstruction task would therefore be inappropriate.

Finally, predictions can be made concerning errors committed during testing. If the free recall semantic category advantage reported previously is due to participants simply guessing words which are connected to a category label, then we would expect more semantic errors (incorrect responses which are semantically related to target words) within categorised lists than in non-categorised lists.

Previous list intrusions may also be relevant to the study. Gregg (1970) reported effects of word frequency upon previous list intrusions (words which have been previously presented, but which are recalled in the wrong list). This is probably due to sophisticated guessing: participants guess at words which have active traces, and it is presumed that high frequency words have lower selection thresholds than low (see Gregg, 1976). It is expected that our results will also show increased previous list intrusions for high frequency words. Semantic categorisation may also affect this type of error, as Nairne & Neumann (1993) point out that semantic information can act as a marker for the whole list. In this case, we would expect low levels of previous list errors to occur in categorised conditions, whereas higher levels should occur in lists for which no overt semantic code is available.

A written recall design was employed throughout this chapter for one reason: the experiments reported were all based upon the design of Delosh & McDaniel (1996), and as such, the introduction of an alternative recall procedure would have undermined any comparisons or contrasts between their study and the present one. Experiment seven in chapter three does explore the difference between spoken and written recall methods.

For all experiments described in this thesis, participants were naive. As the word lists used were necessarily similar each time, using the same participants again for another experiment would have jeopardised the status of the low-frequency words, as by definition, these words would cease to have a low frequency of exposure if they were repeatedly presented to the participants involved.

Method

Participants were 64 University students from Aston University, aged between 18 and 28 years of age. They were drawn from the Psychology undergraduate corpus and completed testing as part of their course requirement.

Design. The experiment had a 2 x 2 design. Independent variables were frequency (high or low) and grouping by semanticity (categorised or uncategorised). Measurement of performance was conducted on a between-subjects basis. Dependent variables were the proportion of correct responses for free recall tasks, and a measure of similarity between the order of target lists and output protocols.

Materials. Word lists were constructed by drawing words from the Oxford Psycholinguistic Database (Coltheart, 1981). 288 bisyllabic words were chosen such that their mean values for age of acquisition, concreteness and familiarity did not differ from any individual items by a value of more than one (scales ranged from 1 to 7). These words were placed in one of four list types, two lists in which words were deemed to be loosely semantically categorised, and two in which words were chosen to have no obvious semantic relationship. All lists were nine words long. The validity of the categorisations was corroborated by presenting the lists to 12 participants who were told to rate the degree of semantic categorisation for each list on a scale from one to seven. Semantically related lists were judged significantly more categorised than non-categorised lists (*paired* $t_{(11)} = 22.40$; $p < 0.001$). Each categorised list was either classed as categorised as a whole (with the assumption that all nine words in the list were related to each other) or as uncategorised.

In addition to lists being divided by semantic category, they were also divided in terms of frequency. Words in the high frequency list had a frequency greater than 50 per million, and low frequency words had a frequency of 7 or less per million (Francis & Kuçera, 1982). This gave four list types: low frequency, non-categorised (LFN); low frequency, categorised (LFC); high frequency, non-categorised (HFN) and high frequency, categorised (HFC). There were eight lists of each type, making 32 nine-word lists in all. The words used in the present experiment, along with a full description of the experimental design which was used to test them are listed in Appendix One.

Apparatus. Stimuli were displayed on an IBM compatible PC, and presented through the Windows interface, using a program written in Microsoft Visual Basic 5.0.

Procedure. Participants were assigned to condition by the order in which they presented themselves at the laboratory. They were instructed that they were to receive a memory test, and that their task was to remember the words which they saw upon the computer screen. They were then to be given a short distracter task involving simple maths problems for a short time, after which they were told they must recall as many of the words as they could from the list. Instructions for recall were written upon the recall sheet such that participants were asked to write down the words as they occurred to the participant at time of recall, and that one word was to be written on each line. At no point was any mention of the order of the items made. Participants were not informed of the nature of the word list they were to view, either in terms of its semantic categorisation status or the frequency of the words themselves.

During each trial, words were presented in the centre of the screen, covering between approximately four and seven degrees visual angle. These measurements are approximate, as no attempt was made to fix the participant's head position. Each word was presented on screen for 2000 msecs, with a 500 msec inter-stimulus interval. Words were shown in lower case in black against a white background. Participants were shown eight word lists, each of which contained nine words. After the final word from each list had been presented, there was a 500 msec blank screen, followed by the simultaneous presentation of 20 simple arithmetic problems. Subjects had been previously instructed to mentally solve and call out the answers to these problems so that the experimenter could record them. These answers, once recorded, were discarded. After 40 seconds, the maths questions disappeared, and were replaced by the prompt to recall the words from the list ('*****'). Participants were given a minimum of one minute to write their answers down, and no maximum limit. This interval was chosen as in the original Delosh & McDaniel study, participants were given unlimited time to attempt order reconstruction of the items from lists. Hence, in the present experiment, care was taken to ensure that participants were not subject to time pressures which had been absent in the Delosh & McDaniel study. Once the subject was ready to proceed, the space bar was pressed, and the next trial began. All instructions given to participants for all experiments reported in this thesis are reported in Appendix Two.

A practice block of two trials preceded the experiment proper. Word lists that matched experimental condition but had been rejected from pilot studies were used to create practice blocks. They contained none of the actual words to be used in the experimental block. In total, the experiment lasted approximately 30 minutes.

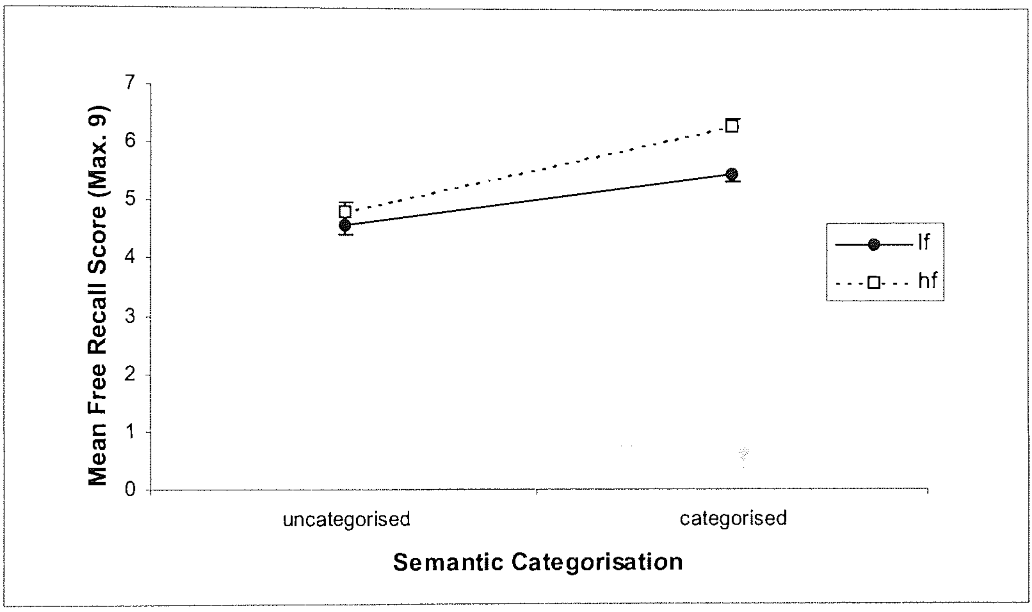
Results

Results were analysed in terms of free recall scores and of input-output correspondence (Asch & Ebenholtz, 1962a). Input-output correspondence scores were obtained by scoring each adjacent pair of responses in the recall protocol for relative order. If a pair was recalled in the correct order, no matter what the distance between the same items at presentation, then it was counted as being correct. The number of all correct pairs was then divided by the total number of pairs of items listed on the response sheet to give a ratio of correct ordered recall to items. Assuming that order encoding involves the coding of associations from one item to the next, then the expected chance value for this ratio of input-output correspondence should be .50. Any values above this would signify that order encoding has taken place, and the higher the value, the more order information has been retained. Summary tables for all statistical analyses carried out in this chapter can be found in Appendix Three.

Free Recall

Mean free recall across all lists was 5.29 words (SD= 1.82). Free recall scores by word type are shown in figure 2.1. An independent measures ANOVA with two factors, each with two levels, was carried out on free recall scores. Factor one, word frequency, had two levels, high and low, and factor two, semantic categorisation, also had two levels, categorised and uncategorised. Mean scores for each participant were entered into the design. The expected advantage for recall of high frequency words over their low frequency counterparts was significant ($F_{(1, 60)} = 4.18$, $p< 0.05$), and the recall enhancement for semantically categorised lists over unrelated lists was highly significant ($F_{(1, 60)} = 18.48$, $p< 0.001$). There was no interaction between the two variables ($F_{(1, 60)} = 1.06$). On breaking down the results to a group-by-group basis, post-hoc Scheffe tests of all six possible pairwise comparisons revealed that only the HFC lists differed from LFN and HFN lists significantly at the .05 level (see Appendix Three). These findings support the notion that semantically related lists are recalled better than unrelated lists, and that within the categorised lists, high frequency words are better recalled than low. The results for free recall by word frequency have replicated the results of Delosh & McDaniel (1996; experiment 1).

Figure 2.1: Mean free recall score per list by semantic categorisation and word frequency (Experiment 1).

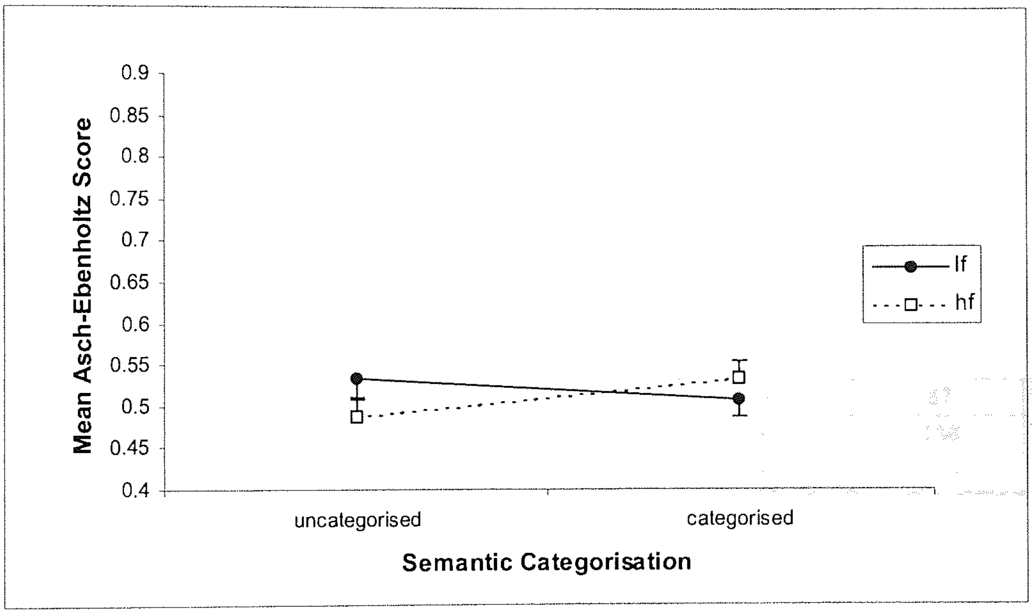


Input-Output Correspondence

Mean Asch-Ebenholtz score across all lists was 0.52 (SD= 0.26), which did not differ from chance (one-sample t $_{(511)}$ = 1.45). Winer (1962) notes that proportional data does not often fit a normal distribution, and that therefore one of the assumptions allowing the use of parametric statistical tests is not met. However, a histogram plot of the Asch-Ebenholtz scores obtained in this study did approximate a normal curve (see Appendix Three), and therefore the scores were not subjected to an arcsine transformation before entering them into a t-test.

Figure 2.2 shows mean Asch-Ebenholtz scores by condition. From the standard error bars, one can see that there is much variability in the data. Analysis of Variance shows there to be no significant effect of either frequency ($F_{(1, 60)} = 0.12$) or of semantic categorisation ($F_{(1, 60)} = 0.12$). Neither was there any interaction between the two variables ($F_{(1, 60)} = 1.42$). Delosh & McDaniels (1996) findings of order encoding have therefore not been replicated.

Figure 2.2: Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency (Experiment 1).



Error Analysis

In total, 329 commission errors were recorded from a grand total of 4096 presented words (an 8% error rate by word). Errors were marked as belonging to one of four categories: semantic errors, phonological errors, previous list intrusions (PLIs), and 'other'. Semantic errors were classed as such if the response in question was directly linked to one of the stimulus words (for example, *bacteria* and *microbe*). Phonological errors were simply incorrect responses which rhymed with stimuli (e.g. *banner* and *manor*). Previous list intrusions were words which were incorrect for the list, but which had been stimulus words for previous trials. Finally, several responses were classed as 'other', as responses were either nonwords, or were ambiguous in terms of the classifications mentioned above (for example, the response *shiver* was deemed to be both phonologically and semantically related to the word *quiver*). Error data are displayed in table 1.

A chi-square test revealed significant differences between conditions for total number of errors per condition ($X^2 = 16.23$; df =3; $p < 0.01$). Separate analyses for each error type individually revealed that semantic errors differed by condition ($X^2 = 77.98$; df = 3; $p < 0.001$), as did PLIs ($X^2 = 18.02$; df = 3; $p < 0.001$). An inspection of table 1 suggests that semantic errors occur largely in the categorised lists, and that HFN lists yield the most PLIs. There was no reliable contingency between condition and the making of phonological errors, however ($X^2 = 4.53$; df = 3).

Table 2.1: Frequency of response errors scored by type and experimental condition.

ERROR TYPE	LFN	LFC	HFN	HFC	**TOTAL**
SEMANTIC	8	47	5	61	**121**
PHONOLOGICAL	5	4	4	0	**13**
PLI	11	10	28	8	**57**
OTHER	29	20	58	31	**138**
TOTAL	**53**	**81**	**95**	**100**	**329**

Discussion

Analysis of the error patterns of this experiment largely confirmed hypotheses, but for the lack of phonological errors. This may have been due to the long retention interval, in which most of the phonological information available to the participant decayed. That categorised lists showed less evidence of previous list intrusion supports theories of pro-active interference laid down by Nairne and colleagues (Nairne, Neath & Serra, 1997; Nairne & Kelley, 1999; Nairne, Whiteman & Kelley, 1999), and supports Crowder's (1993) conceptualisation of item memory as "coarse-grained order memory". Semantic errors occurring in semantically categorised lists point to the predicted sophisticated guessing strategy. These issues will be returned to in chapter four. The remainder of this discussion is focussed upon the primary results.

Free recall results generally confirm expectations: both frequency and categorisation have an effect on free recall and on the strategies used by participants to remember words. For free recall, word lists were better recalled when items were all of the same semantic category, and when they were all of high frequency (figure 2.1). Furthermore, these effects of semantic categorisation and frequency were additive[4]: the best recall of all four lists was for the categorised, high frequency word lists, followed by the intermediate lists, and finally, low frequency, non-categorised lists were the worst. This supports the well-founded notion that participants use general semantic associations to remember lists of words. However, when taken in conjunction with the results for order encoding, the persistence of the high-frequency recall advantage in the absence of any notable effects of order encoding casts doubt upon the order-encoding hypothesis.

Contrary to expectations, the amount of order information retained in this experiment was not different from chance. The Asch-Ebenholtz value obtained was also considerably less than the 60-81% average reported in Delosh & McDaniel's (1996) series of experiments. There were only two major differences between their study and the present one if the semantic categorisation variable is disregarded. The first is that there were tighter controls placed on the word lists in the current study: meaningfulness, age of acquisition and imagability were all controlled for in the present design, and Delosh & McDaniels' results for order encoding may have been partly dependent upon these factors. However, a more obvious difference was that in their study, participants knew that they were going to be tested on an order reconstruction task after the free recall task, whereas in the current study, participants were never aware that the order of their responses was to be analysed. Knowledge of such a task may have prompted an encoding bias towards order retention in their sample.

The order-encoding theorists reported the finding that in mixed lists, low frequency words are better recalled than in pure lists (May & Tryk, 1970; Duncan, 1974; Gregg,

---

[4] The use of the word "additive" throughout this thesis is intended to convey the notion that the effects of semantic categorisation and word frequency together were greater than the effects of either individually. It is not intended to

Montgomery & Castano, 1980). Low frequency words can be said to be advantaged in these situations because order cues will be enhanced for low frequency words, compared to pure lists, and high frequency words will be disadvantaged. This is due to the premise that order encoding for any individual item can only be as good as the order encoding for items around it. Thus, high frequency words in mixed lists are processed for order information as much as in pure lists, but because the surrounding low frequency words suffer an order penalty, as described previously, overall order encoding is not as salient for high frequency words. The opposite is true for the low frequency words, which will have an overall enhancement of order encoding, due to the order information available from high frequency words. Low frequency words will be advantaged as, although order encoding will be equivalent for LF and HF words, LF words will be more episodically salient, due to their unusualness. This has been shown to occur consistently in recognition studies (e.g., Shepard, 1967; McCormack & Swanson, 1972). Delosh & McDaniel hypothesise that this effect occurs because of richer elaboration afforded by resources being used to process item representations for words. The redintegration account (Hulme *et al*, 1997; Walker & Hulme, 1999) does not explicitly predict the low frequency advantage for mixed lists, predicting as it does that the traces of high-frequency words will always be more easy to redintegrate than low-frequency words, and so even in mixed lists, high-frequency words should still gain an advantage.

Experiment Two

Experiment two examined the mixed list effect reported by Delosh & McDaniel (1996; experiment 1A). The design was similar to that of experiment one, except that the lists of words were each comprised of four LF words and four HF words. If the results of the previous experiment are robust, then one would expect an advantage for semantically related mixed lists in free recall. Within lists, one would also expect the LF advantage to be

---

imply that the joint effect was precisely the sum of the two separate effects.

supported, in accordance with Delosh & McDaniel's encoding/ elaboration hypothesis. On the other hand, if the straightforward redintegration account is correct, then one would predict a high-frequency advantage to occur in mixed lists. In either case, it would be of use to examine the effect of semantic categorisation on the frequency effect within mixed lists. As in the first experiment, the effects of semantic and frequency categorisation on overall recall were expected to be additive.

According to the order encoding hypothesis, lists of mixed frequency should demonstrate equivalent levels of order encoding. Low frequency words should benefit from the order information gained by simply being amongst high frequency words. High frequency words, although supposedly encoded for order information, should lose some of that location information due to their reference to weakly order-encoded LF words. Again, on the strength of findings from Bridges & May (1997), it was expected that semantic categorisation would increase input-output correspondence.

Method

Participants were 36 University students from Aston University, aged between 19 and 24 years of age. None of the participants had taken part in previous experiments..

Design. For analysis of free recall scores, a mixed 2 x 2 design was used with list categorisation as the independent measure and word frequency as the within-subjects factor. For input-output correspondence, a between-subjects t-test was employed (it being inappropriate to attempt order analysis of words by frequency in mixed lists). It was inappropriate to attempt order analysis of words by frequency in the mixed lists, as the measure of order used was a relative one. Hence, high and low frequency words would be assessed for order relative to each other and so each word type would yield an equivalent score.

Apparatus & Materials. Apparatus was the same as that used in the previous experiment. Mixed word lists were created in the same way as before, but this time, each list was composed of four high frequency and four low frequency words, randomised so that no words of similar frequency were repeated more than twice. 12 mixed-frequency, semantically categorised lists (MC) and 12 mixed-frequency, semantically non-categorised (MN) lists of eight words each were created. Categorisation was tested with a further 10 participants who were asked to rate lists for semantic relatedness as before. There was found to be a significant difference between list types (*paired* $t_{(9)} = 29.16$, $p < 0.001$). These lists are available for inspection, along with detailed methodology for their validation in Appendix Four.

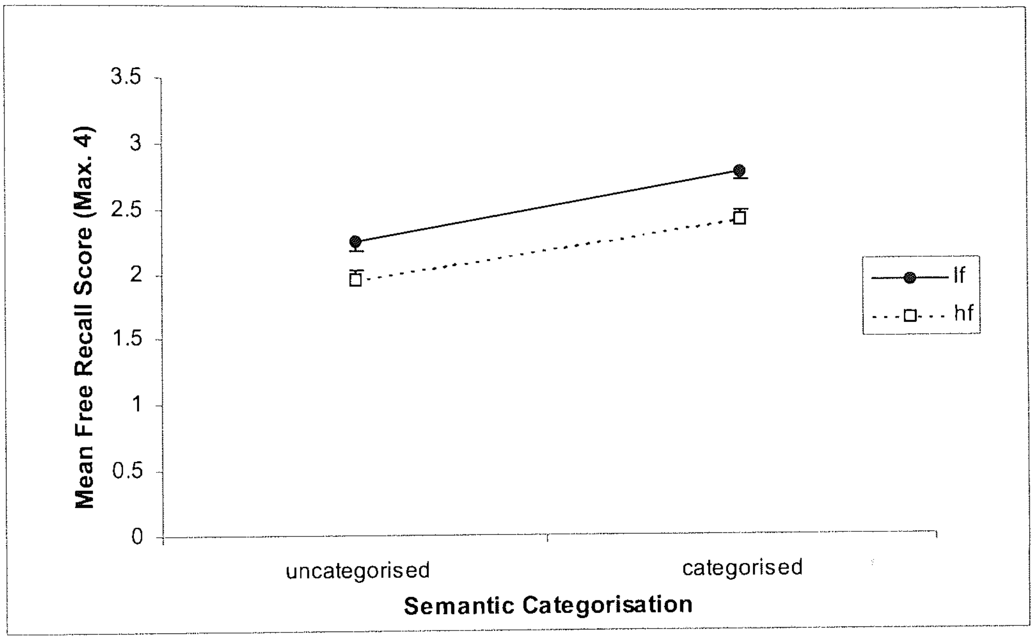
Procedure. Participants were allocated to conditions alternately. Each individual saw either the 12 mixed frequency related lists, or the 12 mixed frequency unrelated lists. The lists were presented in random order by computer. All other details are as described in experiment one.

Results

Free Recall

Mean free recall score overall was 4.73 words per list (SD= 1.63). Mean recall scores were entered into a mixed ANOVA model, with semantic categorisation as an independent factor with two levels (categorised and uncategorised) and word frequency as a repeated measures factor with two levels (HF and LF). Scores broken down by list type and frequency are displayed in figure 2.3, which shows a consistent LF advantage over HF words. The mean number of words correctly recalled for LF lists was 2.53 (SD= 1.06), and for HF lists, 2.20 (SD= 1.10). This small difference was reliable ($F_{(1, 34)} = 17.58$; $p < 0.001$). The figure also demonstrates a strong semantic category advantage ($F_{(1, 34)} = 13.31$; $p < 0.001$). Semantically categorised lists (mean = 5.25; SD= 1.55) were recalled 12.5% better than unrelated lists (mean = 4.21; SD= 1.55). No interaction between semantic relatedness[5] and word frequency was found ($F_{(1, 34)} = 0.19$).

Figure 2.3: Mean free recall score per list by semantic categorisation and word frequency (Experiment 2).



Input-Output Correspondence

The mean Asch-Ebenholtz score overall was 0.49 (SD= 0.29), which was not significantly different from chance (one sample $t_{(431)} = 0.59$). Mean scores for MC lists (mean = 0.50; SD= 0.24) and MN lists (mean = 0.49; SD= 0.30) did not differ significantly from one another ($t_{(34)} = 0.51$).

Error Analysis

From a total of 3456 words displayed in the experiment as a whole, 299 incorrect responses were recorded (8.7%). As before, these errors of commission were classified into phonological, semantic and previous list intrusion errors, and can be seen in tabulated

---

[5] Semantic relatedness and semantic categorisation are used interchangeably throughout this thesis, as the

form in table 2. A chi-squared analysis shows a significant effect of list type on error type (X^2 = 97.08; df =3; $p < 0.001$). Individual analyses for semantic and PLI errors show significant effects of list type also (X^2 = 59.11; df =1; $p < 0.001$ and X^2 = 13.4; df =1; $p < 0.001$, respectively). More semantic errors occur in the categorised lists and more previous list intrusions occur in the non-categorised lists. The analysis of intrusion errors described here refers solely to those items that could only be described as PLIs. Any previous item which may have also been semantically or phonologically related was recorded as "other". Further analysis of PLIs by breaking down errors into subdivisions was deemed inappropriate, considering the low cell counts in a chi square analysis.

Table 2.2: Frequency of response errors scored by type and experimental condition.

ERROR TYPE	MN	MC	TOTAL
SEMANTIC	15	96	**111**
PHONOLOGICAL	18	11	**29**
PLI	29	7	**36**
OTHER	88	35	**123**
TOTAL	**149**	**150**	**299**

Discussion

Errors followed the same pattern as experiment one, with semantically categorised lists benefiting from a reduction in intrusions from other lists, but suffering a penalty from semantic confusions or sophisticated guessing errors. More will be said about these results in chapter four.

As expected, the low-frequency free recall advantage for words in mixed lists was obtained. This supports the notion of the bizarreness effect, with words which stand out from the norm (in this case, LF words) being more readily memorable than common words (e.g. Merry, 1980; Hirshman & Bjork, 1988). Unfortunately, it is impossible to ascertain just how important this effect is in terms of implicit order encoding, as none was observed. This

---

method of creating categorised lists was based upon participant's perceptions of semantic relatedness.

is somewhat surprising, as in both the Delosh & McDaniel (1996) studies, and in Bridges & May (1997), a strong order-encoding effect was found. Again, this casts doubt upon an order-encoding explanation for the frequency effects.

It could be argued that the reason for the non-occurrence of order retention as measured by the Asch-Ebenholtz index may be due to strong recency effects, which, because of the way the index is calculated, would cancel out any order encoded in the early part of the list. This is unlikely, as recency effects have almost always been found to disappear after a filled delay (Postman & Phillips, 1965). Even if recency effects did occur in this paradigm, then they must have occurred equally for semantically categorised and uncategorised lists, as there was no absolute difference between Asch-Ebenholtz scores for these lists. This is a valid inference if one accepts the assumption that recency effects would impact each list to a similar degree. As previously mentioned, by their choice of design, Delosh & McDaniel allowed their participants to explicitly make the link between recalling and ordering their material, and this could very easily give rise to extended use of order information as a viable recall strategy.

In the pilot studies, however, subjects were still not aware of the experimenter's interest in order encoding, and yet still displayed a certain amount of order retention in their output protocols. There was one big difference between the pilot studies and Experiments 1 & 2, and that was that the trial runs used an immediate recall paradigm, whereas the experiment proper used the delayed recall design of Delosh & McDaniel (1996). It may be that immediate recall depends on order information more than delayed recall, which depends less on surface features of the word (like frequency and sound), and more on 'deeper' factors such as meaning and imagery. It is to this possibility that this chapter now turns.

Short Term and Longer Term Memory

As mentioned in chapter one, there has existed a classic dichotomy between short- and long-term storage. The main reason for assuming this distinction was the finding that we can only remember around seven items in the short term, but that long term memories seem to have boundless capacity (Miller, 1956). Existence of serial position effects in list recall (Murdock, 1962) can also be interpreted as evidence for the existence of at least two distinct stores. Better recall for initial items in the list is attributed to long-term storage: an episodic, long term memory trace, whereas augmented probability of recall for later items in the list can be attributed to a limited capacity store with access to highly salient, yet rapidly decaying, memory traces (e.g. Broadbent, 1958; Atkinson &Shiffrin, 1968; Waugh & Norman, 1968). Delays between presentation and recall are found to abolish the recency effect without harming the primacy effect for spoken items (Glanzer & Cunitz, 1966), which is entirely compatible with the idea of a separate short-term store. Primacy, on the other hand, can be radically affected by variables which leave recency unaffected: for example, word familiarity, rate of presentation and frequency (Glanzer, 1972).

One particular study by Watkins (1977) showed that word frequency affected primacy, rather than recency. Word lists were organised such that high frequency and low frequency words were blocked together at either end of mixed frequency word lists. The otherwise identical lists were always recalled better if high frequency words were at the start of the list, rather than the end. However, these findings were interpreted as evidence that information stored in long-term memory can affect the recall of items in the short term. This interpretation was bolstered by the further discovery that high frequency words showed an advantage compared to low frequency words across the whole list, including recency positions.

Hulme, Maughan & Brown (1991) reported similar data. Lexicality was found to have a long-term effect upon short-term recall. Words and pseudowords were presented for recall, resulting in a strong bias towards words. This bias was removed, however, when participants were taught meanings for previously incomprehensible words, drastically improving recall. This illustrates that the boundaries between short and long-term memory are not hardwired, but are dynamic and intertwined.

Findings that long-term information is relevant to short-term recall are common throughout the literature (e.g. Bourassa & Besner, 1994; Crowder, 1979; Delosh & McDaniel, 1996; Gregg, 1976; Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart, 1997; LeCompte & Watkins, 1993). Crowder (1982) argues the case against the existence of separate stores for short and long term memories, based on this type of finding and upon other data which undermines concepts of memory as temporary storage, followed by rehearsal, followed by long-term storage in strict linear order. For example, researchers have found that long-term recall does not depend solely upon the amount of time an item is rehearsed for (Craik & Watkins, 1973), or that repeated rehearsal leads to better recall (Morton, 1967).

If it is true that order encoding is primarily an immediate recall strategy, it may be that it is dependent on the phonological loop (cf. Wicklegren, 1977a). With delayed recall, as the phonological trace decays more and more phonological information is lost, and this means that in a list, individual items have less unique phonological information associated with them. Therefore semantic influences on trace interpretation will have less and less effectiveness the more uniform the information they are provided with. Immediate recall should not suffer such a problem.

Evidence that suggests short-term storage does not necessarily precede long-term memory is complemented in the literature by data that hint at similar processes occurring in short and long term memory tasks. Nairne (1990*a*, experiment one) aimed to gather data which bridged the gap between short and long term positional recall. He obtained distance functions for positional recall of words at delays of two minutes and ten minutes[6]. These curves were not only very similar to each other, but also to original data collected from

---

[6] Postman & Phillips (1965) and Glanzer & Cunitz (1966) each show putative long-term memory effects occurring in a serial recall curve for recall periods as short as 30 seconds, as discussed above.

immediate recall studies such as that of Estes (1972). That errors made on short and long-term recall follow very similar patterns is strong evidence against the two-store model. As a caveat to this kind of inference, Cowan (1997) notes that many short and long-term effects in memory can be made to seem similar, and he terms this phenomenon "virtual memory". In the same way that the hard-drive of a computer can behave as though it were dedicated fast memory, so too can long-term memory structures behave as though they were short-term memory structures. But, although LTM can exhibit STM properties such as serial position effects, there still remain important differences between the two types of memory. As evidence for this, he points to the rapid decay of phonological information over a two second period as a purely STM phenomenon which does not influence LTM processes.

Crowder & Neath (1991) point out that a recency effect can be obtained in long-term recall, such as for recalling the list of American presidents. They argue that recency effects are induced by coding processes, rather than the existence of different physical stores: it is the rate of processing individual items (inter-stimulus interval) compared to the delay between presentation and recall which influences serial position effects. Further, Crowder & Neath also discuss existing data in terms of retrieval processes: placing a delay between items and recall can change the strategy with which participants recall items. Hence, they focus their approach to memory research in terms of processes, rather than storage. This transformation has occurred over recent years due in part to the evidence mentioned above, and in part to the paper written by Craik & Lockhart (1972), who urged such a conceptual change in cognitive psychology.

Short-term verbal recall is increasingly becoming characterised as simply a part of language processing. Crowder (1982) suggests that computerised models of memory exist which can account for both short-term and long-term recall, without having to overtly specify a difference between the two processes. Estes' (1972) model of feature perturbation fits this criterion, as his theory fits data from both long-term (Nairne, 1990*a*) and short-term memory (e.g. Lee & Estes, 1981). Murdock's theory of distributed associative memory (TODAM) also applies, with items being represented by a code which includes local context features too (Lewandowsky & Murdock, 1989; Murdock, 1997). Data from Murdock's simulations fit both immediate and delayed recall curves obtained from human beings.

As an extension of this new brand of theory, Cowan (1988; 1997) has discussed in depth ideas for combining these two approaches. He imagines short term memory as being the part of long-term knowledge which is currently activated above a threshold level. This harmonises well with the data discussed above, as it can easily explain the effects of long-term knowledge upon immediate recall.

In turn, this model appears to have much in common with the redintegration account of the behaviour of short-term memory. Some theorists have suggested that short-term recall shows recency because the best strategy for immediate recall is to attempt to read the most recent memory trace (assuming it is still active within the brain). As time gradually goes by, then the trace decays more and more, resulting eventually in it being indistinguishable from the background noise. The more a trace has decayed, then the more chance it has of being lost, and the more likely it is that other, alternate recall strategies take over (Crowder, 1982). However, it has been suggested (Nairne, 1990*a*; Hulme *et al*, 1997; Estes, 1991; St-Aubin & Poirier, 1999) that long term information can act as another source of evidence in interpreting the trace, thus effectively lengthening its usefulness. One would expect, therefore that the more strategies available to a participant in interpreting those decaying traces, the more likely that those traces will be interpreted correctly, leading to greater performance should the participant adopt order encoding as a recall framework.

This forms the underlying impetus for experiments 3 and 4. Immediate recall with no delays should result in a greater degree of order encoding if models such as trace redintegration or the order encoding model are correct. In the former, order should naturally be output from the phonological loop (see TODAM; Murdock, 1997), and in the latter, mental resources will not be diverted to other tasks, alleviating the effects of interference on both item and order information.

Experiment Three

The third experiment aimed to replicate the design of Experiment 1 in all but the recall duration. In this study, participants were asked to recall the words from each list immediately after the disappearance of the final word in each list. Although it was expected that there would be some disruption in order encoding due to the recency effect (which was not a consideration with the previous filled delay), it was also expected that words were more likely to be remembered in order. Should order retention be demonstrated, it is also supposed on the basis of results from Bridges & May (1997) and two pilot studies that HF and semantically categorised lists will show evidence of more such encoding than LF and unrelated lists, as trace redintegration for HF words is enhanced. Another prediction was that the free recall advantages for high frequency and semantically related words, irrespective of order encoding, would remain in short-term recall, and that this effect would still be additive.

Method

Participants were 64 University students from Aston University, aged between 18 and 24 years of age. They were drawn from the Psychology undergraduate corpus and completed testing as part of their course requirement. None of the participants had taken part in previous experiments.

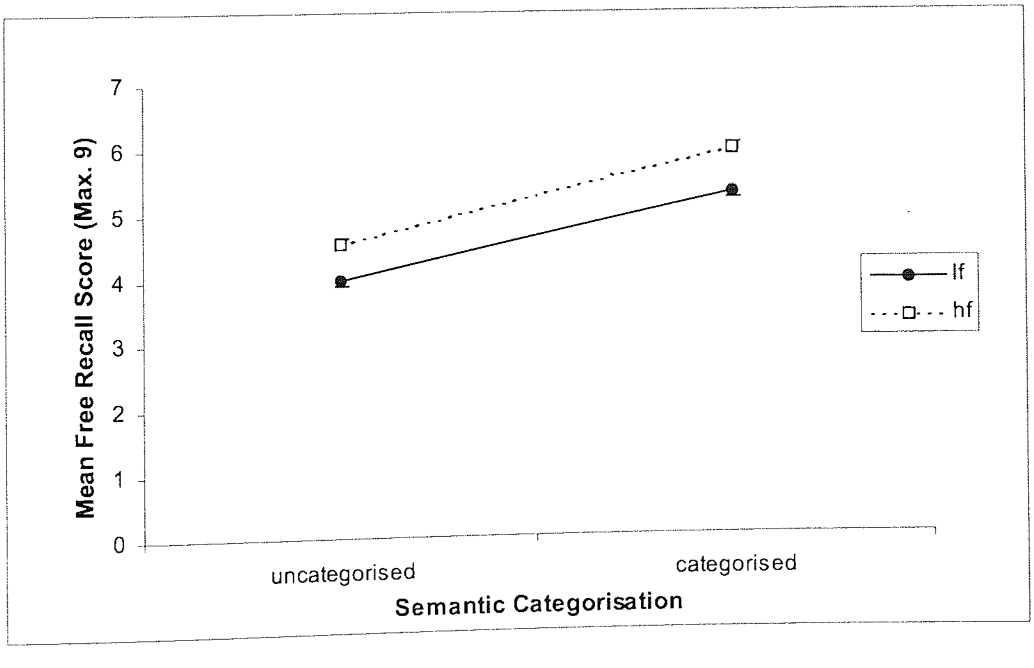
Design. The experiment had a 2 x 2 design. Independent variables were frequency (high or low) and grouping by semanticity (categorised or uncategorised). Measurement of performance was conducted on a between-subjects basis. Dependent variables were proportion of correct responses for free recall tasks and input-output correspondence.

Materials & Apparatus were identical to those reported in experiment 1. The nine-word lists used in this experiment can be found in Appendix One.

Procedure. Participants were assigned to condition by the order in which they presented themselves at the laboratory. They were instructed that they were to receive a memory test, and that their task was to remember the words which they saw upon the computer screen. They were told they must recall as many of the words as they could from the list. Instructions for recall were written upon the recall sheet such that participants were asked to write down the words as they occurred to the participant at time of recall, and that one word was to be written on each line. At no point was any mention of the order of the items made.

Visual presentation followed the same format as in experiment 1. After the final word from each list had been presented, there was a 500 msec blank screen, followed by the prompt, which was a line of five asterisks. Participants were given a minimum of one minute to write their answers down with no maximum limit. Once the participant was ready to proceed, the space bar was pressed, and the next trial began automatically. In total, each participant viewed eight of the word lists.

Figure 2.4: Mean free recall score per list by semantic categorisation and word frequency (Experiment 3).



Results

Free Recall

Mean free recall overall was 4.94 (SD= 1.32), suggesting that marginally fewer words were recalled on average than in the delayed recall experiment. The scores by condition are displayed in figure 2.4. It appears that high frequency words are better recalled than low frequency ($F_{(1, 60)} = 25.31$; $p < 0.001$), and that semantically related lists are better recalled than semantically unrelated lists ($F_{(1, 60)} = 123.36$; $p < 0.001$). Means collapsed across list frequency and across semantic categorisation are available in table 2.3. There was no interaction between the two variables ($F_{(1, 60)} = 0.23$), and indeed, it can be seen that the recall enhancement effect is additive for each variable: HFC lists are enhanced to a greater degree than either of the two list types with only one of the improving factors, and the list type with neither (LFN) is the worst.

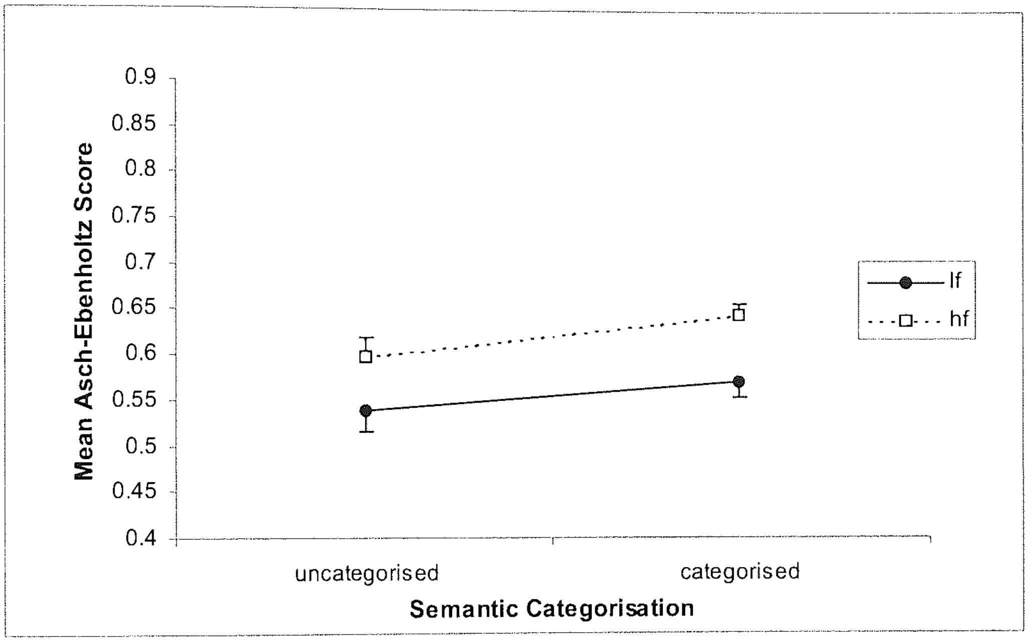
Table 2.3: Main Effects of Frequency and Semantic Relatedness on Free Recall and Input-Output Correspondence.

	FREQUENCY		SEMANTIC CATEGORISATION	
	LF	HF	UNRELATED	RELATED
MEAN FREE RECALL (MAX 9)	4.63	5.25	4.26	5.63
SD	*1.26*	*1.31*	*1.18*	*1.08*
MEAN ASCH-EBENHOLTZ	.55	.62	.57	.61
SD	*.23*	*.20*	*.25*	*.16*

Input-Output Correspondence

Mean Asch-Ebenholtz score across all lists was 0.58 (SD= 0.21), which differed from chance (one-sample $t_{(511)} = 9.32$, $p < 0.001$). Figure 2.5 shows mean Asch-Ebenholtz scores by condition, and table 2.3 gives mean scores by variable. Lists comprising of high-frequency words were remembered in a closer approximation of the original order than lists of low frequency words ($F_{(1, 60)} = 15.16$, $p < 0.001$) and semantically related lists had greater input-output correspondence than non-categorised lists ($F_{(1, 60)} = 4.94$, $p < 0.05$). There was no interaction between the two variables ($F_{(1, 60)} = 0.17$). Effects of semantic categorisation and of word frequency have therefore been demonstrated upon the implicit retention of order information.

Figure 2.5: Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency (Experiment 3).



Error Analysis

The total number of errors recorded was 100 from a total of 4096 words (2.4%). Table 4 shows a detailed breakdown of errors, which were scored in the same way as before. A chi-square test revealed significant differences between conditions for total number of errors per condition ($X^2 = 12.08$; df =3; $p < 0.01$). Individual analyses comparing error types across condition against expected chance frequencies revealed that semantic errors differed by condition ($X^2 = 17.49$; df = 3; $p < 0.001$), but that neither PLIs or phonological errors did. Table 4 shows that the bias towards semantic errors occurs very strongly in the semantically categorised lists, and barely at all in the non-categorised lists.

Table 2.4: Frequency of response errors scored by type and experimental condition.

ERROR TYPE	LFN	LFC	HFN	HFC	**TOTAL**
SEMANTIC	2	19	7	17	**45**
PHONOLOGICAL	2	4	2	3	**11**
PLI	2	3	5	6	**16**
OTHER	8	8	5	7	**31**
TOTAL	**14**	**34**	**19**	**33**	**100**

Discussion

The results for experiment three demonstrate once again how the factors of word frequency and semantic categorisation can have an influence on the recall of individual items from short-term memory. The effects of semantic and frequential categorisation on immediate recall appear to be remarkably similar to those for delayed recall, and taken with the improved level of order retention in the current design, suggest again that order encoding in itself does not have a direct influence on the advantage conferred on *free recall performance* by frequency and semantic effects. However, semantic and frequency effects were found on pure measures of order performance. The fact that word list effects for free recall performance are present across both immediate and delayed recall, and that word list effects on order only occur in immediate recall may raise some serious implications for the order-encoding hypothesis.

The results replicate those of Delosh & McDaniel (1996) in that order encoding is enhanced by the use of high- frequency words as opposed to low. This is notably different from delayed recall, illustrating that order encoding is available for immediate recall in a way that it is not for delayed recall. This reaffirms the connection between phonological rehearsal and order retention, as both appear to decay over delays of more than a few seconds (Baddeley, 1986; Baddeley & Hitch, 1974). This explanation is tested further in chapter three.

The fact that few phonological errors were found in this delayed recall design supports the notion that the phonological code rapidly decays in short term memory. What is of interest in the present experiment is that semantic errors continued to be made over a filled delay for semantically categorised lists, whereas non-categorised lists did not induce these errors. As before, it is supposed that the reasons behind this effect are due to what Walker (1999) terms semantic redintegration, where the specific target word representations are being degraded after presentation, and the matching process at recall finds the best semantic fit for the available degraded representation. Categorised lists would have more evidence pointing to related words than unrelated words, and would perhaps therefore be more susceptible to the recaller choosing a wrong, but closely related word than for non-categorised lists. It could also be that participants were more likely to present guesses for the semantically categorised lists, as there would always be partial evidence (i.e. related meaning) that the guess was correct. For the non-categorised lists, this would not be the case and the participant would be less likely to feel that his interpretation of the degraded trace was likely to be correct.

Where these results are most interesting is where they show an effect of semantic categorisation for the order in which items are recalled. This result fits the hypothesis that long term factors other than frequency play a role in order encoding, and indeed, support the notion that semantic factors can influence interpretation of memory traces in the short-term store. It appears that a relative measure of order encoding is able to detect a semantic effect on order that positional measures have missed (Baddeley, 1966; Poirier & St-Aubin, 1996, 1999b).

As the positive effect of semantic similarity on order encoding is rare (if not non-existent; St-Aubin & Poirier, 1999b), a replication of this effect was sought. Experiment four sought to replicate experiment two, using mixed lists, but as in experiment three, recall was immediate rather than delayed. It was supposed that in addition to the low-frequency advantage described by Delosh & McDaniel (1996), order retention would also be improved by semantic categorisation as it had been for experiment three.

Experiment Four

Method

Participants were 48 University students from Aston University, aged between 19 and 29 years of age. Participants were drawn from the student participant pool at Aston University. None of the participants had taken part in previous experiments.

Design. For analysis of free recall scores, a mixed 2 x 2 design was used with list categorisation as independent measure and word frequency as the within-subjects factor. For input-output correspondence, a between-subjects t-test was employed.

Apparatus & Materials. Apparatus was the same as that used in the previous experiments. The mixed word lists used were the same as those in experiment 2. These lists are available for inspection in Appendix Four.

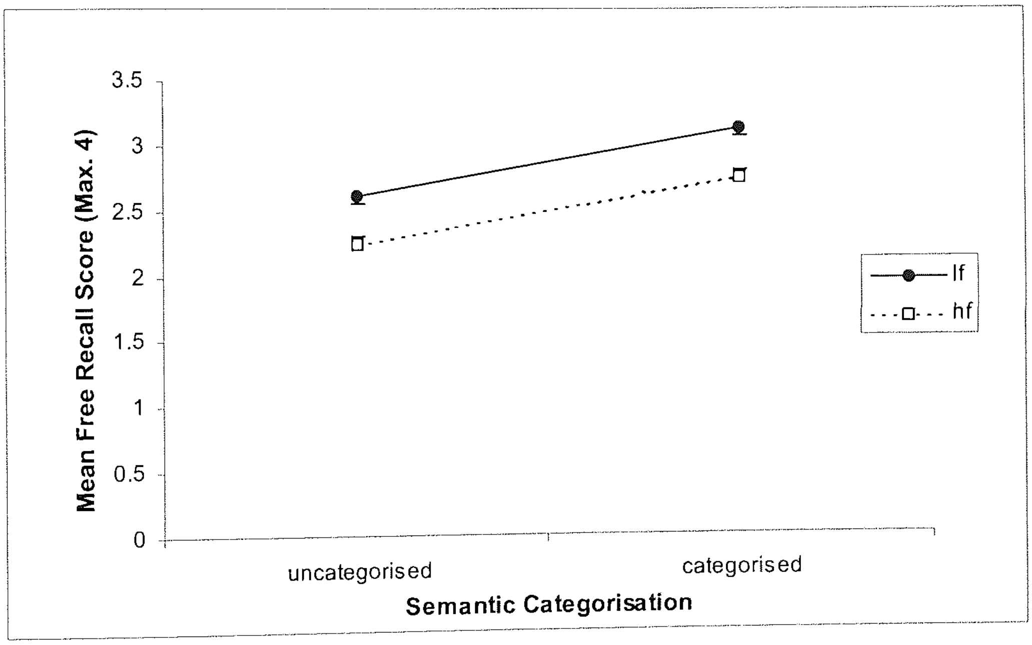
Procedure. Participants were allocated to the two conditions alternately. Each individual saw either the 12 mixed semantically categorised (MC) lists or the 12 mixed semantically non-categorised (MN) lists. After the final word had been presented, there was a 500 msec delay, followed by the recall prompt. All other details were as described in Experiment 2.

Results

Free Recall

Mean free recall score overall was 5.37 words per list (SD= 1.41). Mean recall scores were entered into a mixed ANOVA model, with semantic categorisation as an independent factor with two levels (categorised and uncategorised) and word frequency as a repeated measures factor with two levels (HF and LF). Scores broken down by list type and frequency are displayed in figure 2.6, which shows a consistent LF advantage over HF words. The mean number of words correctly recalled for LF lists was 2.87 (SD= 1.05), and for HF lists, 2.49 (SD= 1.04). This small difference was reliable ($F_{(1, 46)}$ =31.14; p< 0.001). The figure also demonstrates a strong semantic category advantage ($F_{(1, 46)}$ = 29.69; p< 0.001). Recall for semantically categorised lists (mean = 5.88; SD= 1.27) was superior to that for unrelated lists (mean = 4.85; SD= 1.35). No interaction between semantic relatedness and word frequency was found ($F_{(1, 46)}$ = 0.00).

Figure 2.6: Mean free recall score per list by semantic categorisation and word frequency (Experiment 4).



Input-Output Correspondence

The mean Asch-Ebenholtz score overall was 0.66 (SD= 0.21), which was significantly different from chance (one sample $t_{(575)}$ = 17.96; p< 0.001). Mean scores for MC lists (mean = 0.66; SD= 0.18) and MN lists (mean = 0.66; SD= 0.24) did not differ significantly from one another ($t_{(46)}$ = 0.08).

Error Analysis

From a total of 3456 words displayed in the experiment as a whole, 123 incorrect responses were recorded (3.6%). As before, these errors of commission were classified into phonological, semantic and previous list intrusion errors, and can be seen in tabulated form in table 5. A chi-squared analysis shows no significant relationship between list type and error type (X^2 = 0.2; df =1). Individual analyses for semantic errors did show significant effects of list type also (X^2 = 14.4; df =1). More semantic errors occur in the categorised lists and more previous list intrusions occur in the non-categorised lists. None of the other error types differed across lists.

Table 2.4: Frequency of response errors scored by type and experimental condition.

ERROR TYPE	MN	MC	**TOTAL**
SEMANTIC	8	32	**40**
PHONOLOGICAL	14	10	**24**
PLI	14	8	**22**
OTHER	23	14	**37**
TOTAL	**59**	**64**	**123**

Discussion

Results of experiment 4 confound expectations somewhat. Whereas in pure lists, semantic categorisation appeared to encourage the use of order information in the recall phase, in mixed lists there was no difference between categorised and uncategorised lists. In contrast to experiment 2, however, both list types exhibited a larger degree of order retention overall, and it could be that these relatively high levels of remembered order reached a ceiling which prevented any further influence of semantic categorisation from being detected. If one assumes that a recency effect did occur in this paradigm with the last few items on the list recalled whilst they were still fresh in the participants mind, then one would expect the Asch-Ebenholtz measure to be reduced as the order for the first few items is effectively reversed. Unfortunately, due to the relatively short list lengths, it is not possible to test this by removing the last few words on the target list from the analysis.

Alternatively, it is possible that semantic effects on order encoding do not occur in lists of mixed frequency but that they do in pure lists. No existing theory could reasonably explain why this might be so, unless it is an artefact of the composition of lists used in this specific case. The results discussed here are very much in opposition to the literature, which suggests that such effects of semantic categorisation are not at all robust (see Poirier & St-Aubin, 1995; St-Aubin & Poirier, 1999). Hence, if we are to place strong faith in these results, further replication of this effect must be sought.

Experiment Five

Delosh & McDaniel (1996) found average order encoding in their experiments to be much higher than those found in the current experiments (on average, they found order encoding to be around 0.70, compared to the figures of around 0.62 reported in experiments three and four above; and values of around 0.51 in the first two experiments in this chapter). Experiment five attempted to assess whether the strong order encoding effect in the Delosh & McDaniel experiments was due specifically to their participants' knowledge that order was going to be assessed. Specifically, a design was set up in which participants were not overtly asked to encode order information in a free recall task, but were afterwards given an order reconstruction task in which no recall of items was necessary, but the order of items was. It was hypothesised that although free recall would remain at a similar level to that of experiment one, input-output correspondence would increase dramatically.

This design returns to a delayed recall paradigm, in order that the original Delosh & McDaniel (1996) paradigm is replicated. In view of Crowder's (1982) assertion that the differences between long and short-term memory are not robust (Crowder & Neath, 1991), and that memory processing is dependent on natural language processes, it would be interesting to see if effects found previously for immediate memory can be induced for delayed recall.

Method

Participants were 20 University students from Aston University, aged between 18 and 28 years of age. None of the participants had taken part in previous experiments.

Design. The experiment had a 2 x 2 design. Independent variables were frequency (high or low) and grouping by semanticity (semantically categorised or random). Measurement of performance was conducted on a repeated measures basis. Dependent variables were proportion of correct responses for free recall tasks, input-output correspondence and a post-trial measure of order reconstruction. This method of measuring order was chosen in order to provide a valid framework with which to compare the results of the present experiment and that of Delosh & McDaniel (1996). In fact, of all experiments reported in this thesis, experiment five represents the most precise replication of their design.

Materials. The nine-word lists were identical to those found in experiment one, and are reproduced in Appendix One.

Apparatus. Stimuli were displayed on an IBM compatible PC, and presented through the Windows interface, using a program written in Microsoft Visual Basic 5.0.

Procedure. Participants were instructed that they were to receive a memory test, and that they were to perform two tasks. The first was to remember the words which they saw upon the computer screen. Instructions for recall were written upon the recall sheet such that participants were asked to write down the words as they occurred to the participant at time of recall, and that one word was to be written on each line. No mention of the order of the items was made with respect to the first task, but participants were told that after each trial, they would be asked to perform the second task, which was to reassemble the words they had seen in the correct order.

The procedure for presentation of the words and that for the first task replicated the procedure of experiment one, including a forty-second filled delay in which participants were required to solve maths problems. There were two differences from the experiment one design, however: the first was that the present experiment followed a repeated measures design, in which the order of presentation of lists was randomised across list types, with no more than two lists of any type occurring in succession. The other difference was the inclusion of the order encoding task.

After each recall phase participants had their response booklets removed, and were given a sheet of paper containing the words which had just been presented to them. Words were arranged at random non-linear locations on the page to reduce the possibility of influencing order choices. Participants were then handed another answer booklet in which they were asked to write the words from the sheet in the exact order that they had been presented. Order reconstruction scores were measured simply as the proportion of words which were recorded at the correct location. Participants were then handed back their free recall answer booklets and the experiment continued with the next trial in the series.

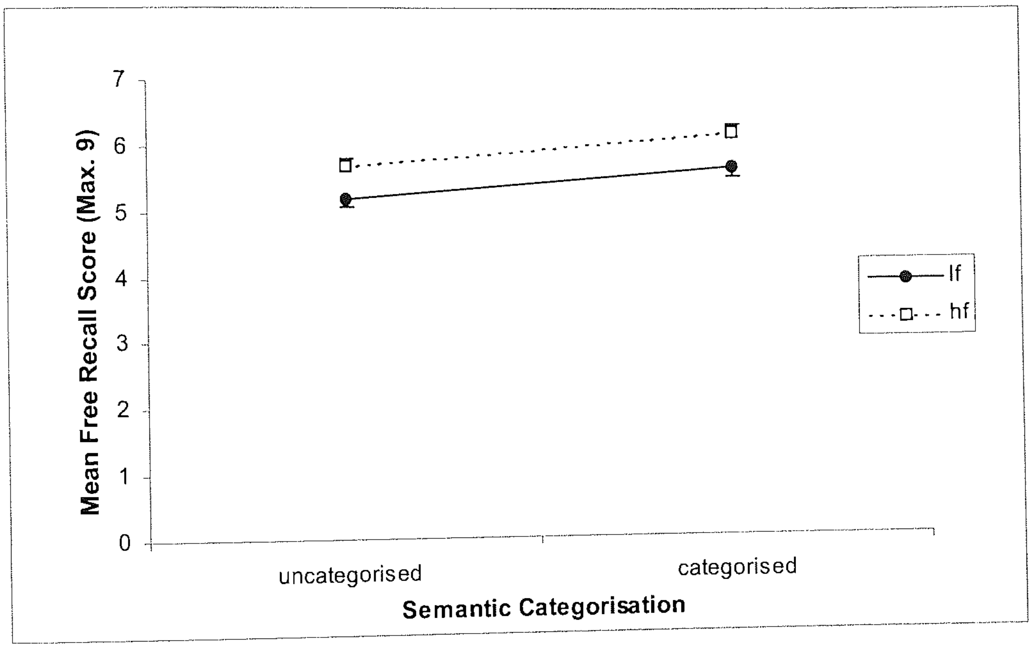
Results

Free Recall

Mean free recall across all lists was 5.70 words (SD= 1.59). Free recall scores by word type are shown in figure 2.7. The expected advantage for recall of high frequency words over their low frequency counterparts was significant ($F_{(1, 19)} = 14.32$, $p< 0.001$), and the recall enhancement for semantically categorised lists over unrelated lists was significant ($F_{(1, 19)} = 9.98$, $p< 0.005$). There was no interaction between the two variables ($F_{(1, 19)} = 0.03$). Post-hoc analysis of homogeneous subsets of list types revealed that LFN and HFC lists were each distinguishable from all other lists. HFN and LFC lists were found to elicit statistically equivalent recall performance (see figure 2.7). This again supports the notion of the effects of word frequency and semantic categorisation being additive. The results for free recall by word frequency have replicated the general pattern of results obtained by

Delosh & McDaniel (1996). In addition, the semantic categorisation effect from experiment one has been replicated.

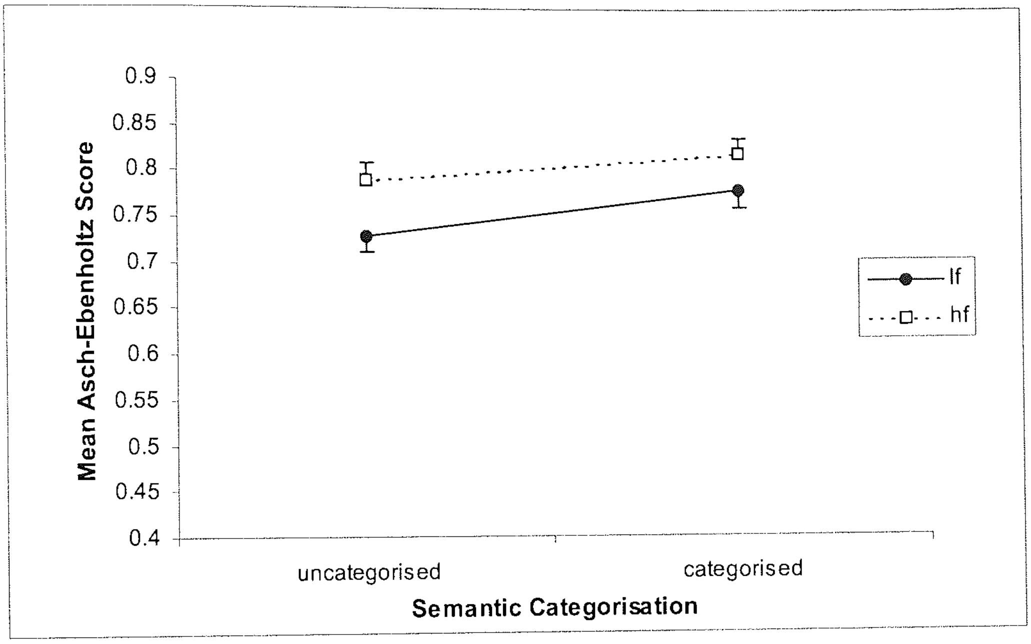
Figure 2.7: Mean free recall score per list by semantic categorisation and word frequency (Experiment 5).



Input-Output Correspondence

The mean Asch-Ebenholtz score across all lists was 0.78 (SD= 0.23), which did differ from chance (one-sample t $_{(639)}$ = 30.28). Figure 2.8 shows mean Asch-Ebenholtz scores broken down by list type. Analysis of Variance showed there to be significant effects of frequency ($F_{(1,19)} = 14.32$, $p < 0.001$) and semantic categorisation ($F_{(1,19)} = 9.98$, $p < 0.005$). There was no interaction between the two variables ($F_{(1,19)} = 0.03$). Delosh & McDaniel's (1996) findings of order encoding have therefore been replicated, as have findings from experiment one of this thesis, in that semantic categorisation does influence encoding of order in this paradigm.

Figure 2.8: Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency (Experiment 5)



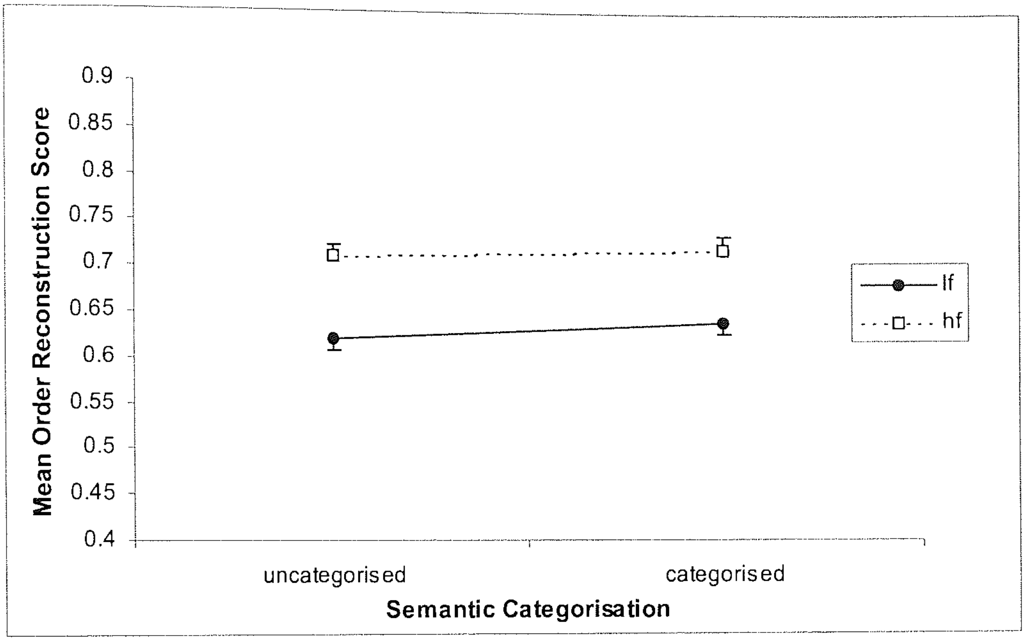
Order Reconstruction

The mean proportion of words within all lists which were placed in the correct order location was 0.67 (SD= 0.17)[7]. Figure 2.9 shows mean order reconstruction scores broken down by list type. Analysis of Variance shows a significant effect of frequency ($F_{(1, 19)}$ = 61.71, $p < 0.001$), with high frequency word lists outperforming low frequency lists. Semantic categorisation had no effect on order reconstruction ($F_{(1, 19)}$ = 1.01) and there was no interaction between the two variables ($F_{(1, 19)}$ = 0.04). Delosh & McDaniels finding that frequency affects order reconstruction has therefore been replicated, but in this task, semantic categorisation has no effect.

---

[7] This was equivalent to an Asch-Ebenholtz score of 0.86.

Figure 2.9: Mean order-reconstruction score per list by semantic categorisation and word frequency (Experiment 5)



Error Analysis

In total, 472 commission errors were recorded from a grand total of 5760 presented words (an 8.2% item error rate by word). Errors were classified in the same manner as for previous experiments. Error data are displayed in table 6. Chi-square tests were used to discover contingencies between error distributions and word list types. There was a significant difference between list types for the total number of errors committed ($X^2 = 52.08$, df = 9, $p < 0.001$). Individual analyses revealed that list types also differed in respect of the total number of semantic errors ($X^2 = 91.88$, df = 3, $p < 0.001$) and PLIs ($X^2 = 17.55$, df = 3, $p < 0.001$). Phonological errors did not appear to be contingent with list type, however ($X^2 = 6.93$, df = 3). The data for item errors therefore replicate the findings of experiment one.

Table 2.6: Frequency of response errors scored by type and experimental condition.

ERROR TYPE	LFN	LFC	HFN	HFC	TOTAL
SEMANTIC	12	87	23	85	**207**
PHONOLOGICAL	9	22	11	15	**57**
PLI	10	15	32	12	**69**
OTHER	38	40	25	36	**139**
TOTAL	**69**	**164**	**91**	**148**	**472**

Discussion

The free recall results obtained follow the same pattern as those obtained for experiment one, with the exception that the average number of words recalled for each list tended to be greater in the present experiment. Again, the post-hoc comparisons showed that the effects of semantic categorisation and item frequency were additive, which is to suggest that each of these effects is complementary to the other. More evidence for this assumption can be found in the results for order retention. In this case, the results did not mirror those of experiment one (in which semantic and frequential factors had no effect on the amount of order information in the output protocols). Instead, Delosh & McDaniel's (1996) results for the frequency effect in order encoding were replicated and extended in that a semantic effect was also found, as in the immediate recall paradigm. This replication of the semantic effect on order encoding reduces the possibility that the results of experiment three are due to chance, and suggests that there is a genuine effect which is elicited by this particular list composition and method of measuring order retention.

The results of the order reconstruction task did not match those findings, however. Although the frequency effect was found, as expected, there was no significant semantic effect on order reconstruction. This potentially casts doubt upon the hypothesised explanation of a semantic effect in order retention in the free recall task being based upon enforced participation in an order reconstruction task. If semantic categorisation does not affect order reconstruction, then there must be some other explanation for the semantic effect on input-output correspondence. This said, it should be noted that the results for order encoding are suggestive of a weak positive semantic effect, and it may be that the power of this experiment was not sufficiently high to detect such a trend.

It could be that the reason for the differences between the effect of semantic information on recall and the effect of the same on order reconstruction come down to the differences between recognition memory and recall. In the order reconstruction task, participants were able to recognise words which had been presented, whereas in free recall, they were asked to generate the items themselves. Thus, it could be that the basis of the semantic effects found previously are created by the search strategy employed by participants to activate the target items, and does not occur when the items are already activated. This would perhaps support a resource-based interpretation such as the order encoding hypothesis which predicts that once the words are activated in the participant's mental representation of the experimental trial, all remaining resources can be allocated to processing contextual information, such as order.

That the effects of semantic and frequential categorisation for immediate memory can be replicated across filled delays lend some support to monistic theories of human memory. The same patterns for free recall and order encoding appear to occur in each case, even though the effects in experiment five appear only when order encoding is made a useful or efficient strategy by the introduction of an order requirement later on. This implies that participants are able to pick and choose between strategies for free recall depending upon the requirements of the task and the information available to them. Order encoding and trace redintegration may each play a part in short-term recall, but neither offers a full account of the processes which occur during recall.

Summary

The results described in this study are varied, but consistent with respect to a model of human memory. Experiments one and two showed that information held in long term memory in one form or another can influence the recall of specific items, but has no effect on order encoding. In fact, with a filled delay between stimulus presentation and recall, no order information is encoded at all unless participants are forced to adopt an order encoding strategy (experiment five). In immediate recall paradigms, the pattern of free recall is the same as that from the delayed recall experiments, with both semantic and frequency categorisation influencing the number of items recalled. Order encoding is in evidence in experiments three and four, suggesting that order encoding may be used as a recall strategy, but that it decays at a rapid rate.

In experiments containing mixed lists, low frequency words were found to be advantageously recalled, in line with the findings of Delosh & McDaniel (1996). They suggested that the reason this effect occurs is as follows. High frequency words in pure lists take up less mental resources to process, and therefore more resources are available to process order information: this being the supposed reason for high frequency words having a recall advantage in pure lists. In mixed frequency lists, the low frequency words are hypothesised to reduce the amount of order encoding for high frequency words, and because low frequency words are more distinctly recognisable (e.g. Merry, 1980), they enjoy an item recall advantage once the order encoding between the two conditions has been neutralised. The evidence reported here suggests that this account is not adequate.

Experiments one and two, which are to all intents and purposes replications of Delosh & McDaniel's (1996) experiments, demonstrate that there is no noticeable encoding of order information at all in delayed recall. Even though this is the case, the LF advantage is still obtained in the mixed list design (experiment two). More damning for the order-encoding hypothesis is that experiments three and four show evidence of order encoding in immediate recall, and yet the pattern of the low frequency advantage remains, irrespective of whether order is overtly encoded or not. A further complication arises from the results of experiment five, in which an order encoding strategy can be adopted when it is appropriate to do so. Participants may adopt strategies which afford them the least cognitive exertion, and order encoding may be an efficient free recall strategy in experiment five because of the extra demand for order information to be retained. This idea is similar to that of Lewandowsky & Murdock (1989), in which associations between one item and the next can in themselves act as a cue to the next item in the list. Experiments one and two suggest that this strategy may not normally be efficient in delayed free recall when semantic representations of the items are available as cues, but experiments three and four suggest that order is used for immediate recall.

The error data from this series of experiments also conforms to a general pattern which extends across the duration of recall. In all five experiments, semantically categorised lists suffered more semantic errors, which is to be expected if one considers that interpreting a degraded memory trace may rely on semantic (as well as phonological) information, and that the unique information available for words in a categorised list is impoverished compared to lists in which items are more likely to be semantically distinct. Previous list intrusions occurred in HFN lists in experiment one and in non-categorised mixed lists in experiment two. This is suggestive of a bias towards high frequency words when redintegrating decaying traces, but largely when there is no other relational constraint imposed upon the list. When semantic categorisation information is inherent in the list, individuals are less likely to insert a word from a previous list.

Duration of Retention

The classical distinction of memory into short and long-term has evolved into a framework where an articulatory loop for phonological processing (Baddeley, 1992) is supplementary to a larger, more monistic form of memory processing (Cowan, 1988; Crowder, 1982). The evidence presented here supports such a framework.

Firstly, it has been demonstrated that the effects of long-term information such as semantic categorisation and word frequency have similar effects on immediate and delayed free recall (compare fig. 2.1 against fig. 2.4, and fig. 2.3 against fig. 2.6). Estes (1972) found this similarity between short and long term recall in serial position effects across even longer delays of 2 and 10 minutes, and Crowder & Neath (1991) report even longer-term similarities between short and long-term recall.

What is interesting and important about the research reported here is that we can see that the usage of semantic and frequency information does differ across one dimension. Although overall recall profiles are strikingly similar across immediate and delayed recall, the use of order encoding as a strategy is only naturally observable at immediate recall. This finding not only supports the common finding of articulatory loop involvement in immediate recall, but also lends weight to accounts of order information being implicitly encoded in the phonological representation of the loop (e.g. Wicklegren, 1977a).

Furthermore, it can be demonstrated that the process of accessing items in the phonological loop in order to recall them is not a strictly modular process (Fodor, 1983), but can be itself influenced by the presence of information about the words to be recalled. The most parsimonious explanation for these effects is that articulatory traces of degraded phonological information are accessed, and that semantic information and word frequency influence the choice of possible interpretations of the trace. This choice is then fed back into the convoluted code of phonological information, and the process begins again. For longer-term recall, all phonological information has already been degraded, and the individual relies on other strategies, such as semantic searching and sophisticated guessing. Inter-item associations may be available to allow the encoding of order information over longer periods in which phonological information has become too degraded to interpret. This strategy may be less efficient than others over extended periods of recall, and consequently is not used unless it becomes efficient. The next chapter attempts to continue this line of analysis, and attempts to assess the extent of phonological involvement in order encoding in immediate memory.

Chapter Three: The Role of the Phonological Loop in Order Encoding

Chapter two demonstrates a number of important points about the use of order information in free recall. Firstly, it is clear that factors from long-term memory such as semantic information and word frequency affect the amount of order information implicitly recalled. Secondly, it appears to be the case that this effect is mainly apparent when recall is required immediately after presentation. A delay of 40 seconds between list presentation and recall is enough to disrupt those effects. This pattern is consistent with the notion explored in chapter one that order is dependent on phonological processing (cf. Wicklegren 1965, 1977a).

There are a number of sources of evidence linking order encoding to phonological memory. It has been accepted across the spectrum of theories of short term memory that phonological information stored in a phonological store degrades rapidly (Baddeley, 1966; 1997; Schweickert, 1993; Hulme *et al*, 1999; Nairne, 1990*a*). The pattern of order information contained in recall output is consistent with this finding in that for immediate recall, order information is largely intact, but after a delay much longer than the 2 seconds for which Baddeley & Hitch (1977) suggested phonological memory persists, evidence of order encoding is obviated.

Another source of evidence for a phonological role in order retention comes from Wicklegren (1965). Conrad & Hull (1964) showed that the recall of items which sounded similar to each other was worse than that for items which were not acoustically similar. Wicklegren extended these findings to show that it was with recalling the order of these items that participants had difficulty, and that acoustic similarity even hinted at enhancing item recall.

In the light of this evidence, it was decided to analyse the relationship between order encoding and LTM factors with respect to phonological processing. A search of the literature reveals that thus far, little evidence has been collected on the relationship between these factors. The standard paradigm used in memory research is that of serial ordered recall, which does not allow the careful scrutiny of differential effects on item and order information (e.g. Poirier & St-Aubin, 1995, 1996). Furthermore, experiments examining the role of word frequency and semantic relatedness have generally not controlled for speech rate (e.g. Hulme, Maughan & Brown, 1991; Baddeley, 1966b; Wethrick, Huttenlocher & Newcombe, 1976; Bourassa & Besner, 1994). The experiments reported in this chapter aimed to redress these issues.

If it is indeed the case that phonological processing influences the use of order information in short term recall, then one would expect articulatory suppression to disrupt the mechanisms which process and utilise order information. Articulatory suppression has been used as a tool in memory research since the development of Baddeley's working memory model. The principle is that although overt speech is not required for verbal recall, the same mechanisms are used in processing target items. When these mechanisms are put in use by another task, such as concurrently speaking, processing of the target items in a phonological manner is disrupted or eradicated. This occurs with both visually presented and spoken words, the former because items cannot be converted into phonological representations, and the latter because material is already filling the phonological store (Baddeley, 1997).

The technique of articulatory suppression has had its critics in the past. For example, Parkin (1988) suggested that performance is impaired because attentional processes are being used up by the task, rather than being used to remember the words presented. However, control tasks which ostensibly require the same attentional demands (such as tapping in the same rhythm as the spoken task) generally do not exhibit the same degree of influence on recall performance (Baddeley, Lewis & Vallar, 1984). There is also evidence from neuropsychological studies which demonstrates that the task specifically affects the articulatory loop (cf. Vallar & Baddeley, 1984).

Experiment Six

The aim of this experiment was to ascertain the extent of the involvement of the articulatory loop in the semantic and lexical processing of words for recall. It was expected that order encoding and subsequent recall output would depend on the articulatory control component of the phonological loop, and therefore that the input-output correspondence of recall in lists where articulatory processes are otherwise engaged will be lower than previously found. Information gathered on the relationship between phonological processes and input-output correspondence will provide useful insight into the way in which order information is encoded and used.

Given that word frequency and phonological processing are linked to lexical activation at a sub-semantic level (e.g. Ellis & Young, 1988; Harley, 1997), one would expect also that articulatory suppression would disturb the recall of low frequency lists. Articulatory suppression should have no differential effect across the semantic categorisation variable under the redintegration hypothesis, as semantic information is hypothesised to affect only the interpretation of traces in which order has already been encoded (Walker & Hulme, 1999).

Method

Participants were 40 University students from Aston University, aged between 18 and 25 years of age. They were drawn exclusively from the Psychology undergraduate intake and completed testing as part of their course requirement. None of the participants had taken part in previous experiments.

Design. The experiment had a 2 x 2 x 2 within-subjects design. Independent variables were suppression (articulatory or a hand-tapping control), word frequency (high or low) and grouping by semanticity (categorised or uncategorised). Dependent variables were proportion of correct responses for free recall scores, and input-output correspondence to assess the degree of order encoding in each recalled list.

Materials. Word lists were constructed by drawing words from the Oxford Psycholinguistic Database (Coltheart, 1981). 256 bisyllabic words were chosen such that their mean values for age of acquisition (in years), concreteness (scaled from 1 to 7) and familiarity (1 to 7) did not differ from any individual items by a value of more than 1.2 for any individual variable. These words were placed in four different list types. Two of these lists were loosely categorised semantically (for example, all words were in some way connected with water- *haddock, canoe, coral, etc.*). The other two lists were chosen to be semantically unrelated. All lists were eight words long. The validity of the categorisations was corroborated by presenting the lists to 13 participants who were told to rate the degree of semantic categorisation for each list on a scale from one to seven. Mean categorisation rating for categorised lists was 5.90 (SD= 0.87) compared to 1.89 (SD= 1.00) for uncategorised. Semantically related lists were judged significantly more categorised than non-categorised lists (*paired* $t_{(12)} = 26.29$; $p < 0.001$).

In addition to lists being divided by semantic category, they were also divided in terms of frequency. Words in the high frequency list had a frequency greater than 50 per million, and low frequency words had a frequency of 10 or less per million (Francis & Kuçera, 1982). This gave four list types: low frequency, non-categorised (LFN); low frequency, categorised (LFC); high frequency, non-categorised (HFN) and high frequency, categorised (HFC). There were eight lists of each type, making 32 eight-word lists in all. The words used in the present experiment are listed in Appendix Five.

Apparatus. Stimuli were displayed on an IBM compatible PC, and presented through the Windows interface, using a program written in Microsoft Visual Basic 5.0.

Procedure. Two pools of 16 lists were created by randomly choosing four of each list type. Lists were counterbalanced such that each pool was used an equal number of times in suppression and control conditions. Furthermore, the order of presentation of list pool and suppression conditions was counterbalanced across participants in a Latin square design. After each cycle of four participants, the list pool was randomly split again. Each time the pools were created, the order of lists and the order of words within lists was randomised.

Participants were given full written instructions on the task they were asked to perform. They were told that they were to receive a memory test and that words would appear on the screen. During presentation, participants were asked to perform one of two tasks: pronouncing the words 'one-two, three-four' in a regular rhythm, or tapping in the same rhythm on the desk. This phrase was chosen because research has shown that simple repetition of one syllable alone is sometimes not sufficient to disrupt phonological loop processes (Macken & Jones, 1995). Macken & Jones' study was based on the irrelevant speech effect, in which participants are presented with an unattended audio stream whilst participants attempt to remember a list of items. If a single syllable was repeated over and over, this had little effect upon recall performance. However, if a stream of changing syllables was presented, recall is disrupted (Jones & Macken, 1995). Macken & Jones (1995)extended this finding to articulatory suppression, where participants asked to voice seven syllables repeatedly fared worse in a serial recall task than did those asked to repeat a single syllable. Whilst at first glance, it may appear that this is an attentional effect rather than a phonological loop effect, Neath, Suprenant & LeCompte (1998) have recently found that irrelevant speech can eliminate the word length effect for both visual and auditory items. This suggests that changing state information is as relevant for the phonological loop as it is for unattended speech, given that Baddeley (1997; Baddeley & Hitch, 1974) cite the word length effect as one of the primary pieces of evidence in favour of the existence of the phonological loop.

The concurrent task commenced before the first word was presented on screen, and stopped when the recall prompt appeared. Participants were then asked to write down the words that they remembered on an answer sheet. At no point was any mention of the order of the words made. If a participant asked whether the words had to be recalled in order of presentation, they were told that this did not matter[8]. Instructions given to participants for experiments in this chapter are given in appendix six.

The experiment was arranged in two halves, each consisting of a 4-trial practice block, followed by the experiment proper. Half of the participants were given the articulatory suppression condition first, followed by the tapping condition, and the other half received the opposite treatment. Each trial within a block consisted of a prompt to begin the suppression or tapping task, followed after a two second delay by the first word in the list. Words were presented in the centre of the screen, covering between approximately four and seven visual degrees. These measurements are approximate, as no attempt was made to fix the participant's head position. The stimuli were displayed on screen for 1500*msecs* each, followed by a 500*msec* blank before the next word appeared. After the eighth word, a 500*msec* delay preceded the prompt for participants to begin recall. This was a row of five asterisks, so as not to confuse the participant into thinking that the cue was yet another word.

Recall was carried out on a separate answer sheet for each trial, with instructions that there should be one word per line, and words should be written directly underneath the last as they came to memory. Each participant chose when they would progress to the next trial by pressing the space bar. The minimum possible recall period was 45 seconds.

---

[8] Participants were told that it did not matter only if they asked whether the order was important. This did not form part of the instructions, and in practice, few actually inquired about the order of responses. Therefore, this was not deemed a substantial bias against ordered recall.

Results

Results were analysed in terms of free recall scores and of input-output correspondence (Asch & Ebenholtz, 1962a). Input-output correspondence scores were obtained by scoring each adjacent pair of responses in the recall protocol for relative order in exactly the same manner as in chapter two. All ANOVA summary tables for experiments in chapter three are given in appendix seven.

Free Recall

The grand mean for free recall scores was 4.73 words (SD= 1.56). Free recall scores by list type and suppression condition are shown in Fig. 3.1. A 3-way repeated measures ANOVA was carried out on the free recall scores. There was found to be no 3-way interaction between articulatory suppression, word frequency and semantic categorisation ($F_{(1,39)}$=0.23). Main effects were found for each of the three variables. Mean values for the main effects can be found in table 3.1. As expected, articulatory suppression was found to have reduced the number of words remembered compared to the tapping control ($F_{(1,39)}$= 68.36, $p< 0.001$). In addition, categorisation of the lists in terms of both word frequency ($F_{(1,39)}$ = 14.26, $p< 0.001$) and semanticity ($F_{(1,39)}$ = 120.88; $p< 0.001$) was found to yield better recall results than non-categorised lists.

Table 3.1a: Main effects of frequency, semantic categorisation and articulatory suppression on free recall.

	ARTICULATORY SUPPRESSION		FREQUENCY		SEMANTIC CATEGORISATION	
	SUPPRESS	CONTROL	LF	HF	UNRELATED	RELATED
MEAN FREE RECALL SCORE (MAX 8)	4.3	5.01	4.52	4.87	4.30	5.09
SD	*1.59*	*1.47*	*1.52*	*1.58*	*1.54*	*1.48*

Table 3.1b: Main effects of frequency, semantic categorisation and articulatory suppression on input-output correspondence.

	ARTICULATORY SUPPRESSION		FREQUENCY		SEMANTIC CATEGORISATION	
	SUPPRESS	CONTROL	LF	HF	UNRELATED	RELATED
MEAN ASCH-EBENHOLTZ SCORE	.40	.50	.43	.47	.42	.48
SD	*.29*	*.27*	*.29*	*.28*	*.30*	*.27*

Figure 3.1: (a) Mean free recall score per list by semantic categorisation and word frequency under articulatory suppression (Experiment 6).

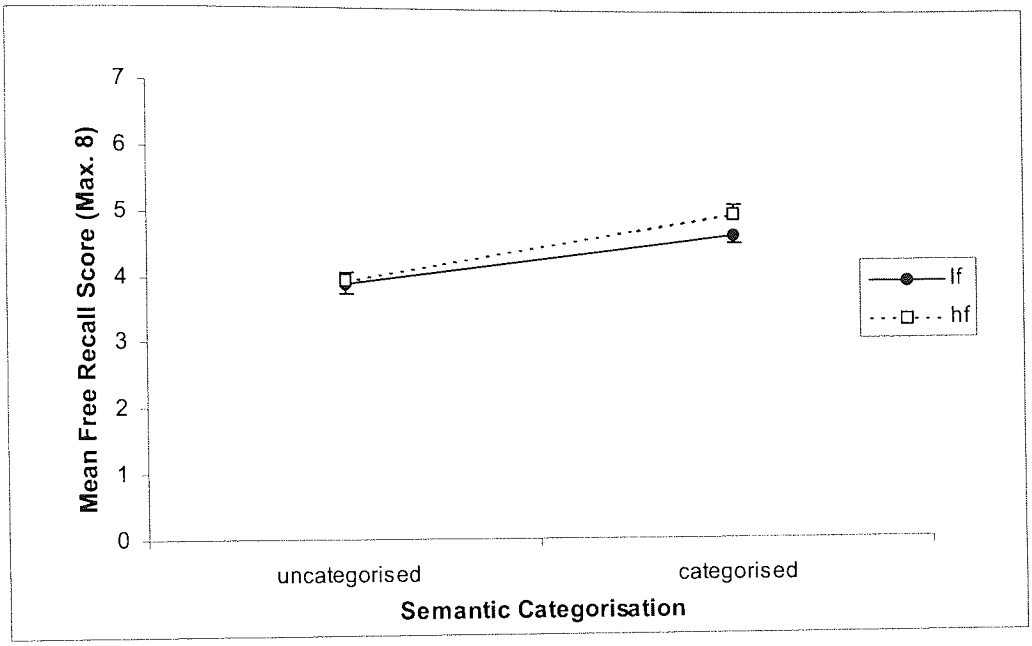
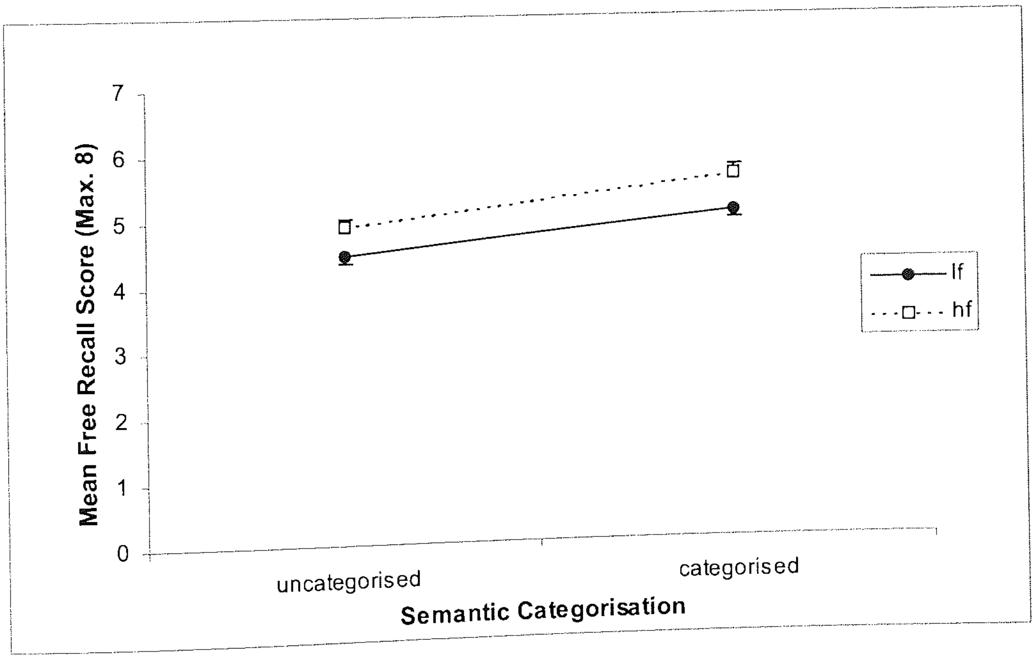


Figure 3.1: (b) Mean free recall score per list by semantic categorisation and word frequency under control condition (Experiment 6).



A marginal two-way interaction was found between articulatory suppression and word frequency ($F_{(1, 39)} = 3.36$, $p < 0.08$). As can be seen from Fig. 3.2, it appears that under the control condition, high frequency words are better recalled than low, but under suppression, the frequency effect is greatly diminished. The interaction between suppression and semantic categorisation was also marginal ($F_{(1, 39)} = 3.37$, $p < 0.08$) with a suggestion that suppression reduced the semantic category advantage. The two-way interaction between frequency and semantic categorisation was not significant ($F_{(1, 39)} =$ 2.51).

Figure 3.2: Interaction between word frequency and articulatory task for free recall score (Experiment 6).

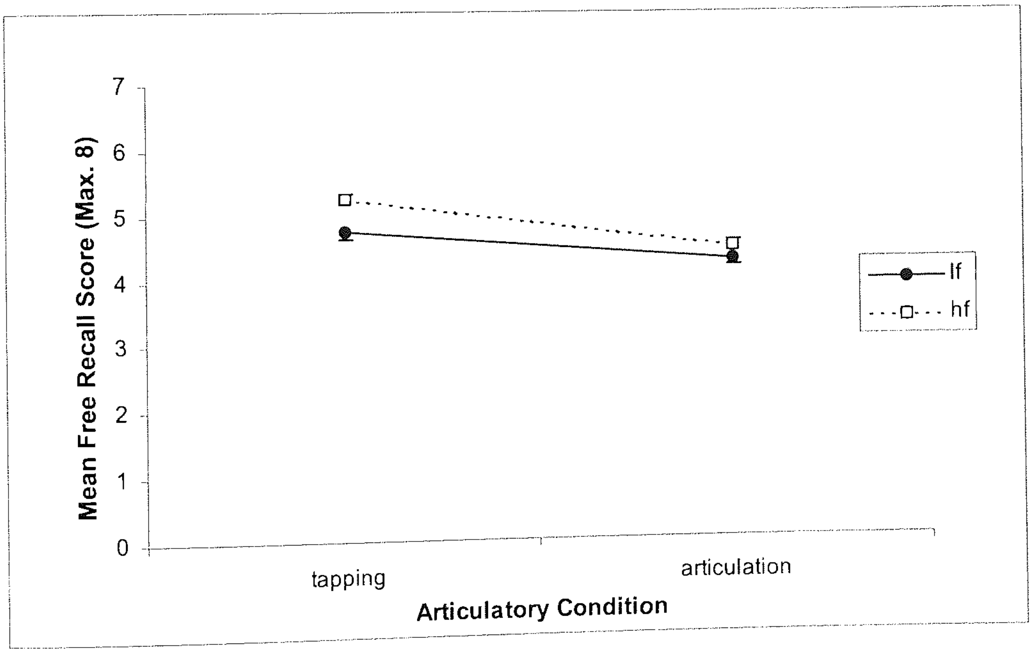


Figure 3.3:(a) Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency under articulatory suppression (Experiment 6).

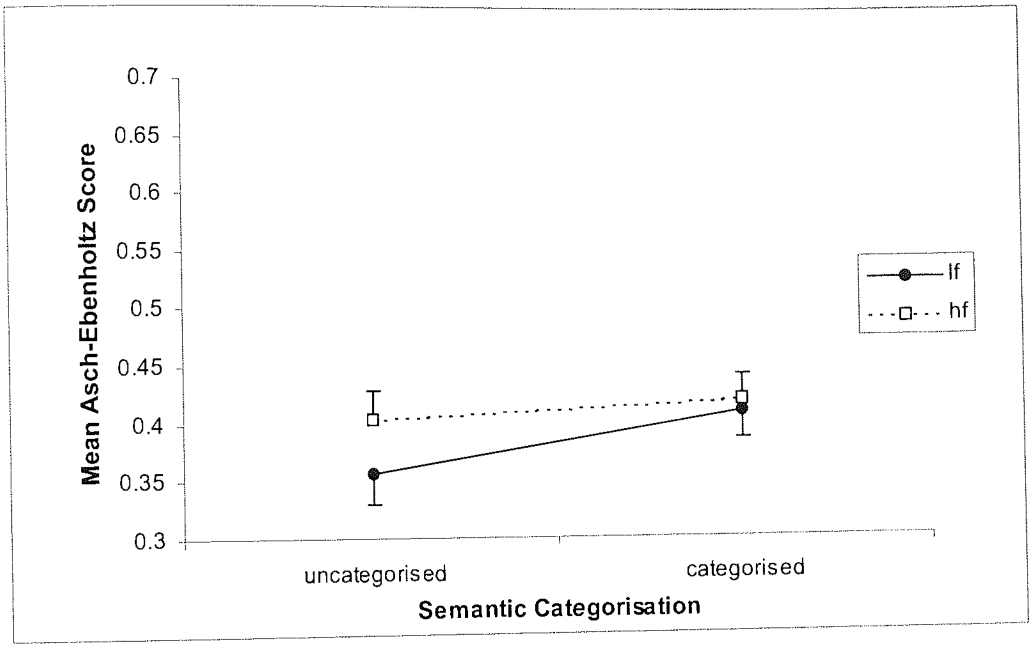
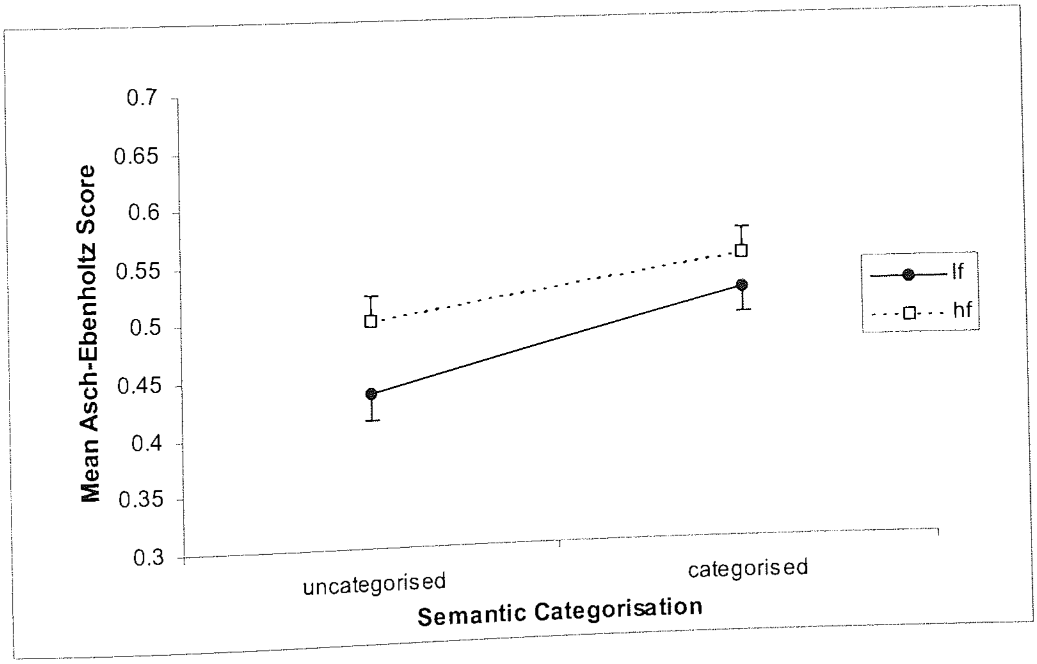


Figure 3.3:(b) Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency under control condition (Experiment 6).



Input-Output Correspondence

The grand mean for Asch-Ebenholtz scores was 0.45 (SD= 0.29). This figure is significantly different from the chance value of 0.50 ($t_{(1279)} = 5.92$, $p < 0.001$). Interestingly, as with the LFN lists from experiment one, this figure represents the case that overall, recall tended towards reversing the presented order of items (Asch & Ebenholtz, 1963). This phenomenon is discussed below. Figure 3.3 shows Asch-Ebenholtz scores broken down across the three variables of suppression, semantic categorisation and word frequency. A 3-way repeated measures ANOVA revealed no 3-way interaction between the variables ($F_{(1, 39)} = 0.01$).

Main effects data are shown in table 3.1. Articulatory suppression elicits less evidence of order encoding than the tapping control ($F_{(1, 39)} = 35.77$, $p < 0.001$). High frequency words have a higher Asch-Ebenholtz score than low frequency lists ($F_{(1, 39)} = 5.83$, $p < 0.05$) and semantically related lists were recalled in a manner more closely resembling the original order than unrelated lists ($F_{(1, 39)} = 11.80$; $p < 0.001$). None of the 2-way interactions were significant (all $Fs_{(1, 39)} < 1.60$).

Error Analysis

In total, 537 commission errors were recorded from a grand total of 10240 presented words (a 5.2% error rate). As in the previous chapter, errors were categorised into four distinct types. Semantic errors were those responses which were synonymous or highly conceptually related to a target word, or to another response made for the same list (e.g. *cow* and *cattle*). Phonological errors were those which rhymed with target words (e.g. *resort* and *report*). Previous list intrusions (PLIs) were target words or responses from previous list presentations. Several responses which overlapped these categories were classed as 'other' (e.g. *brotherhood* was a response to a list containing the target words *childhood* and *brother*). For each error type and each list type, no significant differences were found between articulatory and tapping conditions (all X^2s <1.5; df = 1). Table 3.2

shows a breakdown of these errors collapsed across articulatory conditions. Chi-square analysis showed that the difference in total errors for each list type was significant (X^2= 8.69; df = 3; p< 0.05). The totals in the table suggest that this effect is due to the greater number of errors in non-categorised lists (LFN and HFN) as compared to the two categorised lists.

Table 3.2: Frequency of response errors scored by type and experimental condition.

	SEMANTIC	PHON	PLI	OTHER	**TOTAL**
LFN	19	55	33	48	**155**
LFC	51	34	14	21	**120**
HFN	13	32	56	46	**147**
HFC	67	6	20	22	**115**
TOTAL	**150**	**127**	**123**	**137**	**537**

Individual analyses by error type yielded a significant difference between lists for semantic errors, with the distribution mainly falling in semantically categorised list types(X^2= 53.2, df = 3, p< 0.001). Phonological errors appeared more often in LFN lists, followed by equivalent error levels for LFC and HFN lists, with HFC lists having the least number of phonological errors (X^2= 38.07, df = 3, p<0.001). Previous list intrusions also demonstrated distributions contingent upon list type, with HFN lists eliciting the greatest number (X^2= 33.78, df = 3, p< 0.001).

Discussion

The pattern of results mirrors that found in experiments 1 and 5. Both the frequency of occurrence of words on a list and the semantic relatedness of items on the list have an effect on processing those lists for free recall. As before, these factors appeared to be additive in their effects, with HFN and LFC lists enjoying equivalent improvement over the LFN lists, and HFC lists outperforming all others. The effects of frequency and semantic categorisation are equally clear when the order of the items recalled is examined, re-emphasising the importance of order encoding in free recall and of order encoding upon the success of free recall.

The fact that overall, recall tended towards a reversal of the original presented order suggests that in this paradigm, participants were likely to output those items which were still consciously activated first (Crowder, 1989), and then to recall the remainder of the items. That this reduction in the Asch-Ebenholtz measure occurred for all lists under suppression and for only the LFN list under the control condition suggests that this strategy was employed by participants when no phonological trace was available as a recall cue. When the trace was available (as in the tapping condition), more use was apparently made of the phonological cues, and the rapid output of list-final items was apparently reduced. Future experiments would do well to take note of this issue and to analyse the recall positions of final items from the lists.

This experiment aimed to assess how much of the order encoding seen in output protocols was due to phonological loop involvement. In line with much of the literature, articulatory suppression whilst attempting to remember a list of words reduces performance in comparison to a non-phonological control task. The frequency effect for free recall is reduced under suppression, supporting the hypothesis that the classical word frequency effect is at least partially dependent on phonological processes, even for orthographic stimuli (cf. Ziegler, Tan, Perry & Montant, 2000).

The notion of the inter-relatedness between order encoding and phonological processing is further supported by the effect of suppression on input-out correspondence. Articulatory suppression reduces order processing by up to 15% in some lists. The picture which emerges from these data fits well with models that suggest that the order of items in a list is implicitly coded into the phonological traces which represent that list (cf. Murdock, 1997; Lewandowsky & Murdock, 1989; Wicklegren, 1977a), and that word frequency affects the process of accessing word representations in the lexicon for conversion into phonological traces. That there is no interaction between articulatory suppression and semantic categorisation is informative. If the effects of categorising the word list occurred during or before phonological loop processing, then one would expect semantic categorisation effects to occur differentially in lists which were subject to articulatory suppression and those which were not. Instead, the data fit the hypothesis that semantic effects happen at a post-phonological level and that semantic information is used to check the output of items to be recalled, either through redintegration or through a process of checking semantic constraints on to-be-output items.

This general picture is supported also by the error analysis. As in the previous chapter, semantic errors were made for semantically categorised word lists, suggesting that the spreading activation of words within the lexicon caused certain words to become highly active, and therefore more likely to be chosen by the participant at recall. For example, seven of the 40 participants recalled the word *distress* for the list containing *damsel*, even though *distress* was not a target item used in the experiment. Phonological errors were made mostly in low frequency, unrelated lists, suggesting that activation of the correct word in each case may not have been as robust as for the HFN lists. If only partial traces for words were available to participants through the phonological route, one would expect rhyming words to be accessed (and subsequently recalled) more often than entirely different words. Occurrences of words from previous lists in recall protocols also seem to fit the hypothesis, as words from previous presentations in the same session would have higher levels of residual activation than words not used at all in the experiment.

In sum, the present experiment supports the notion that order encoding is primarily a phonological phenomenon. But, the effects of frequency and semantic categorisation persist when the phonological loop is suppressed. This necessitates changes to current thinking. Redintegration supposedly acts only on phonological traces which can then be interpreted by top-down processes such as lexical and semantic information. If a unique match is found between a trace and a higher-level representation, then that word is output for recall. Both phonological and semantic information can be used to probe the lexicon and provide cues for activation. It may be that complete and degraded phonological traces inherently contain order information, and the more complete the traces, the more order is naturally retained.

However, it seems possible that this is not the only way in which order is encoded. In the same way that order encoding was demonstrated over a filled delay in experiment five the present experiment demonstrates that order encoding can occur without the involvement of the phonological loop. For example, in the case of semantic categorisation, natural associations exist between semantically associated words. Co-presentation of related items could conceivably temporarily boost their inter-item association strength (cf. Sheull & Keppel, 1967), and provide a chain of inter-related items, each a cue to the next. In this way, it could be supposed that it is the natural mechanisms of language processing which result in these short-term memory effects (Crowder, 1993). Further, given this perspective, it may be that this mechanism is the dominant one in encoding order, and that the phonological suppression effects on order encoding outlined above may occur simply as a result of reduced cognitive resources whilst carrying out the suppression task. This interpretation is reminiscent of the original order-encoding hypothesis (Delosh & McDaniel, 1996).

That even in the articulatory condition, some lists (e.g. HFC) had higher Asch-Ebenholtz scores than others (LFN) raises an important question. Are the lower scores representative of an initial output of recent items, as discussed above, followed by an essentially random process of recall, and could the slightly higher values from the HFC list represent an initial output of activated items, followed by an ordered recall of the remainder of the words? This could still create an Asch-Ebenholtz value of less than 0.5, whilst retaining some order information that the LFN lists did not. With the output lists often being so short for the suppressed LFN lists, it is effectively impossible to rigorously test this hypothesis, but this notion would form an interesting basis for investigation in the future.

If this latter notion is indeed what is occurring in the data presented here, then in the same way that order encoding was demonstrated over a filled delay in experiment five, the present experiment might point to the possibility that order encoding can occur without the involvement of the phonological loop.

In order to study the issue of phonological effects in more detail, experiment seven measured the speech rate of individual items on the word lists, and assessed its impact upon semantic and frequency effects in recall and order encoding in the absence of phonological suppression.

Experiment Seven

Experiment seven examined the role of speech rate in recall of categorised word lists. In measuring the speech rate of individual items, it was hoped that a more sophisticated analysis of the importance of phonological processing on the semantic and phonological routes mentioned above could be achieved. In addition, the possibility of confounding speech processes and attention was eliminated in this design. The various theories of short-term memory predict different outcomes for this design. The working memory model of Baddeley & Hitch (1974) predicts that semantic and frequency effects will come down to differences in speech rate between list types.

In the Baddeley & Hitch (1974) model, the phonological loop can hold about 2 seconds-worth of information. The longer it takes to pronounce a word, the less of those words a participant can remember. For example, Naveh-Benjamin & Ayres (1986) showed a clear relationship between the time it takes to pronounce the digits of a language and the digit memory span for native speakers of that language. Categorised lists and frequently-encountered words may be easier to access (cf. Kruschke, 1996; Humphreys, Bain & Pike, 1989), and Hitch (personal communication) suggests that this may be the basis of the semantic and frequency effects discussed thus far.

Consequently, when speech rate is entered as a covariate in the design, it should effectively account for most of the variance ascribed to semantic and frequential categorisation. The redintegration account predicts that speech rate *per se* should not directly account for any differences, but that semantic and frequential categorisation will affect the success of free recall. It will also predict frequency, but not semantic effects on order encoding (St-Aubin & Poirier, 1999b). The natural language account would suggest that semantic effects would occur in order encoding, but that speech rate would not directly affect the semantic and frequency effects in free recall.

Method

Participants were 24 University students from Aston University, aged between 18 and 25 years of age. None of the participants had taken part in previous experiments.

Design. The experiment followed a 2x2 within-subjects design. Independent variables were within-list word frequency (high or low) and grouping by semanticity (categorised or uncategorised). Dependent variables were proportion of correct responses for free recall scores, and input-output correspondence to assess the degree of order encoding in each recalled list. In addition, the speech rate for each word presented was obtained for each participant.

Apparatus and Materials were identical to those used in experiment six. Word lists used are available in appendix five.

Procedure. Participants were instructed that they would be visually presented with 32 lists of eight words which they were to verbally recall immediately after each list was presented. At no point in the instructions was there any indication that the participants were to recall the words in order.

For each participant, the order of the 32 word lists was randomised by computer, as was the order of words within each list. Commencement of each trial was initiated by the participant, and after a 2000 msec cue, words were presented one at a time in the middle of the screen for 1500 msecs (with a 500msec inter-stimulus interval). After the last word was presented, there was a 500 msec gap before recall cue was displayed (*****). Verbal responses were then recorded. Participants had a minimum recall time of 30 seconds, with no maximum limit imposed.

Once all lists had been presented, the speech rates[9] for each of the words used in the experiment were calculated. Each participant was presented with a list of all 256 words, and was asked to read each word out loud five times. As it was often the case that participants paused, stumbled or skipped words during this process, it was decided to measure the time taken to say each word individually by digitising the sound record and using computer software to measure the length of each sound wave. These values were then compiled into an average for each word for each person.

Results

Initially, results were analysed irrespective of speech rate, for comparison with earlier experiments. Subsequently, speech rates were added as covariates to assess the impact of articulatory processes on free recall and order encoding.

Table 3.3: Main effects of frequency and semantic categorisation on free recall and input-output correspondence.

	FREQUENCY		SEMANTIC CATEGORISATION	
	LF	HF	UNRELATED	RELATED
MEAN FREE RECALL (MAX 8)	4.03	4.44	3.83	4.65
SD	*1.26*	*1.22*	*1.15*	*1.22*
MEAN ASCH-EBENHOLTZ	.45	.54	.48	.51
SD	*.33*	*.29*	*.35*	*.28*

Free Recall

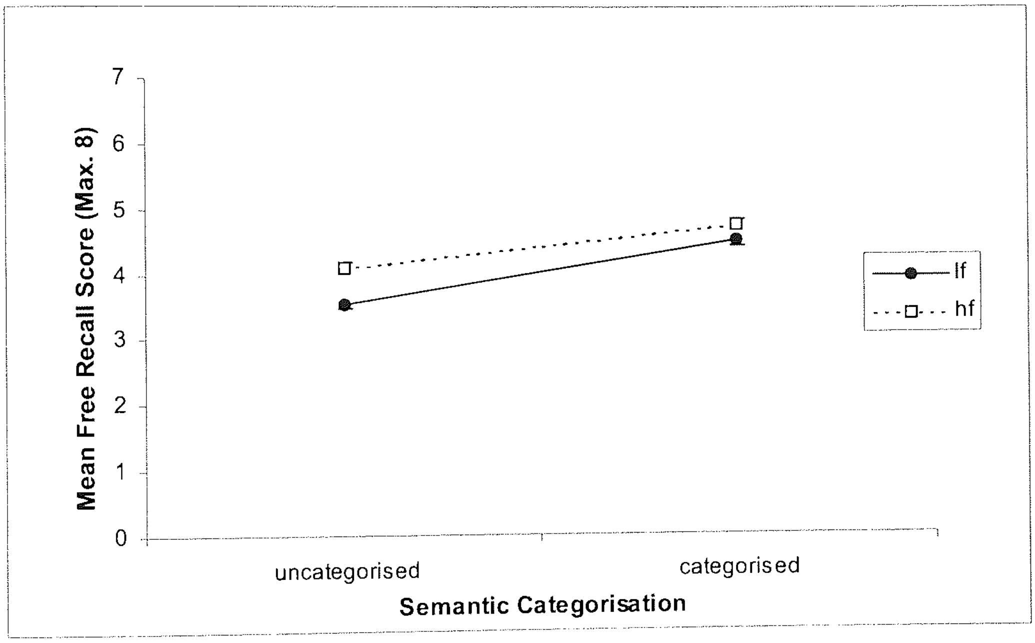
The grand mean for free recall was 4.24 words (SD=1.25). Figure 3.4 shows the mean number of words recalled for each list type. The general pattern is very similar to findings in the previous experiment and in the previous chapter. A repeated measures

---

[9] Speech rate, as used in experiment seven refers to the speed with which participants could pronounce the target words in the study, and not their rate of speech for all the words they knew. It is likely, however that these two different measurements are strongly correlated.

ANOVA was calculated and supported the weak interaction shown above ($F_{(1,23)}$= 4.84, $p < 0.05$). Main effects for word frequency ($F_{(1,23)}$= 31.19, $p < 0.001$) and semantic categorisation ($F_{(1,23)}$= 111.37, $p < 0.001$) were also found. Table 3.3 shows that high frequency words were recalled better than low, and categorised word lists were recalled better than uncategorised.

Figure 3.4: Mean free recall score per list by semantic categorisation and word frequency (Experiment 7).



Input-Output Correspondence

The grand mean Asch-Ebenholtz score was 0.495 (SD= 0.32). These scores are broken down by condition and list type in table 3.3 and figure 3.5 respectively. A main effect of word frequency was found, with low frequency lists having a lower Asch-Ebenholtz score than high frequency word lists ($F_{(1,23)}$= 11.64, $p < 0.001$). There was no main effect for semantic categorisation ($F_{(1,23)}$= 0.742), and the interaction between semantic relatedness and word frequency was marginal ($F_{(1,23)}$= 3.71, $p < 0.08$).

Figure 3.5: Mean Asch-Ebenholtz score per list by semantic categorisation and word frequency (Experiment 7).

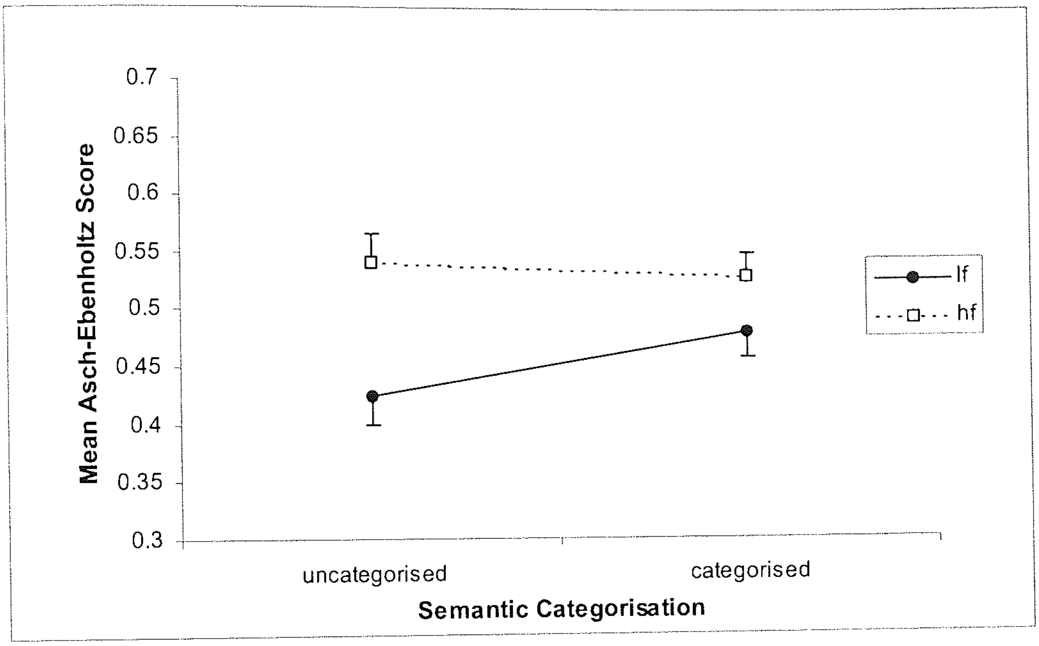
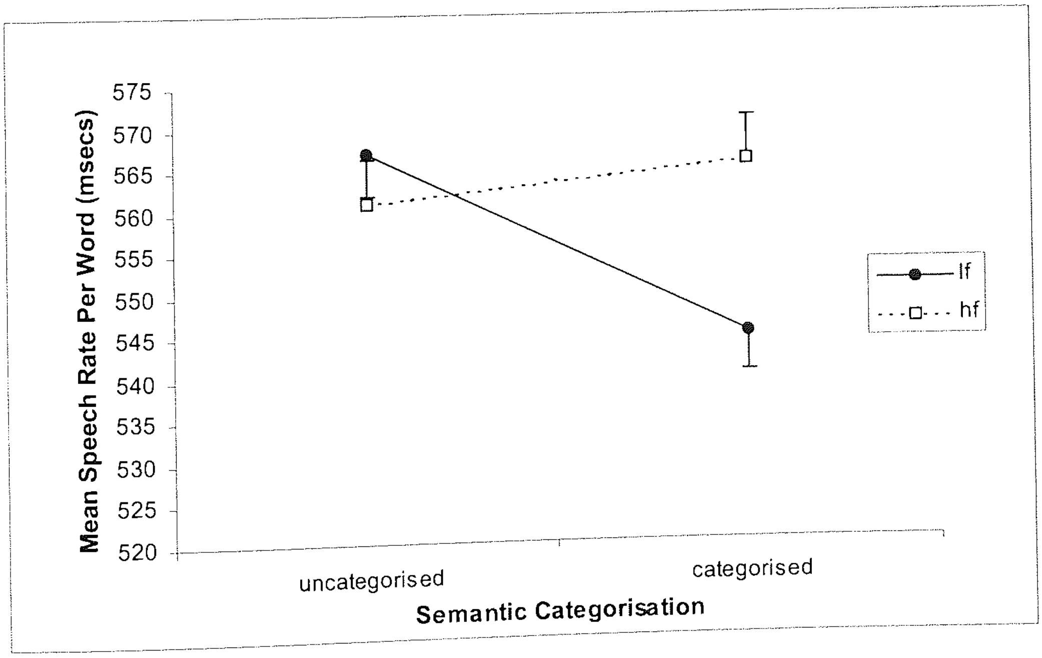


Figure 3.6: Mean Speech rate (in msecs) for stimulus lists by semantic categorisation and word frequency (Experiment 7).



Effects of Speech Rate

The grand mean over all 256 words across all 24 participants for pronunciation time of words used in the experiment was 560 msecs (SD= 71.2). Contrary to expectations, figure 3.6 shows no difference between pronounciation times for high and low frequency words ($F_{(1,23)}$= 1.54), and demonstrates that words in semantically categorised lists take longer to pronounce than uncategorised words ($F_{(1,23)}$= 7.73, $p< 0.05$). The interaction is also significant ($F_{(1,23)}$= 33.13, $p< 0.001$). However, close inspection of the mean speech rates for each list type reveals that all are within one 30th of a second of each other, and therefore caution must be applied whilst interpreting these tiny differences.

Nevertheless, the effect of speech rate on both free recall and input-output correspondence was assessed by entering speech rate as a covariate in a 2 x 2 independent measures ANOVA. Speech rate for a given list was calculated by averaging the individual pronunciation times for each of the eight words in that list.

A repeated measures analysis of covariance model was chosen over a regression model for this analysis due to the nature of the raw data involved. The two main factors, semantic and frequential categorisation, were divided over four lists and would have required a certain amount of unnecessary manipulation to be entered into a regression model to extract the factors from the repeated measures design. It was therefore considered that an ANCOVA would provide a cleaner and clearer interpretation of the data, and would also allow the direct comparison of effect sizes with previous experiments.

Results for free recall data suggest that when speech rate is removed from the model, the frequency effect becomes non-significant ($F_{(1,22)}$= 0.82) but the semantic categorisation effect remains ($F_{(1,22)}$= 5.84, $p< 0.05$), as shown in table 3.3. The interaction between semantic categorisation and word frequency is not significant ($F_{(1,22)}$= 1.41). Speech rate itself is not a significant predictor of free recall performance in this study ($F_{(1,22)}$= 0.79).

Results for input-output correspondence are also affected by the inclusion of speech rate as a predictor of order retention. The main effect of frequency discussed above is non-significant ($F_{(1,22)}$= 0.18). However, no effect of semantic categorisation is emergent ($F_{(1,22)}$= 0.43) and neither is the interaction significant ($F_{(1,763)}$= 0.09). Again, average list speech rate is not a significant predictor of input-output correspondence in its own right ($F_{(1,22)}$= 0.03).

Error Analysis

Commissive errors totalled 186 from a total of 6144 presented items (3%). Table 3.4 shows a breakdown of errors by list type and error type. The total number of errors did not differ across lists (X^2= 3.98, df = 3). However, semantic (X^2= 16.5, df = 3, $p<0.001$) and phonological errors (X^2= 16.5, df = 3, $p<0.001$) did differ between lists. The table suggests much the same pattern as previously outlined, with semantically categorised word lists being susceptible to semantic errors, and low frequency word lists having a tendency towards phonological errors. Differences between lists for the number of previous list intrusions were not significant (X^2= 3.42, df = 3).

Table 3.4: Frequency of response errors scored by type and experimental condition.

	SEMANTIC	PHON	PLI	OTHER	TOTAL
LFN	5	18	20	13	**56**
LFC	16	9	10	13	**48**
HFN	2	4	17	14	**37**
HFC	16	1	15	13	**45**
TOTAL	**39**	**32**	**62**	**53**	**186**

Discussion

The results presented above largely support the findings of the series of experiments reported in this thesis. Word frequency influences not only free recall performance, but the *order* of recall. Semantic categorisation of lists affects free recall performance, but does not affect order encoding in this experiment (despite it being a replication of experiment three in chapter one, which did find such an effect). This null semantic effect may have been due to the nature of the recall task. Experiment three used a written recall procedure, whereas the present experiment was based upon spoken recall by the participants. Little evidence exists for the disruption of order encoding by spoken processing (indeed, the opposite is often suggested), nevertheless, it is a possibility which warrants further investigation.

These data also show that the main effect of semantic categorisation could not be explained by reference to corresponding differences in the articulatory rate of the words in each list, there being very little relative difference between mean word pronunciation times for each list. Although analysis of speech rates for HF and LF words did not differ, speech rate effects still reduced the effect size of word frequency on order encoding to a non-significant level. Thus, there exists some degree of ambiguity in these results. Other researchers have found that articulatory rate *per se* is less useful a predictor of memory span than the syllable length of words (e.g. Bosshardt & Laug, 1995), and considering that all items used in the present design were matched for syllable length, this could explain the lack of any direct effect on memory performance.

As previously discussed, the fact that low frequency words had a tendency to be recalled in reverse order, and the high frequency words in forward serial order may well fit the order-encoding theory's hypothesis that the issue is a resource-based one, and that high frequency words are more quickly accessed in the lexicon at presentation, leaving more resources free to encode the contextual information. At the recall phase for low frequency words, it might be possible that a better strategy is to instantly output items that are still active within the lexicon (i.e. the most recent items presented), and then to attend to the redintegrating memory trace as a poor-quality cue. In the case of high frequency words, lexical representations of target words might be accessed that bit faster (see Hulme et al, 1997), and allow more target items to be successfully matched. This would mean that matching a redintegrated (phonological) trace is a better recall strategy for high frequency words than for low frequency. This approach offers some synthesis between the two major research directions outlined in this thesis, and the basis for this synthesis is discussed in chapter four.

Errors followed the same pattern as for many of the previous experiments, with the majority of phonological errors being made in lists containing low frequency words, perhaps suggesting that word frequency, phonology and lexical access are linked in short-term recall. This is considered more fully below. Semantic errors also adhered to previous distributions, as more semantic errors were made in lists which were semantically categorised, suggesting that word selection is based upon super-threshold activity of individual items within the lexicon.

Summary

The experiments reported in this chapter attempted to examine the relationship between the phonological processing of individual words and the order-encoding effects caused by word frequency and semantic factors found in chapter two. From the data reported here, it can be observed that the word frequency order effect is contingent upon the phonological encoding process (experiment six), and that this dependency may extend to the length of time taken to rehearse each word (experiment seven). When this data is combined with findings from other sources, a slightly clearer picture of the order encoding process begins to emerge.

Firstly, evidence presented here suggests that the locus of the frequency effect at the level of language input, rather than phonological output. Speech rate is a measure of the rate of language output, and in the current study, speech rate did not directly influence the frequency effect in terms of free recall, but did confuse the effect of frequency on order encoding. Whilst this may have been due to the similarity in the syllable length of the words used for this study, Wright (1979) and Walker (1999) have each obtained similar evidence in different designs in which the syllable lengths of target words did differ. However, interrupting normal phonological encoding by an articulatory suppression task interferes with the frequency effect at free recall and prevents the high-frequency advantage found when a non-verbal task of similar attentional complexity is concurrently performed. This supports models in which the initial processing mechanisms of verbal short term memory overlap, and are influenced by, language-processing mechanisms before rehearsal takes place (Crowder, 1993; Cowan, 1997).

The effects of phonological influences upon the frequency effect in order encoding are less clear-cut. Articulatory suppression does not inhibit the order encoding effect, whereas speech rate does have an influence upon it. Possible reasons for this are hypothetical, but if the redintegration model is correct, one would assume that at least part of the frequency effect on output order is based upon the completion of degraded traces processed at encoding. It is possible that because of their higher speed of lexical access at encoding (cf. Plaut *et al,* 1996), high frequency words have stronger traces, and are therefore less degraded than low frequency words. This would leave them more easily reconstructed at recall, and stronger individual traces are more likely to provide better cues to subsequent items in the list (cf Murdock, 1997; Lewandowsky & Li, 1994).

That semantic effects on free recall are not mediated by articulatory suppression reaffirms beliefs in a language processing system whereby semantic information can be processed with a mechanism distinct from phonological lexical activation. Results reported in this chapter correspond with a diffuse semantic network in which nodes sharing similar semantic characteristics are strongly connected and those which are unrelated are weakly connected (Plaut, *et al*, 1996). When a word is presented in a semantically categorised list, spreading activation from the target word node primes semantically related nodes around it, and thus speeds lexical access for subsequent words in the list. This would result in a processing speed advantage. However, the benefit of this speedier access is unaffected by phonological processing mechanisms, and must therefore depend on an alternative mechanism.

There are numerous possibilities for this alternative recall process. One is that after the input of a probe (such as the first word in the list), semantic links within the lexicon provide a chain of strong links following from one word to the next, strengthened by their juxtaposition at stimulus presentation. This process would neatly account for ordering effects found in semantically categorised word lists. The other is based on trace redintegration, and the possibility that as words can be accessed faster during lexical processing (and by extension, memory encoding), stronger traces may develop and stronger links between items are also able to develop. Again, this explanation can comfortably account for the increased order information evident in the output of lists. Whichever of these accounts may be true, one can be reasonably sure that these semantic effects do not occur in phonological rehearsal (Baddeley, 1997).

The patterns of errors in these two experiments also sit well with a retrieval-based account such as the redintegration model. If one were to assume that a memory trace for a word consists not only of phonological information, but also semantic, orthographic and syntactic information (and many other factors associated with language comprehension), then one would expect retrieval of a decaying trace to yield occasional errors which were closely related in meaning to target words. When targets which share semantic or phonological features appear in the same list, retrieving those traces becomes more difficult due to interference from the other items. This position is most coherent if one assumes Crowder's (1993) stance on the mechanisms of recall: retrieval is based upon the processes involved with language comprehension itself. Words on semantic lists may contain a number of features in common, and the resulting trace for those features in particular will be stronger, perhaps causing some of the finer resolution to be lost, and hence hindering distinctions between similar words in memory.

Chapter four examines the findings of this and the previous chapter in more detail, and discusses the implications of data presented here for the diverse range models of short-term memory.

Chapter Four: General Discussion and Conclusions

The research described here aimed to obtain a clearer picture of the processes which occur in short-term memory. The experiments reported were designed to probe the issues of order encoding as a cue to recall and the effect of information from long-term memory on the process of order encoding. In addition, two of the experiments attempted to assess the importance of phonological mechanisms in the order-encoding process. The series also addressed the issues of a distinction between short and long-term memory processes. Three of the experiments examined delayed recall, whereas four required immediate recall. By repeatedly manipulating the variables of word frequency and semantic categorisation in this manner, it was hoped that a more coherent representation of short-term memory processes might be obtained.

Summary of Findings

A concise summary of the findings from all experiments can be seen in table 4.1. Experiments one and two were replications of studies by order encoding theorists (Delosh & McDaniel, 1996). They had found that word frequency information affected the efficiency with which participants were able to use order information as an aid to free recall. In the experiments which have been reported here, there was no such effect on order information. This may have been because there was no evidence of order encoding at all in either lists of pure word frequency or of mixed high and low frequencies. Neither was any evidence found for the influence of semantic categorisation of word lists on order encoding, for the same reasons- neither semantically categorised nor uncategorised lists showed any difference from chance levels of order retention.

Frequency and semantic relatedness did however show an effect in free recall measures, with a semantic relatedness advantage for lists of pure frequency and mixed frequency. Support for part of the order encoding hypothesis was found in the mixed list low frequency advantage, which could not be predicted by the trace redintegration theory, but can be explained through the extra elaboration of item representations for rarer words, as compared to the more easily-accessed high-frequency words. However, the hypothesised freedom of resources to process order information when reading and remembering high frequency words did not reveal itself, undermining the most important aspect of the order-encoding theory.

Table 4.1: A summary of the experimental findings of the thesis.

EXPERIMENT	LIST TYPE (BY WORD FREQUENCY)	DELAY	SEMANTIC ITEM ADVANTAGE	SEMANTIC ORDER ADVANTAGE	FREQUENCY ITEM ADVANTAGE	FREQUENCY ORDER ADVANTAGE	PHONOLOGICAL EFFECT
1	PURE	40 SECS	YES	*NONE*	HF[1]	*NONE*	-
2	MIXED	40 SECS	YES	*NONE*	LF	-	-
3	PURE	IMMEDIATE	YES	YES	HF	HF	-
4	MIXED	IMMEDIATE	YES	*NONE*	LF	-	-
5	PURE	40 SECS	YES	YES	HF	HF	-
6	PURE	IMMEDIATE	YES	YES	HF	HF	ITEM & ORDER ENCODING REDUCED BY SUPPRESSION
7	PURE	IMMEDIATE	YES	*NONE*	HF	HF	NO DIRECT SPEECH RATE EFFECT

[1] HF- HIGH FREQUENCY, LF- LOW FREQUENCY

Experiments three and four aimed to replicate experiments one and two in an effort to discover the conditions in which order encoding does occur. To this end, an immediate recall design was chosen, on the basis of evidence that phonological coding and order information are closely related to immediate recall (Baddeley, 1986; Wicklegren, 1977a). In addition to the expected item advantages for high frequency and semantically categorised sets in pure lists, and low frequency advantages for mixed lists, order encoding was demonstrated to have taken place in one of the two experiments. This time, support for the order-encoding interpretation was complete, with high-frequency lists demonstrating improved levels of order retention, compared to low frequency lists. Semantic categorisation showed a similar advantage. This finding does not sit well with the redintegration hypothesis, which assumes order encoding to be a product of the way traces are stored in the phonological loop. If items in the loop degrade to such an extent that they lose their unique identifying features, then the trace redintegration process will select any of the items which offer a partial match to the degraded trace. Thus, for this theory, item recall should be enhanced, but order information should not be influenced by word frequency or by semantic categorisation.

A modified order-encoding hypothesis fits the data from experiment three much more closely. If one assumes that both semantic relatedness and high-frequency status allow words on a list to be accessed faster at presentation, then cognitive resources may then be free to process the temporal relationship between individual items. However, it is unclear from the order encoding hypothesis why this type of order encoding disappears over a filled delay of a few seconds, as demonstrated in experiments one and two.

For experiments one to four as a set, neither the trace redintegration theory nor the order encoding hypothesis can offer a full account of all of the data obtained. One suggestion is that participants in any given experiment utilise the most efficient strategy available to them in order to reduce cognitive load. Experiment five found that if one placed a requirement on the recallers to also remember order information over longer periods, then they are able to. It would appear that strategy selection plays an important part in memory tasks, and this is rarely specified by theoretical models, which for the most part concentrate almost exclusively upon a single hypothesised process or a single component of mental architecture.

Experiments six and seven aimed to examine the phonological component of order encoding. Experiment six demonstrated that suppression of articulatory activity reduced both item and order encoding. In addition, it was discovered that the frequency effect for item recall was entirely removed under articulatory suppression, but was not for the order of the items. Again, there does not appear to be an existing model which would predict this result. Experiment seven demonstrated that the speech rate for items on a list was not responsible for the item advantage for semantically categorised and high frequency lists. Nor was it related to the effect of word frequency on order encoding. Unfortunately, the semantic effect on order encoding was not replicated in this experiment, so it was impossible to assess whether the semantic-order effect reported in previous experiments was due to speech rate or to other factors, such as an enhancement of activation based on inter-item association strength.

The overall picture of research presented here is that order encoding appears to play a greater role in short-term memory than in delayed recall tasks, and that the process is related to the frequency of the words presented. Semantic and frequency categorisation both affect item recall, and each shows some evidence of influencing the order of items recalled, but the effect for frequential categorisation appears to be more robust than that for semantic categorisation. Finally, order encoding appears to be related to one's ability to process items with a phonological code.

Implications of the Data

Trace Redintegration

The data reported here have implications for the validity of all models discussed so far. The trace redintegration hypothesis is partially supported by the findings, in that order information is influenced by phonological factors, as demonstrated by the fact that articulatory suppression reduces the amount of order information encoded compared to a control condition. However, experiment seven demonstrates that this effect is not necessarily contingent upon the speech rate of items presented. Trace redintegration uses Baddeley's (1986) model as a basis, suggesting that short-term order encoding is a phonological loop phenomenon, and that the locus of the effects of long-term influence on short-term recall is in the post-retention phase. If the phonological loop itself were solely responsible for short-term order retention, then one would expect speech rate to have an effect on order encoding, and this was not found. However, some studies have suggested that syllable length of items is a better indicator of phonological loop retention than speech rate (e.g. Bosshardt & Laug, 1995). Future studies in this area might benefit from manipulation of the syllable length of stimuli.

On the other hand, St-Aubin & Poirier (2000) suggest that it is not the quantity of information in the phonological loop which affects order retention, but the quality of information. They see order information as specified solely by the uniqueness of items in the phonological loop. At recall, the degraded traces are automatically output in the order of presentation. The problem comes when those degraded traces are compared to existing lexical representations. If the traces have degraded to such an extent that a unique match to one of the list items is not possible, then other list items may be recalled in their place. This approach may explain why articulatory suppression reduces evidence of order retention in output protocols through the degrading of the phonological trace. However, an approach such as this does not allow for findings which suggest that order information itself is subject to perturbations in some very specific patterns. Estes (1972; Lee & Estes, 1981) and Wicklegren (1977a) each present data which suggest that the error profile of the recall of the order of items is more likely to involve the simple swapping of the position of adjacent items, and less and less likely to occur the further apart in the presentation list the two words occur. Such a phenomenon is easily explained by a theory in which order encoding is explained by associative recall (e.g. Wicklegren, 1977a), but very difficult to explain if order errors are solely the product of unreliable interpretation of correctly ordered, but homogeneous traces.

A small modification can be made to the redintegration theory in order for it to successfully account for the effect of semantic categorisation on item recall (see for example, Walker & Hulme, 1999; St-Aubin & Poirier 1999b). The reason for a word frequency advantage is cited as there being a small advantage in lexical access for high frequency words, which leads to more efficient access of the phonological representation of that word in long-term memory. This latter representation is compared to the phonological trace, and if a unique match is found, that item is recalled. Walker (1999; Walker & Hulme, 1999) suggests that the existence of semantic traces and semantic pattern-matching may be introduced to account for the effect of imagability on short-term recall. Temporary semantic traces may be set up which can be compared to the long-term semantic store, with similar constraints to phonological traces with respect to speed of decay. This notion could be an explanation as to why some of the experiments here have demonstrated an effect of semantic categorisation on order encoding: the lists used in this body of work were very loosely categorised as a result of balancing concreteness, imagability, age of acquisition and frequency factors. Therefore, the words used were initially more distinct from each other than most semantically categorised word lists used in the literature (cf. Poirier & St-Aubin, 1995; St-Aubin & Poirier, 1999b; Baddeley, 1966a). If one assumes that degrading the semantic trace acts in the same way as Walker suggests, then it could be that the degrading of standard categorised lists removes most of the unique semantic features of words, making the process of distinguishing between them correctly an impossible task. This would in turn make the retention of order information unlikely. However, in the lists used here, it could be that because of the initial distinctiveness between items, the degrading of the trace may not remove all unique information for items, and so boosts the amount of order information retained.

This account of the semantic categorisation effect seems unlikely for two reasons. Firstly, it is unparsimonious, an accusation which may be levelled at the whole retrieval-based hypothesis. It suggests that two traces are needed for each comparison- for example, the phonological trace and the long-term phonological representation. If this is extended to a semantic trace and a semantic LTM representation, and perhaps an orthographic pair too, the picture of human verbal memory becomes rapidly cluttered with simultaneous trace comparison processes, with no specification in the retrieval hypothesis as to how responses from these various systems are selected for output. The notion of parsimony is revisited below.

A second criticism of the semantic trace comparison theory is that semantic memories have been traditionally associated with long-term recall. There is no evidence available to suggest that semantic information decays rapidly. An alternative explanation of the semantic-item effect is that it may have its basis in lateral priming within the lexicon, with related words mutually increasing each other's activation levels, and therefore increasing the chance of lexical activation for semantically related stimulus items. The locus of the semantic categorisation effect having a different basis from that of the frequency effect seems intuitively correct, given the additive nature of the separate item and order effects shown throughout this experimental series.

The Order-Encoding Hypothesis

The data reported provide some evidence for the order encoding hypothesis, most apparently in the replication of the frequency effect on input-output correspondence in all but one of the experiments in which it was tested. Experiment one was the closest in design to the original study (Delosh & McDaniel, 1996), and yet provided evidence that participants do not preferentially use order encoding over longer delays. This would suggest that order encoding is not the most efficient cue to aid recall of individual items, as hypothesised by the order-encoding theory. However, it is apparent from the results of experiment five that participants are certainly able to access order information over longer delays, prompting speculation upon the richness of potential frameworks available to recallers with which to underpin the process of short-term recall.

This said, other researchers have found that order encoding remains constant over a variety of delays. Nairne, Whitman & Kelley (1999) reported that a filled delay did nothing to influence the recall of item information in an order reconstruction task. This only occurred for open list pools (i.e. stimulus sets in which no words were repeated across lists), and when closed pools, in which items used in the experiment were repeatedly drawn from the same small group, order encoding was more degraded across longer intervals. The explanation ascribed to this effect is that of proactive interference, in which it is supposed that information about lists themselves is remembered, and if that information is similar to representations of previous lists, the recaller will have a harder time distinguishing between the two, and therefore a harder time generating the appropriate context for recall. This idea may be the reason why order encoding falls off at longer intervals in the current study, but if it were, one would possibly expect the semantically categorised lists to demonstrate an advantage, since their representation would be more obviously distinct from other lists. This advantage occurs for item recall in all of the experiments reported here, but less frequently for order retention.

Both the order-encoding account and the redintegration account neatly explain the effects of semantic categorisation on item recall, and so the role of semantic information in order recall is of great interest to the order-encoding hypothesis. Whereas the trace redintegration account places the locus of order encoding at a level before lexical access, in the phonetic stream, the order encoding model predicts that order is processed only after lexical access has taken place. It is more likely, then that the order encoding account would predict *a priori* that semantic relationships between items would improve order encoding, as it is the associations between items which give rise to the encoding of order information. Semantically related items are likely to be more closely associated than unrelated items, and one might suspect that this associational information could be used to encode order.

Although the semantic-order effect reported in these studies is inconsistent, it has still been demonstrated often enough to warrant closer scrutiny. Of particular interest in the debate between the retrieval-based and the encoding-based explanation for the semantic effect is that of inter-item associative strength (see Postman & Keppel, 1970; Meyer & Schvaneveldt, 1971). A pre-lexical mechanism of order encoding should not be expected to be influenced by intra-lexical associations, and therefore inter-item association strength would not affect the role semantic information plays in order encoding. If, on the other hand, inter-item associations can be demonstrated to account for the semantic-order effect, the post-lexical account of order encoding would be supported. A study which tested the relationship between inter-item associations and the semantic categorisation effect on order encoding would therefore be invaluable.

It is not specified in the order encoding framework to what level phonological processing is involved in the recall of item and order information. One might expect the role of phonetics to be less important in the encoding model, as order information is primarily based upon associations between semantic lexical representations. However, it is perfectly possible that the representations being used for association are the phonological representations, which are stored in a separate phonological lexicon (Harley, 1997). The effects of both semantic categorisation and of articulatory suppression on item and order recall suggest that under the encoding approach, each of these mechanisms is active.

A major problem for the order encoding theory is that the degree of item recall demonstrated for filled delays and immediate recall is very similar, irrespective of how much order information has been retained. The order encoding hypothesis states that the recall of order information directly influences the success of item recall, and the findings here suggest that it does not. However, this was not directly assessed within one experimental design. It could be that as time passes, information available at short intervals is gradually re-processed in different elaborative codes, from the orthographic, through phonological to semantic representations, each of which has different characteristics, and therefore different modus operandi when it comes to recall strategies. This gradual progression of the form and organisation of the code used could occur as a progression of the natural processing of words in perceptual and linguistic systems, and perhaps offers a more parsimonious and more widely applicable explanation of verbal short-term memory.

The Naturalistic Approach

As each of the theories above focuses on a different part of the memory process, it is conceivable that a degree of synthesis may be achieved between the two differing points of view. The pertinent question is one of just how much of the long-term information processing is pre-lexical and how much occurs after the pronunciation of the word is accessed. The answer perhaps lies in models of the language-processing system.

There are a number of theorists (e.g. Crowder, 1993; Ellis & Young, 1988; Harley, 1997) who support the notion that verbal short term memory is very much a process extending from the natural processing of language. As such, there may be no memory store per se, but rather activity within the mechanisms designed for interpreting and generating language can be used to retain the kinds of information that are being processed. Thus, one may expect the influence of various factors upon the processing (and subsequent memory) of a word to occur at various stages of processing, from pre-lexical perception of word forms to lexical access and linguistic output. In this way, it is clear that both order encoding theory and the retrieval hypothesis have something to offer an all-encompassing theory of memory.

Support for the pre-lexical effects of semantic categorisation comes from priming studies which demonstrate that lexical access is faster for words which are associated (e.g. *doctor, nurse*) than for those which are not (Hodgson, 1991), although there is some evidence that lexical priming is more to do with association than with semantic relatedness (Shelton & Martin, 1992). Word frequency also demonstrates a strong lexical access effect (Whalley, 1978; Forster & Chambers, 1973). This is the assumption for which the order encoding hypothesis is based. If the speed of lexical access of related word lists and high frequency words is faster, then linguistic processing is able to move on to the next stage faster and process those words for meaning before low frequency and uncategorised word lists.

Equally, evidence exists for the effects of word frequency and semantic categorisation at a post-lexical level. Goldiamond & Hawkins (1958) found that if nonwords were frequently presented to participants, those participants were likely to spontaneously generate those nonwords as responses in a recognition task in which no recognisable word was presented. Frequent words demonstrate a response bias which has often been interpreted in the sense of activation thresholds in the lexicon. High frequency words require less evidence from a source than low to be selected for output. This is the basis for the redintegration account.

Recall of items on a list may be related to the activation levels of words in the lexicon. Specifically, the presentation of a group of words in a list will briefly augment the connections between items listed. Neurobiological evidence for this has been demonstrated by Hebb (1949). When it comes to recall, each individual item recalled may have a greater chance of cueing that item with which it was associated in the list, thus enhancing order recall (see Lee & Estes, 1981).

An account such as this can explain both the item and order advantages for semantically related and high frequency words; is psychologically plausible (Phillips & Singer, 1997); and can offer a more parsimonious account of trace redintegration. In their review, Phillips & Singer describe a computational algorithm which they believe can be applied to a great many human cognitive functions. They suggest that throughout perception, language and mental representation, the same basic function operates, namely the comparison of a representation of a letter, word or colour with its local and distal context. They show how, in a variety of neurophysiological and psychological paradigms, this effect of contextual input laterally and hierarchically influences the processing of the individual item which is then automatically bound by the algorithm into a Gestalt. Approaches such as this offer an exciting basis on which to build links which bridge classical distinctions between cognitive disciplines.

Where the trace redintegration theory requires a memory trace and a referential trace for comparison (Walker & Hulme, 1999), the interactive activation approach has all of the long-term information built in to the network. Trace decay at sub-lexical levels of word processing (phonological and orthographic traces, for example) involves certain of the words features (phonemes and graphemes) becoming less active due to interference or decay. Top-down processes are available from the lexicon through modality-specific feedback connections (Plaut *et al*, 1997), which are able to reactivate the correct sub-lexical units and consequently the correct word, through repeated iterations. The end result is that long-term information such as inter-item associativity and word frequency influence the probability of correct recall for item information.

Associative order encoding is also explained by this approach if one assumes that the connectivity between lexical units extends beyond the merely adjacent, and spreads over a wider range of list items, decreasing in strength (and therefore importance) the further one gets from the word itself. This approach has been characterised in Estes' (1972) model for order retention, and seems to follow fundamental principles of neurological organisation (Phillips & Singer, 1997).

The natural language approach offers a wider framework within which to study memory, and offers a level of contextual relevance often not found in models of specific aspects of recall. It is not as well specified in terms of its predictions of memory phenomena (Walker, 1999), yet it does offer a rich basis from which to build a more elaborated theory of recall. It is hoped that this approach will tie together some of the more fractionated theories of human memory.

Criticisms and Directions for Ongoing Research

As mentioned above, there are some aspects of the present series which would benefit from improvements in design. For one, the subtle difference between semantic categorisation and inter-item associative strength needs to be addressed. It may be that the format of the word lists used was confounded in terms of these two measures. Future experiments will need to assess just how much of the semantic advantage reported for order information here is due to the relatedness of the words as opposed to the categorisation per se. A model which hypothesises a lexical basis to the long-term order encoding effects discussed here would predict that the relatedness of the two words is more important than the categorisation itself, since it is the strength of the connection between the two words which is responsible for correct order encoding. The trace redintegration hypothesis does not specify just how the semantic categorisation item effect occurs, but would certainly predict that neither relatedness nor categorisation would play a role in the retention of order information.

This brings us to the second major criticism of the designs reported here. The issue of the point of contact between stimuli and lexical access is to some extent confounded in this series by the mode of presentation chosen for the stimuli. In many of the papers supporting trace redintegration theory, the presentation of words has occurred in the auditory modality. This means that the phonological code of the words is necessarily available before lexical access occurs. In the experiments reported in chapters two and three, a visual method of presentation was chosen, in which it is probable that the lexical activation occurs before the phonological code is accessed (Harley, 1997; see also Taft & von Graan, 1998). If this is the case, then obviously one should be extremely careful about drawing inferences on the locus of order encoding when comparing experiments based on different sensory modalities. In the experiments reported here, it is entirely plausible that the way in which information entered the phonological loop was influenced by lexical factors before the phonological representation of each word was active. In this way, the contents of the phonological loop could have been influenced by the LTM factors in a way which the verbally presented stimuli of trace redintegration studies could not (e.g. St-Aubin & Poirier, 1999b, Hulme *et al*, 1997; Walker & Hulme, 1999).

Any serious analysis of the order-encoding effects described here must therefore attempt to compare the presentation of visual and verbal stimuli. If it is lexical association which is the prime locus of the semantic-order effect, then one would expect the effect to be diminished under verbal presentation of stimuli. If the effect could be demonstrated under such conditions, however, it would offer a serious challenge to the existing trace redintegration account.

This said, evidence does exist for a non-lexical route for word pronunciation, in which word representations are not accessed at a lexical level, but are processed directly from written form into speech sounds (Gough, 1972; Coltheart & Rastle, 1994). If this is the case, then it may be responsible in part for the frequency effect reported throughout this thesis. Some of the words used in the current design were often unknown to the participants (e.g. *stoic*, *boatswain*: see Appendices), and were unlikely to have pre-existing lexical representations. The only possible route open to them for pronunciation was that of the grapheme-phoneme conversion route outlined above. This process is much slower and much less efficient than the direct lexical route (see Coltheart & Rastle, 1994), and may have been responsible for the resource drain for low frequency words hypothesised by Delosh & McDaniel (1996). However, although it was not tested directly in the current design, the words used were all matched for familiarity across all lists (Coltheart, 1981; Toglia & Battig, 1978), and should therefore have had some degree of lexical representation across participants. Due to the constraints imposed by attempting to neutralise the effects of concreteness, age of acquisition and imagability, matching of individual words was not possible. Future studies might benefit from the use of a low-frequency list in which the mean frequencies of words were not so vastly different from the mean frequencies of the HF lists.

As noted previously, the mode of response can affect the ability of participants to recall items successfully. In the first six experiments outlined in this thesis, participants were asked to recall items by writing them down in a list. This method was chosen in order that the design of the experiment replicated that of Delosh & McDaniel (1996) as closely as possible. In both of the immediate recall designs using pure lists, a semantic order effect was found. However, in experiment seven, recall was verbal rather than written, and although the semantic item effect remained, the semantic order effect obtained in some of the previous experiments was eliminated. This finding could be due to the fact that written answers enforce a slower response rate, and given that phonological traces rapidly decay, this may have interrupted the participant's primary method of storing order information. Hence, in some of the earlier experiments, the semantic order effect may have occurred as a result of participants using secondary strategies that were not used when a phonological cue to order was available. It would seem that this idea is worthy of study, and any further experiments would benefit from a study comparing the impact of spoken and written recall across immediate and filled delays on the semantic order effect.

The error analysis used in this thesis was only ever intended to be indicative, and if similar experiments were to be repeated at a future point, improvements in this form of data collection could be made. For example, it would be ideal to formulate a method of defining those errors currently marked as "other" in terms of which specific factors have influenced them. This includes errors such as the reporting of *shiver* instead of *quiver*: perhaps a post-trial test would be able to determine whether the participant was more sure of the sound of the word or the meaning. Also, it would be useful to obtain a measure of confidence for each word in the post-trial break in order to inform the analysis of errors further, for at present it is impossible to determine which errors (and indeed, correct answers) were the result of guesses.

Conclusions

In sum, this thesis has achieved its aims of gathering more data on the effects of long-term information on short-term recall. It has been demonstrated that semantic categorisation and word frequency affect both the recall of words, and more importantly, the order in which they are recalled. Two competing explanations for this effect have been examined, and each has been found wanting. Trace redintegration theory cannot explain the influence of semantic categorisation upon the order of items recalled, as it claims that the semantic effect occurs after the order of items has been recalled. Support for the straightforward order encoding hypothesis is undermined by the differences between immediate and delayed recall, which is suggestive of participants being able to selectively choose from a variety of strategies. The model proposed to account for the findings reported here is a more naturalistic approach which places the locus of all verbal short-term memory effects in the language comprehension and production system. The findings of this study support calls for a more Gestalt view of human memory processes, in which the complexity of the brain is given more importance in descriptions of processes and models of memory. Specifically, the research here advocates closer scrutiny of the relationship between lexical access and memory for words in future short-term recall studies.

References

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

Anderson, J.R., Bothell, D., Lebiere, C. & Matessa, M. (1998): An integrated theory of list memory. *Journal of Memory and Language,* **38**: 341-380.

Asch, S.E.& Ebenholtz, S.M. (1962a): The principle of associative symmetry. *American Philiosophical Society,* **106:** 135-163.

Asch, S.E.& Ebenholtz, S.M. (1962b): The process of free recall: Evidence for non-associative factors in acquisition and retention. *The Journal of Psychology,* **54:** 3-31.

Atkinson, R.C. & Shiffrin, R.M. (1968): Human memory: A proposed system and its control processes. In K.W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory, Volume 2.* New York: Academic Press.

Baddeley, A.D. (1966*a*): Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology,* **18:** 362-365.

Baddeley, A.D. (1966*b*): The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology,* **18:** 302-309.

Baddeley, A.D. (1986): *Working Memory.* Oxford: OUP.

Baddeley, A.D. (1997): *Human Memory: Theory and Practice (Revised Edition).* Hove: Psychology Press.

Baddeley, A.D. & Hitch, G. (1974): Working memory. In, G.A. Bower (Ed.), *Recent Advances in Learning and Motivation: Volume 8.* New York: Academic Press.

Baddeley, A.D. & Hitch, G. (1977): Recency re-examined. In, S. Dornic (Ed.), *Attention and Performance VI.* Hilldale, N.J.: LEA.

Baddeley, A.D. & Lewis, V.J. (1981): Inner active processes in reading: The inner voice, the inner ear and the inner. In, A.M. Lesgold & C.A. Perfetti (Eds.), *Interactive Processes in Reading.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Baddeley, A.D., Lewis, V.J. & Vallar, G. (1984): Exploring the articulatory loop. *Quarterly journal of experimental psychology,* **36**: 233-252.

Baddeley, A.D., Thomson, N. & Buchanan, M. (1975): Word length and structure of short-term memory. *Journal of Verbal Learning and Verbal Behaviour,* **14:** 575-589.

Baddeley, A.D. & Warrington, E.K. (1970): Amnesia and the distinction between short-term and long-term memory. *Journal of Verbal Learning and Verbal Behaviour,* **9:** 176-189.

Basso, A., Spinler, H., Vallar, G. & Zanobio, E. (1982): Left hemisphere damage and selective impairment of auditory verbal short-term memory: A case study. *Neuropsychologica,* **20:** 263-274.

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

Besner, D. & Davelaar (1982): Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology,* **36:** 701-711.

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

Bosshardt, H.G. & Laug, C. (1995): Relationships between word length, lexicality and articulatory rate. *Zeitschrift fur Experimentelle Psychologie*, **42:** 576-593. [Abstract]

Bourassa, D.C. & Besner, D. (1994): Beyond the articulatory loop- a semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin and Review,* **1**: 122-125.

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

Bousfield, W.A. & Cohen, B.H. (1955): The occurrence of clustering in the recall of randomly arranged words of different frequencies of usage. *Journal of General Psychology,* **52**: 83-95.

Bridges, A.M. & May, S.L. (1997): The role of order encoding in the free recall of words. *Proceedings of the British Psychological Society,* **6**: 37. [Abstract]

Broadbent, D.E. (1958): *Perception and communication.* London: Pergammon Press.

Brown, G.D.A. & Hulme, C. (1992): Cognitive psychology and second language processing: The role of short term memory. In, R.J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 105-122). North Holland: Elsevier Science Publishers.

Brown, G.D.A. & Hulme, C. (1995): Modelling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language,* **34:** 594-621.

Brown, J. (1958): Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology,* **10:** 12-21.

Cherry, E.C. (1953): Some experiments on the recognition of speech with one and two ears. *Journal of Acoustic Society of America*, **25:** 975-979.

Coltheart, M. (1981): The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology,* **33A:** 497-505.

Coltheart, M. & Rastle, K. (1994): Serial processing in reading aloud: Evidence for dual route models of reading. *Journal of Experimental Psychology: Human Perception and Performance,* **20:** 1197-1211.

Conrad, R. (1964): Acoustic confusion in immediate memory. *British Journal of Psychology,* **55:** 75-84.

Conrad, R. & Hull, A.J. (1964): Information, acoustic confusion and memory span. *British Journal of Psychology,* **55:** 429-432.

Cowan, N. (1988): Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin,* **104:** 163-191.

Cowan, N. (1997): *Attention and memory: An integrated framework.* New York: Oxford University Press.

Cowan, N., Day, L., Saults, J.S., Keller, T.A., Johnson, T., Flores, L. (1992): The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language*, **31:** 1-17.

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.

Craik, F.I.M. & Watkins, M.J. (1973): The role of rehearsal in long-term memory. *Journal of Verbal Learning and Verbal Behavior,* **12:** 599-607.

Crowder, R.G. (1969):Behavioural strategies in immediate memory. *Journal of Verbal Learning and Verbal Behaviour,* **8:** 153-162.

Crowder, R.G. (1979): Similarity and order in memory. In G.Bower (Ed.), *The Psychology of Learning and Motivation, volume 13.* London: Academic Press.

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

Crowder, R.G. (1989): Modularity and dissociations in memory systems. In, H.L. Roediger & F.I.M. Craik (Eds.), *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving.* Hillsdale, N.J.: LEA.

Crowder, R.G. (1993): Short-term memory: where do we stand? *Memory and Cognition,* **21:** 142-145.

Crowder, R.G. & Morton, J. (1969): Precategorical acoustic storage (PAS). *Perception and Psychophysics,* **5:** 365-373.

Crowder, R.G. & Neath, I. (1991): The microscope metaphor in human memory. In W.E. Hockley & S. Lewandowsky (Eds.), *Relating Theory and Data: Essays in honor of Bennet B. Murdock.* Hillsdale, N.J.: LEA.

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 on free recall. *Psychological Reports,* **5:** 305-312.

Deese, J. (1960): Frequency of usage and number of words in free recall: the role of association. *Psychological Reports,* **7**: 337-344.

Delosh, E.L. & McDaniel, M.A. (1996): The role of serial order information in free recall: Application to the word frequency effect. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **22**: 1136-1146.

Dennis, S. & Kruschke, J.K. (1998): Shifitng attention in cued recall. *Australian Journal of Psychology,* **50:** 131-138.

Duncan, C.P. (1974): Retrieval of low frequency words from mixed lists. *Bulletin of the Psychonomic Society,* **4:** 137-138.

Ebbinghaus, H. (1885): *Über das Gedächtnis.* Leipzig: Dunker. (Translation by H. Ruger and C.E. Bussenius (1913), *Memory.* New York : Teachers College.)

Ebenholtz, S.M. (1963): Serial position learning and sequential associations. *Journal of Experimental Psychology,* **66**: 353-362.

Ellis, A.W. & Young, A.W. (1998): *Human cognitive neuropsychology.* Hove: LEA.

Engle, R.W., Nations, J.K. & Cantor, J. (1990): Is working memory capacity just another name for word knowledge? *Journal of Educational Psychology,* **82:** 799-804.

Estes, W.K. (1972): An associative basis for coding and organisation in memory. In A.W. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington, D.C.: Winston.

Estes, W.K. (1991): On types of item coding and sources of recall in short-term memory. In, W.E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays in honour of Bennet B. Murdock.* Hillsdale, N.J.: Erlbaum.

Eysenck, M.W. (1978): Verbal remembering. In, B.M. Foss (Ed.), *Psychology Survey, No. 1.* London: Allen & Unwin.

Fodor, J.A. (1983): *The modularity of mind.* Cambridge, MA: MIT.

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

Francis, W.N. & Kuçera, H. (1982): *Frequency analysis of English usage: Lexicon and grammar.* Boston, M.A.: Houghton Mifflin.

Galton, F. (1883): *Inquiries into human faculty and its development.* London: Dent.

Gillund, G. & Shiffrin, R.M. (1984): A retrieval model for both recognition and recall. *Psychological Review,* **91:** 1-67.

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

Goldiamond, I. & Hawkins, W.F. (1958): Vexiversuch: The logarithmic relationship between word frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology,* **56:** 457-463.

Goldman-Rakic, P.S. (1995): Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences,* **769:** 71-83.

Gough, P.B. (1972): One second of reading. In, J.F. Kavanaugh & I.G. Mattingly (Eds.), *Language by Ear and by Eye.* Cambridge, MA: MIT Press.

Greene, R.L., Thapar, A. & Westerman, D.L. (1998): Effects of generation on memory for order. *Journal of Memory and Language,* **38**: 225-264.

Gregg, V. (1970): *Word frequency effects in human memory.* Unpublished PhD thesis, Birkbeck College, University of London.

Gregg, V. (1976): Word frequency, recognition and recall. In J.Brown (Ed.), *Recall and Recognition.* London: Wiley.

Gregg, V., Montgomery, D.C. & Castano, D. (1980): Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behaviour,* **19:** 240-245.

Harley, T. (1997): *The Psychology of Language: From Data to Theory.* Hove: Erlbaum, Taylor & Francis.

Healy, A.F. (1974): Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behaviour,* **13:** 644-655.

Hebb, D.O. (1949): *Organisation of Behaviour.* New York: Wiley.

Hebb, D.O. (1961): Distinctive features of learning in the higher animal. In, J.F. Delafresnaye (Ed.), *Brain Mechanisms and Learning.* London: OUP.

Hinton, G.E. & Shallice, T. (1991): Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review,* **93:** 411-428.

Hirshman, E. & Bjork, R.A. (1988): The generation effect: support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **14:** 484-494.

Hodgson, J.M. (1991): Informational constraints on pre-lexical priming. *Language and Cognitive Processes* **6:** 169-205.

Hoosain, R. & Salili, F. (1988): Language differences, working memory and mathematical ability. In, M.M. Gruenberg, P.E. Morris & R.N. Sykes (Eds.), *Practical aspects of memory: Current research and issues. Volume 2: Clinical and educational implications.* Chichester: Wiley.

Horowitz, L.M. & Prytulak, L.S. (1969): Redintegrative memory. *Psychological Review,* **76:** 519-531.

Hulme, C., Maughan, S. & Brown, G.D.A. (1991): Memory for familiar and unfamiliar words: Evidence for a long-term contribution to short-term memory span. *Journal of Memory and Language,* **30**: 685-701.

Hulme, C., Roodenrys, S., Brown, G.D.A & Mercer, R. (1995): The role of long-term memory mechanisms in memory span. *British Journal of Psychology,* **86:** 527-536.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G.D.A., Martin, S. & Stuart, G. (1997): Word frequency effects on short term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **23**: 1217-1232.

Humphreys, M.S., Bain, J.D. & Pike, R. (1989): Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review*, **96:** 208-253.

Huttenlocher, J. & Newcombe, N. (1976): Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behaviour,* **15:** 387-399.

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 Behaviour,* **12:** 471-480.

James, W. (1890): *Principles of Psychology, Vol 1.* New York: Holt.

Jenkins, J.J. & Russell, W.A.(1952): Associative clustering during recall. *Journal of Abnormal and Social Psychology,* **47:** 818-821.

Jones, D.M. & Macken, W.J. (1993): Irrelevant tones produced an "irrelevant speech effect": Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **19:** 369-381.

Jones, D.M. & Macken, W.J. (1995): Phonological similarity in the irrelevant speech effect: Within or between stream similarity? *Journal of Experimental Psychology: Learning Memory and Cognition,* **21:** 103-115.

Kruschke, J.K. (1992): Alcove: An exemplar-based connectionist model of category learning. *Psychological Review,* **99:** 22-44.

Kruschke, J.K. (1996): Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **22:** 3-26.

LeCompte, D.C. & Watkins, M.J. (1993): Similarity as an organizing principle in short-term memory. *Memory,* **1**: 3-22.

Lee, C.L. & Estes, W.K. (1981): Item and order information in short term memory: Evidence for multi-level feature perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory,* **7**: 149-169.

Lewandowsky, S. & Li, S.C. (1994): Memory for serial order revisited. *Psychological Review,* **101**: 534-538.

Lewandowsky, S. & Murdock, B.B. (1989): Memory for serial order. *Psychological Review,* **96**: 25-37.

McClelland, J. (1998): Connectionist models and Bayesian inference. In, M. Oaksford & N.Chater (Eds.), *Rational Models of Cognition.* pp 21-53. Oxford: OUP.

McCormack, P.D. & Swanson, A.L. (1972): Recognition memory for common and rare words. *Journal of Experimental Psychology,* **95:** 72-77.

Macken, W.J. & Jones, D.M. (1995): Functional characteristics of the inner voice and the inner ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21:** 436-448.

May, R.B. & Tryk, H.E. (1970): Word sequence, word frequency and free recall. *Canadian Journal of Psychology,* **24:** 299-304.

Melton, A.W. (1963): Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behaviour,* **9:** 596-606.

Merry, R. (1980): Image bizareness in incidental learning. *Psychological Reports,* **46:** 427-430.

Metcalf, S. (1990): Composite holographic associative recall model (CHARM) and blended memories in eye-witness testimony. *Journal of Experimental Psychology: General,* **119:** 145-160.

Meyer, D.E. & Schvaneveldt, R.W. (1971): Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operation. *Journal of Experimental Psychology,* **90:** 227-235.

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.

Milner, B. (1966): Amnesia following operation on the temporal lobes. In, C.W.M. Whitty & O.L. Zangwill (Eds.) *Amnesia.* London: Butterworths.

Morris, C.D., Bransford, J.D. & Franks, J.J. (1977): Levels of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behaviour,* **16:** 519-533.

Morton, J. (1967): A singular lack of incidental learning. *Nature,* **215:** 203-204.

Murdock, B.B. (1962): The serial position effect of free recall. *Journal of Experimental Psychology,* **64**: 482-488.

Murdock, B.B. (1975): Item and order information in short-term serial memory. *Journal of Experimental Psychology: General,* **105**: 191-216.

Murdock, B.B. (1982): A theory for the storage and retrieval of item and associative information. *Psychological Review,* **89:** 609-626.

Murdock, B.B. (1997): Context and mediators in a theory in a theory of distributed associative memory (TODAM2). *Psychological Review,* **104**: 839-862.

Nairne, J.S. (1990*a*): Similarity and long-term memory for order. *Journal of Memory and Language,* **29**: 733-746.

Nairne, J.S. (1990*b*): A feature model of immediate memory. *Memory and Cognition,* **16:** 343-352.

Nairne, J.S. & Kelley, M.R. (1999): Reversing the phonological similarity effect. *Memory and Cognition,* **27:** 45-53.

Nairne, J.S., Neath, I. & Serra, M. (1997): Proactive interference plays a role in the word length effect. *Psychonomic Bulletin and Review,* **4:** 541-545.

Nairne, J.S. & Neumann, C. (1993): Enhancing effects of similarity on long term memory for order. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **19**: 329-337.

Nairne, J.S., Riegler, G.L. & Serra, M. (1991): Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **17**: 702-709.

Nairne, J.S., Whiteman, H.L. & Kelley, M.R. (1999): Short-term forgetting of order under conditions of reduced interference. *Quarterly Journal of Experimental Psychology*, **52A**: 241-251.

Naveh-Benjamin, M. & Ayres, T.J. (1986): Digit span, reading rate and linguistic relativity. *Quarterly Journal of Experimental Psychology,* **38:** 739-751.

Neath, I., Suprenant, A.M. & LeCompte, D.C. (1998): Irrelevant speech eliminates the word length effect. *Memory & Cognition,* **26:** 343-354.

Nickerson, R.S. & Adams, M.J. (1979): Long-term memory for a common object. *Cognitive Psychology*, **11:** 287-307.

Norman, D.A. & Shallice, T. (1986): Attention to action: Willed and automatic control of behaviour. In, R.J. Davidson, G.E. Schwarts & D. Shapiro (Eds.), *Consciousness and self-regulation. Advances in research and theory, Volume 4.* New York: Plenum Press.

Parkin, A.J. (1988): Review of "working memory". *Quarterly Journal of Experimental Psychology*, **40A**: 187-189.

Patterson, K. & Hodges, R. (1992): Deterioration of word meaning – Implications for reading. *Neuropsychologia,* **30:** 1025-1040.

Peterson, L.R. & Peterson, M.J. (1959): Short-term retention of individual verbal items. *Journal of Experimental Psychology,* **58:** 193-498.

Phillips, W.A. & Singer, W. (1997): In search of common foundations for cortical computation. *Behavioural and Brain Sciences,* **20**: 657-706.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S. & Patterson, K. (1996): Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review,* **103:** 56-115.

Plaut, D.C. & Shallice, T. (1993): Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, **10:** 377-500.

Poirier, M. & St-Aubin, J. (1995): Memory for related and unrelated words: Further evidence on the influence of semantic factors in immediate serial recall. *Quarterly Journal of Experimental Psychology,* **48A**: 384-404.

Poirier, M. & St-Aubin, J. (1996): Immediate serial recall, word frequency, item identity and serial position. *Canadian Journal of Experimental Psychology,* **50**: 408-412.

Postman, L. (1975): Verbal learning and memory. *Annual Review of Psychology*, **26:** 291-335.

Postman, L. & Keppel, G. (1970): *Norms of word association.* New York : Academic Press.

Postman, L. & Phillips, L.W. (1965): Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology,* **17**: 132-138.

Robbins, T., Anderson, E., Barker, D., Bradley, A., Fearneyhough, c., Henson, R., Hudson, S. & Baddeley, A. (1996): Working Memory in Chess. *Memory and Cognition* **24:** 83-93.

Roodenrys, S., Hulme, C., Alban, J., Ellis, A.W. & Brown, G.D.A. (1994): Effects of word frequency and age of acquisition on short-term memory span. *Memory and Cognition,* **22**: 695-701.

Rugg, M.D., Cox, C.J.C, Doyle, M.C. & Wells, T. (1995): Event-related potentials and the recollection of high and low frequency words. *Neuropsychologia,* **33:** 471-484.

Rumelhart, D.E. & McClelland, J.L. (1986): *Parallel distributed processing: Volume 2, Psychological and biological models.* Cambridge, U.S.A: MIT Press.

Rundus, D. (1977): Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behaviour,* **16:** 665-681.

Rundus, D. & Atkinson, R.C. (1970): Rehearsal processes in free recall, procedure for direct observation. *Journal of Verbal Learning and Verbal Behaviour,* **9:** 99-105.

Sachs, J.S. (1967): Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics,* **2:** 437-442.

St-Aubin, J. & Poirier, M. (1999*a*): The influence of long-term memory factors on immediate serial recall: An item and order analysis. *International Journal of Psychology,* **34:** 347-352.

St-Aubin, J. & Poirier, M. (1999*b*): Semantic similarity and immediate serial recall: is there a detrimental effect on order information? *Quarterly Journal of Experimental Psychology,* **52A:** 367-394.

Salamé, P. & Baddeley, A.D. (1982): Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behaviour*, **21:** 150-164.

Schweickert, R. (1993): A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory and Cognition,* **21**: 168-175.

Schweickert, R., Hayt, C., Hersberger, L. & Guentert, L. (1996): How many words can be held in working memory? A model and method. In S. Gathercole (Ed.) *Models of short-term memory* (pp. 267-293). Hove: Lawrence Erlbaum.

Serra, M. & Nairne, J.S. (1993): Design controversies and the generation effect. *Memory & Cognition,* **21**: 34-40.

Shalllice, T. & Warrington, E.K. (1970): Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology,* **22:** 261-273.

Shallice, T. (1988): *From neuropsychology to mental structure.* Cambridge: CUP.

Shallice, T. & Warrington, E.K. (1974): The dissociation between long-term retention of meaningful sounds and verbal material. *Neuropsychologia,* **12:** 553-555.

Shelton, J.R. & Martin, R.C. (1992): How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory and Cognition,* **18:** 1191-1209.

Shepard, R.N. (1967): Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behaviour,* **6:** 156-163.

Shepard, R.N. & Metzler, J. (1971): Mental rotation of three-dimensional objects. *Science,* **171:** 701-703.

Shuell, T. & Keppel, G. (1967): A further test of the chaining hypothesis of serial learning. *Journal of Verbal Learning and Verbal Behaviour,* **6:** 439-445.

Smith, K.H. & Mynatt, B.T. (1982): Construction and representation of orderings in memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation, volume 16.* London: Academic Press.

Sperling, G. (1960): The information available in brief visual presentations. *Psychological Monographs: General and Applied,* **74:** 1-29.

Squire, L.R. & Cohen, N.J. (1984): Human memory and amnesia. In, J. McGaugh, G. Lynch & N. Weinberger (Eds.), *Proceedings of the conference on neurobiology of learning and memory.* New York: Guildford Press.

Squire, L.R. & Zola, S.M. (1998): Episodic memory, semantic memory and amnesia. *Hippocampus,* **8:** 205-211.

Sumby, W.H. (1963): Word frequency and serial position effects. *Journal of Verbal Learning and Verbal Behavior,* **1**: 443-450.

Taft, M & von Graan, F. (1998): Lack of phonological mediation in a semantic categorisation task. *Journal of Memory and Language,* **38:** 203-224.

Tehan, G. & Humphries, M.S. (1988): Articulatory loop explanations of memory span and pronunciation rate correspondences- a cautionary note. *Bulletin of the Psychonomic Society,* **26:** 293-296.

Toglia, M.P. & Battig, W.R. (1978): *Handbook of Semantic Word Norms.* New York: Erlbaum.

Toglia, M.P. & Kimble, G.A. (1976): Recall and use of serial position information. *Journal of Experimental Psychology: Human Learning and Memory,* **2:** 431-445.

Tulving, E. (1966): Subjective organisation and effects of repetition and multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavoiur,* **5:** 193-197.

Tulving, E. (1985): Memory and Consciousness. *Canadian Psychology,* **26:** 1-12.

Tulving, E., Schacter, D.L. & Stark, H.A. (1982): Priming effects of word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **8:** 336-342.

Tzeng, O.J.L. & Cotton, B. (1980): A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory,* **6:** 705-716.

Underwood, B.J. & Postman, L. (1960): Extra-experimental sources of interference in forgetting. *Psychological Review,* **67**: 73-95.

Vallar, G. & Baddeley, A.D. (1984): Fractionation of working memory. Neuropsychological evidence for a phonological short term store. *Journal of Verbal Learning and Verbal Behaviour*, **23**: 151-161.

Walker, I. (1999): *The role of language processing mechanisms in verbal short term memory.* Unpublished thesis. York University, 1999.

Walker, I. & Hulme, C. (1999): Concrete words are easier to recall than abstract: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **25:** 1256-1271.

Warrington, E.K. & Shallice, T. (1972): Neuropsychological evidence of visual storage in short-term memory tasks. *Quarterly Journal of Experimental Psychology,* **24:** 30-40.

Watkins, M.J. (1977): The intricacy of memory span. *Memory and Cognition,* **5**: 529-534.

Watkins, M.J. & Watkins, O.C. (1977): Serial recall and the modality effect: Effects of word frequency. *Journal of Experimental Psychology: Human Learning and Memory,* **3:** 712-718.

Waugh, N.C. (1970): Retrieval time in short-term memory. *British Journal of Psychology,* **61:** 1-12.

Waugh, N.C. & Norman, D.A. (1965): Primary memory. *Psychological Review,* **72:** 80-104.

Whiteman, H.L., Nairne, J.S. & Serra, M. (1994): Recognition and recall-like processes in the long-term reconstruction of order. *Memory,* **2**: 275-294.

Wicklegren, W.A. (1965): Short-term memory for phonemically similar lists. *American Journal of Psychology*, **78**: 567-574.

Wicklegren, W.A. (1967): Rehearsal grouping and heirarchical organisation of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology,* **19:** 97-102.

Wicklegren, W.A. (1977*a*): *Learning and Memory.* Englewood Cliffs, N.J.: Prentice Hall.

Wicklegren, W.A. (1977*b*): Rehearsal grouping and hierarchical organization of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology,* **19:** 97-102.

Winer, B.J. (1962) : *Statistical principles in Experimental Design.* New York: McGraw Hill.

Winzenz, D. & Johnson, K. (1980): Effects of input order and category cues on serial recall. *Psychological Reports,* **47**: 1079-1085.

Wright, C.C. (1979): Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory and Cognition,* **7:** 411-419.

Zeigler, J.C., Tan, L.H., Perry, C. & Montant, M. (2000): Phonology matters: The phonological frequency effect in written Chinese. *Psychological Science*, **11:** 234-238.

Appendix One: Pilot Study for creation of Pure Word Lists Used in Chapter Two

Pilot Study 1:

Aim: To create four types of word list, for use in the experiments described in chapter one. The four different types were represented by lists containing purely low-frequency words in semantically uncategorised lists (LFN), purely low-frequency words in semantically categorised lists (LFC), purely high-frequency words in semantically uncategorised lists (HFN) and purely high-frequency words in semantically categorised lists (HFC).

Introduction: Each of the four word list types were to be created using two distinct procedures. Firstly, the word frequency aspects of list creation were to be resolved using the Oxford Psycholinguistic database (Coltheart, 1981). This allowed the experimenter to download all words with a low frequency of occurrence for which information on other, confounding variables was available (concreteness, imagability and age of acquisition). Similarly, all high-frequency words with similar available information was obtained. These two source lists then formed the pools from which the experimenter created semantically categorised and uncategorised lists.

Participants: Twelve participants were drawn from the Staff and postgraduate student corpus at Aston University. None of the participants were used in any other study or pilot study.

Design: There was only one variable of interest in this study, namely categorisation status of the word lists. This had two levels: categorised or uncategorised. The experiment followed a repeated measures design.

Materials: Firstly, two pools of 144 bisyllabic words were obtained from the Oxford Psycholinguistic Database (Coltheart, 1981). All words with a frequency of occurrence less than 8 per million were obtained, with the proviso that information about each word existed on the following criteria: concreteness, age of acquisition and imagability. This formed the low-frequency word pool. Similarly, words with a frequency of occurrence of 50 or more per million were gathered, also subject to the availability of information on the three factors of concreteness, age of acquisition and imagability. This formed the high-frequency word pool.

Procedure: For each pool of words, the experimenter created lists of nine words which were subjectively thought to be related to each other semantically (for example, *science* and *knowledge* were thought to be related, as were *nature* and *country*). No formal method was applied in determining semantic relatedness at this stage. A further group of lists was created for each pool by purposely choosing words which appeared subjectively to have no obvious semantic relationship with each other. Across all four list types, care was taken to ensure that the mean concreteness, age of acquisition and imagability scores for each list did not greatly differ from any other list. Each of these three factors are scored on the Oxford Psycholinguistic Database on a scale from 0 to 700. The means and standard deviations for mean scores across all 32 lists for these factors were: concreteness, mean rating = 479.3 (SD= 57.2); age of acquisition, mean rating = 401.8 (SD= 57.4) and imagability mean rating = 483.6 (SD= 35.5). Despite an attempt to ensure that all scores for all of the confounding factors were within 100 rating points of the mean values across all lists, six lists of the 32 did not conform to this standard. Three lists had concreteness scores outside this margin, and three had larger absolute age of acquisition scores. However, for each of these six lists, all values were within 120 rating points of the grand mean. All lists created can be found in tables A.1-4, below.

The next step of the process was to assess how effective the subjective semantic categorisation had been for each of the lists created. The participants were each given a booklet containing the 32 word lists in a random order, four lists to a page. Each booklet used a different randomised order for the word lists and for words within lists. Participants were asked to rate each list on the basis of semantic relatedness alone. The exact instructions they were given were as follows:

INSTRUCTIONS:

On the following pages, you will find 32 lists of words. Your task is to simply decide how categorised each word list is in terms of the meaning of the words. You are asked to give each list a score between 1 and 7. If you think that the words on the list are all highly related, then give that list a 7. If you think that the words are not at all related, then give that list a 1. Use numbers between 1 and 7 for lists that are not so obvious. You have as much time as you require to complete this booklet.

Results:

Mean ratings across list types were obtained. Categorised lists (mean: 5.50; SD= .55) were rated as more semantically related than non-categorised lists (mean: 1.66; SD= .49). This difference was evident in all 12 participants, and was significant in a repeated measures t-test design, using means across list types as data points (*paired* t = 22.402, df = 11, p<.001).

Conclusions:

This experiment created four types of word list, covering two variables, word frequency and semantic categorisation. Furthermore, these stimuli were controlled in terms of the effect of a number of confounding factors, namely, concreteness, age of acquisition and imagability. Hence, it was hoped that these lists would prove a valuable tool in determining the processes involved in memory for linguistic items. The lists themselves are presented in the following tables.

Table A.1.1: Low-Frequency, Noncategorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
CORNICE	0	490	302	653
QUIVER	0	485	505	460
PROGRAMME	0	411	458	332
SENDER	0	482	370	406
HERDSMAN	0	543	500	458
SIXPENCE	0	583	602	249
NEIGHBOUR	0	515	548	323
JINGLE	0	437	497	386
POLYP	0	435	235	678
BROWNIE	0	535	553	381
LOWLAND	0	497	453	447
FISSURE	0	477	381	631
FORELOCK	0	565	421	542
FLAVOUR	0	449	472	342
AIRSHIP	0	585	545	406
HANDMAID	0	505	397	508
DEFENCE	1	346	413	447
MILLSTONE	1	578	526	464
JEWEL	1	594	621	292
THINKER	6	403	405	408
SYNOD	1	399	229	683
BAGPIPE	1	601	594	377
OFFAL	1	583	402	633
DREAMER	2	442	507	350
FIDDLE	2	582	555	367
TREMOR	2	487	491	511
FLUTTER	2	386	459	463
PLAZA	2	479	351	642
CIDER	2	626	626	431
CASTOR	2	535	429	453
BENZENE	2	535	418	675
TURNER	6	462	345	517
NAPKIN	3	585	582	342
LIAR	3	409	425	308
SPASM	3	439	486	558
SHEPHERD	3	598	600	275

Table A.1.1 (continued): Low-Frequency, Non-categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
MONSOON	3	508	498	508
SLUMBER	3	386	500	419
FOREARM	3	608	529	436
FORFEIT	3	295	310	533
FOOTSTEP	3	470	526	317
TUMBLE	3	433	461	350
RULER	3	555	543	311
KETTLE	3	602	594	274
KEEPER	3	459	421	392
KNUCKLE	3	586	520	356
FIELDER	3	509	472	458
CORKSCREW	3	614	580	419
BRISTLE	3	558	562	383
CANAL	3	598	588	350
BOATSWAIN	3	543	343	578
FENCING	4	525	518	497
FAIRY	4	433	536	242
FLYER	4	467	478	406
INCLINE	4	376	429	500
COMRADE	4	497	515	492
LABOUR	4	406	424	506
HOBBY	4	449	494	361
PEDAL	4	602	556	306
GRAVY	4	606	594	269
RECEIPT	4	474	432	517
TWILIGHT	4	467	588	411
ABYSS	4	450	453	597
MERMAID	1	494	578	322
MISTRESS	5	530	535	517
PALETTE	5	565	437	542
BANKER	5	547	565	392
SHUTTER	5	562	533	389
RATTLE	5	549	554	261
RACKET	5	562	530	386
PRELUDE	5	364	340	572
ULCER	5	558	516	517

Table A.1.2: Low-Frequency, Categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
PURSER	0	479	362	592
DESPOT	2	364	297	650
FRIAR	1	543	497	483
TAILOR	2	535	499	378
BEGGAR	2	533	593	364
CYNIC	0	379	377	619
TRAVELLER	3	492	491	364
SELLER	6	444	427	381
NOMAD	0	512	516	511

DAMSEL	1	544	551	406
HENCHMAN	1	503	447	578
GENTRY	1	452	462	556
TUNIC	1	563	508	406
BARON	2	513	498	472
ARMOUR	2	591	536	400
MAIDEN	2	545	554	429
MONARCH	3	525	572	456
MANOR	5	523	508	419

MOHAIR	0	583	508	583
SAMPLER	0	426	378	539
KNITTING	1	583	578	286
TUNIC	1	563	508	406
ARMOUR	2	591	536	400
TAILOR	2	535	499	378
LAUNDRY	5	576	559	367
GARMENT	6	552	507	453
LINEN	6	581	551	386

OUTPOST	3	462	378	481
DINER	0	515	497	444
CAVERN	1	534	548	433
OUTHOUSE	1	573	494	461
KENNEL	3	611	580	322
CANTEEN	2	587	540	436
WAREHOUSE	4	578	502	492
ABODE	4	498	458	553
MANOR	5	523	508	419

Table A.1.2 (continued): Low-Frequency, Categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
PHIAL	0	545	394	653
ODOUR	0	472	556	461
MICROBE	0	482	394	622
FLAVOUR	0	449	472	342
VAPOUR	0	499	493	533
PIPING	5	538	491	436
ALGAE	7	545	424	631
NOZZLE	4	555	513	411
GRAPHITE	5	562	443	586

CLEAVER	0	526	478	561
CARNAGE	0	376	351	628
DEFENCE	1	346	413	447
ARMOUR	2	591	536	400
TYRANT	2	467	494	492
MADMAN	2	470	545	417
HAVOC	3	338	505	469
TORTURE	3	437	533	408
SHIVER	4	455	578	308

ASTER	0	447	224	631
BLUEBELL	0	621	605	303
ROSEBUD	0	593	586	369
THISTLE	0	611	624	333
LILY	1	609	541	317
CHERRY	6	611	582	317
VIOLET	7	541	560	344
PAMPAS	0	570	410	597
CEDAR	1	608	516	425

WOODLAND	2	585	608	381
FLORA	1	557	472	569
CREEPER	1	555	526	431
DUNGHILL	0	576	475	483
ADDER	0	615	583	356
SKYLARK	1	614	548	394
ROBIN	2	637	615	233
THICKET	1	571	511	469
THISTLE	0	611	624	333

Table A.1.3: High-Frequency, Non-categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
PICTURE	162	579	581	219
WELCOME	50	350	470	342
CONGRESS	152	384	356	575
BEDROOM	52	615	629	206
SILENCE	52	352	470	333
BREAKFAST	53	576	586	233
CONTENT	53	300	391	389
MINUTE	53	361	473	264
DEVICE	55	444	391	472
FIFTEEN	56	379	491	289
VISION	56	395	440	411
REVIEW	56	388	345	547
MOTOR	56	565	521	344
ENTRANCE	57	484	493	372
METAL	61	582	541	308
SHOULDER	61	589	577	264
BUILDING	160	589	578	300
MUSIC	216	512	549	272
APPEAL	62	333	402	431
FELLOW	63	502	435	378
SIGNAL	63	464	513	367
WOMAN	224	580	626	258
OBJECT	65	487	408	339
ESCAPE	65	341	459	353
RELIEF	66	303	432	443
WONDER	67	305	402	339
FASHION	69	356	474	467
WEATHER	69	439	537	292
EXCHANGE	70	356	394	443
BEAUTY	71	336	513	344
FACTOR	71	328	269	474
COLUMN	71	520	491	436
PLATFORM	72	547	529	386
PRESENCE	76	326	339	456
TITLE	77	384	413	375
COFFEE	78	613	618	292

Table A.1.3 (continued): High-Frequency, Non-categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
LADY	80	564	571	231
SURFACE	200	447	453	378
PATTERN	113	472	453	319
PATIENT	86	487	526	366
ATTEMPT	95	313	302	392
CONCERN	98	509	353	444
COUNCIL	103	435	405	464
SEASON	105	445	495	328
STATION	105	572	554	300
DISTANCE	108	353	432	344
STANDARD	110	324	319	472
BUILDING	160	589	578	300
WINTER	83	499	621	236
FUNCTION	113	343	294	511
JUSTICE	114	307	379	500
DESIGN	114	444	407	403
PEOPLE	847	540	548	281
WINDOW	119	609	602	231
APPROACH	123	323	329	444
RESPECT	125	280	343	433
HOTEL	126	591	597	308
TREATMENT	127	343	408	433
COUNTRY	324	465	539	329
WATER	442	616	632	153
TROUBLE	134	310	395	322
SUMMER	134	439	618	253
TRIAL	134	446	516	433
VOLUME	135	418	464	461
NORMAL	136	237	294	375
MEMBER	137	455	399	392
NATION	139	415	436	425
METHOD	142	303	304	481
HUMAN	299	583	543	369
MEETING	159	403	451	356
ISLAND	167	596	643	289
EVENING	133	439	559	303

Table A.1.4: High-Frequency, Categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
MEMBER	137	455	399	392
COUSIN	51	502	478	278
UNCLE	57	580	574	192
BROTHER	73	585	589	219
CIRCLE	60	515	591	214
MARRIAGE	95	398	556	383
BACKGROUND	67	383	427	400
WOMAN	224	580	626	258
FAMILY	331	525	577	280
CAREER	67	373	418	514
SUCCESS	93	295	443	411
DEGREE	125	406	521	508
THEORY	129	287	317	557
SCIENCE	131	366	423	458
STUDENT	131	549	603	481
KNOWLEDGE	145	278	348	477
ADVICE	51	291	352	425
SUBJECT	161	406	418	417
CHILDHOOD	50	335	489	372
PAINTING	59	615	602	258
WRITING	117	467	540	256
QUESTION	257	387	398	314
NUMBER	472	395	489	239
TEACHER	80	569	575	247
MEASURE	91	366	379	344
LEARNING	60	303	370	322
ANSWER	152	397	368	294
GARDEN	60	602	635	186
FOREST	66	609	633	297
VILLAGE	72	576	578	317
VALLEY	73	575	600	339
PRODUCE	82	432	396	431
MARKET	155	551	583	328
NATURE	191	414	513	342
COUNTRY	324	465	539	329
CATTLE	97	600	619	261

Table A.1.4(continued): High-Frequency, Categorised words used in Chapter Two.

WORD	FREQUENCY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
NUMBER	472	395	489	239
ORDER	376	344	352	344
MILLION	204	364	440	419
DOZEN	52	396	525	319
SERIES	130	373	398	461
MEASURE	91	366	379	344
PORTION	62	384	399	411
DISTANCE	108	353	432	344
UNIT	103	389	334	411
LEADER	74	487	502	353
HERO	52	428	483	361
CONFLICT	52	305	432	497
VICTORY	61	376	461	411
RIFLE	63	606	581	322
BODY	276	568	614	267
MURDER	75	445	549	381
ARMY	132	543	578	317
DANGER	70	338	505	300
PAYMENT	53	432	472	439
BUDGET	59	366	394	497
MONEY	265	574	604	247
AMOUNT	172	335	316	392
INTEREST	330	305	359	411
INCOME	109	429	475	506
STATEMENT	141	379	386	481
EXCHANGE	70	356	394	443
BALANCE	90	366	429	367
PAPER	157	599	590	229
LETTER	145	577	595	256
REPORT	174	417	411	386
NOVEL	59	529	547	475
PICTURE	162	579	581	219
NOTICE	59	479	467	369
PERMIT	77	399	388	458
ORDER	376	344	352	344
DESIGN	114	444	407	403

Appendix Two: Instructions given to participants in Chapter Two

EXPERIMENT ONE AND TWO INSTRUCTIONS:

In each trial, you will be shown **[nine/ eight]** words, one at a time. Your task is to remember the words you are shown. After each group of words has been presented, you will be shown some simple maths problems, and will be asked to solve them in your head. You will only get half a minute for these problems, and you are not expected to be able to answer them all. Just try to answer as many questions as possible in the time given.

Once you have performed this task, a row of asterisks will appear (*****), and this is your cue to begin writing down all the words you can remember on the sheets provided. Use a new sheet for every trial. You will be informed when you are allowed to move on to the next trial by a message on screen.

There will be **[eight/ twelve]** trials in total. If you have any questions, ask the experimenter now.

PRESS ANY KEY WHEN YOU ARE READY TO BEGIN.

EXPERIMENT THREE AND FOUR INSTRUCTIONS:

In each trial, you will be shown **[nine/ eight]** words, one at a time. Your task is to remember the words you are shown. Once the words have disappeared, a row of asterisks will appear (*****). When you see this sign, this is your cue to write down all the words you can remember on the sheets provided. Use a new sheet for every trial. You will be informed when you are allowed to move on to the next trial by a message on screen.

There will be **[eight/ twelve]** trials in total. If you have any questions, ask the experimenter now.

PRESS ANY KEY WHEN YOU ARE READY TO BEGIN.

EXPERIMENT FIVE INSTRUCTIONS

In each trial, you will be shown eight words, one at a time. Your task is to remember the words you are shown. After each group of words has been presented, you will be shown some simple maths problems, and will be asked to solve them in your head. You will only get half a minute for these problems, and you are not expected to be able to answer them all. Just try to answer as many questions as possible in the time given.

Once you have performed this task, a row of asterisks will appear (*****), and this is your cue to begin writing down all the words you can remember on the sheets provided. Use a new sheet for every trial. You will be informed when you are allowed to move on to the next task by a message on screen.

The next task requires you to place the words you have just seen *in the correct order*. All the words will be given to you on screen. Simply write them *in the correct order* on the sheets provided.

There will be eight trials in total. If you have any questions, ask the experimenter now.

PRESS ANY KEY WHEN YOU ARE READY TO BEGIN.

Appendix Three: Statistical Summaries for Chapter Two

Table A.3.1: Experiment One: 2x2 Independent Measures ANOVA for Free Recall means.

Tests of Between-Subjects Effects

Dependent Variable: FRMEAN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	29.138[a]	3	9.713	7.904	.000
Intercept	1775.885	1	1775.885	1445.149	.000
FREQ	5.130	1	5.130	4.175	.045
CAT	22.705	1	22.705	18.477	.000
FREQ * CAT	1.302	1	1.302	1.060	.307
Error	73.732	60	1.229		
Total	1878.754	64			
Corrected Total	102.869	63			

a. R Squared = .283 (Adjusted R Squared = .247)

Table A.3.2: Experiment One: 2x2 Independent Measures ANOVA for Asch-Ebenholtz means.

Tests of Between-Subjects Effects

Dependent Variable: AEMEAN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.653E-02[a]	3	8.844E-03	.555	.646
Intercept	17.079	1	17.079	1072.794	.000
FREQ	1.945E-03	1	1.945E-03	.122	.728
CAT	1.935E-03	1	1.935E-03	.122	.729
FREQ * CAT	2.265E-02	1	2.265E-02	1.423	.238
Error	.955	60	1.592E-02		
Total	18.061	64			
Corrected Total	.982	63			

a. R Squared = .027 (Adjusted R Squared = -.022)

Table A.3.3: Experiment One: Scheffe post-hoc comparisons for free recall means.

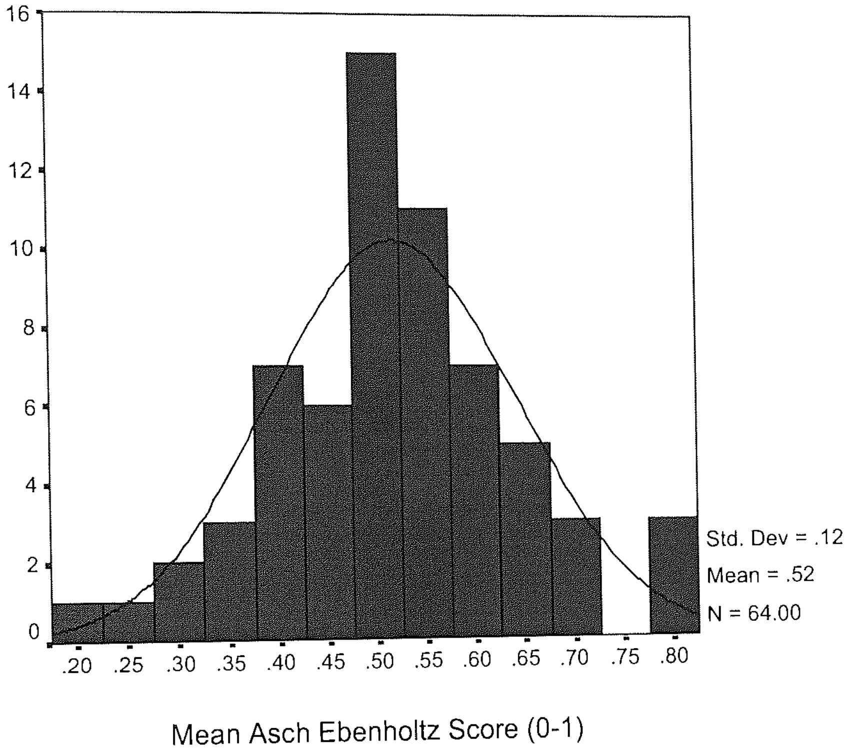
Multiple Comparisons

Dependent Variable: FRMEAN
Scheffe

(I) FRECAT	(J) FRECAT	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
lfn	lfc	-.91	.3919	.160	-2.03	.22
	hfn	-.28	.3919	.915	-1.41	.85
	hfc	-1.76*	.3919	.001	-2.88	-.63
lfc	lfn	.91	.3919	.160	-.22	2.03
	hfn	.63	.3919	.473	-.50	1.75
	hfc	-.85	.3919	.205	-1.98	.28
hfn	lfn	.28	.3919	.915	-.85	1.41
	lfc	-.63	.3919	.473	-1.75	.50
	hfc	-1.48*	.3919	.005	-2.60	-.35
hfc	lfn	1.76*	.3919	.001	.63	2.88
	lfc	.85	.3919	.205	-.28	1.98
	hfn	1.48*	.3919	.005	.35	2.60

*. The mean difference is significant at the .05 level.

*Fig A.3.1. **Demonstration of Normal Distribution of Asch-Ebenholtz scores for Expt 1***



EXPERIMENT TWO

Table A.3.4: Experiment Two: Unpaired t-test on Free Recall and Asch-Ebenholtz means.

Independent Samples Test

	t-test for Equality of Means						
						95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
FRMEAN	-3.649	34	.001	-1.0417	.28549	-1.62185	-.46148
AEMEAN	-.056	34	.956	-.0023	.04068	-.08496	.08039

Table A.3.5a: Experiment Two: 2x (2) Mixed ANOVA SummaryTable- Repeated Measures Effects on Free Recall means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	FREQ	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	Linear	1.972	1	1.972	17.578	.000
FREQ * LISTTYPE	Linear	2.170E-02	1	2.170E-02	.193	.663
Error(FREQ)	Linear	3.815	34	.112		

Table A.3.5b: Experiment Two: 2x (2) Mixed ANOVA SummaryTable- Independent Measures Effects on Free Recall means

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	403.359	1	403.359	1099.767	.000
LISTTYPE	4.883	1	4.883	13.313	.001
Error	12.470	34	.367		

EXPERIMENT THREE

Table A.3.6. Experiment Three: 2x2 Independent Measures ANOVA on Free Recall means

Tests of Between-Subjects Effects

Dependent Variable: FRMEAN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	36.305[a]	3	12.102	49.632	.000
Intercept	1563.955	1	1563.955	6414.086	.000
FREQ	6.172	1	6.172	25.313	.000
CAT	30.078	1	30.078	123.357	.000
FREQ * CAT	5.493E-02	1	5.493E-02	.225	.637
Error	14.630	60	.244		
Total	1614.891	64			
Corrected Total	50.935	63			

a. R Squared = .713 (Adjusted R Squared = .698)

Table A.3.7. Experiment Three: 2x2 Independent Measures ANOVA on Asch-Ebenholtz means

Tests of Between-Subjects Effects

Dependent Variable: AEMEAN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.769E-02[a]	3	3.256E-02	6.754	.001
Intercept	22.141	1	22.141	4592.264	.000
FREQ	7.307E-02	1	7.307E-02	15.155	.000
CAT	2.381E-02	1	2.381E-02	4.939	.030
FREQ * CAT	8.078E-04	1	8.078E-04	.168	.684
Error	.289	60	4.821E-03		
Total	22.528	64			
Corrected Total	.387	63			

a. R Squared = .252 (Adjusted R Squared = .215)

EXPERIMENT FOUR

Table A.3.8. Experiment Four: Unpaired t-test on Free Recall and Asch-Ebenholtz means

Independent Samples Test

	t-test for Equality of Means						
						95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
FRMEAN	5.449	46	.000	1.0278	.18863	.64808	1.40747
AEMEAN	.083	46	.934	.0016	.01902	-.03670	.03988

Table A.3.9a. Experiment Four: 2x (2) Mixed ANOVA SummaryTable- Repeated Measures Effects on Free Recall means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	FREQ	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	Linear	3.501	1	3.501	31.137	.000
FREQ * LISTTYPI	Linear	.000	1	.000	.000	1.000
Error(FREQ)	Linear	5.172	46	.112		

Table A.3.9b. Experiment Four: 2x (2) Mixed ANOVA SummaryTable- Independent Measures Effects on Free Recall means

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	691.584	1	691.584	3239.403	.000
LISTTYPE	6.338	1	6.338	29.687	.000
Error	9.821	46	.213		

EXPERIMENT FIVE

Table A.3.10. Experiment Five: 2x2 Repeated Measures ANOVA on Free Recall means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	FREQ	CAT	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	Linear		5.578	1	5.578	14.319	.001
Error(FREQ)	Linear		7.402	19	.390		
CAT		Linear	4.572	1	4.572	9.979	.005
Error(CAT)		Linear	8.705	19	.458		
FREQ * CAT	Linear	Linear	4.883E-03	1	4.883E-03	.025	.875
Error(FREQ*CA	Linear	Linear	3.647	19	.192		

Table A.3.11. Experiment Five: 2x2 Repeated Measures ANOVA on Asch-Ebenholtz means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	FREQ	CAT	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	Linear		5.578	1	5.578	14.319	.001
Error(FREQ)	Linear		7.402	19	.390		
CAT		Linear	4.572	1	4.572	9.979	.005
Error(CAT)		Linear	8.705	19	.458		
FREQ * CAT	Linear	Linear	4.883E-03	1	4.883E-03	.025	.875
Error(FREQ*CA	Linear	Linear	3.647	19	.192		

Table A.3.12. Experiment Five: 2x2 Repeated Measures ANOVA on Ordinal means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	FREQ	CAT	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	Linear		12.012	1	12.012	61.712	.000
Error(FREQ)	Linear		3.698	19	.195		
CAT		Linear	.378	1	.378	1.005	.329
Error(CAT)		Linear	7.145	19	.376		
FREQ * CAT	Linear	Linear	7.031E-03	1	7.031E-03	.039	.846
Error(FREQ*CA	Linear	Linear	3.454	19	.182		

Appendix Four: Pilot Study for creation of Mixed Word Lists Used in Chapter Two

Pilot Study 2:

Aim: To create two types of word list, for use in the experiments described in chapter one. The two different types were represented by lists containing a mixture of high and low frequency words. Mixed categorised (MC) lists were to contain four low frequency and four high frequency words, and all eight words in each list were to be semantically related to each other. The mixed non-categorised (MN) word lists were to contain the same balance of high and low frequency words, but in these lists, each word was unrelated to the other.

Introduction: Each of the word list types were to be created using two distinct procedures. Firstly, the word frequency aspects of list creation were to be resolved using the MRC Psycholinguistic database (Coltheart, 1981). This allowed the experimenter to download all words with a low frequency of occurrence for which information on other, confounding variables was available (concreteness, imagability and age of acquisition). Similarly, all high-frequency words with similar available information was obtained. These two source lists then formed the pool from which the experimenter created semantically categorised and uncategorised lists.

Participants: Ten participants were drawn from the staff and postgraduate student corpus at Aston University. None of the participants were used in any other study or pilot study.

Design: There was only one variable of interest in this study, namely categorisation status of the word lists. This had two levels: categorised or uncategorised. The experiment followed a repeated measures design.

Materials: Two sets of words were obtained from the MRC Psycholinguistic Database (Coltheart, 1981). All words with a frequency of occurrence of 10 or less per million were obtained, with the proviso that information about each word existed on the following criteria:

concreteness, age of acquisition, familiarity and imagability. These were the low-frequency words. Similarly, words with a frequency of occurrence of 50 or more per million were gathered, also subject to the availability of information on the three factors of concreteness, age of acquisition and imagability. These were the high-frequency words.

Procedure: The experimenter manipulated some of these words into 12 eight-word lists which were subjectively thought to contain words which were related to each other semantically (for example, *pleasure* and *humour* were thought to be related, as were *sunshine* and *summer*). No formal method was applied in determining semantic relatedness at this stage. Each list was comprised of four high frequency words and four low frequency words. A further 12 lists were created by purposely choosing words which appeared subjectively to have no obvious semantic relationship with each other. Across both list types, care was taken to ensure that the mean concreteness, age of acquisition and imagability scores for each list did not greatly differ from any other list. Each of these factors are scored on the MRC Psycholinguistic Database on a scale from 0 to 700. The means and standard deviations for mean scores across all 24 lists for these factors were: concreteness, mean rating = 458.4 (SD= 42.5); age of acquisition, mean rating = 412.6 (SD= 41.2); familiarity, mean rating = 487.5 (SD= 27.2) and imagability mean rating = 468.5 (SD= 46.2). All scores for all of the confounding factors were within 100 rating points of the mean values across all lists. The word lists created can be found in tables A3.1 and A3.2.

The next step of the process was to assess how effective the subjective semantic categorisation had been for each of the lists created. Participants were each given a booklet containing the 24 word lists in a random order, four lists to a page. Each booklet used a different randomised order for the word lists and for the words witihin lists. Participants were asked to rate each list on the basis of semantic relatedness alone. The exact instructions they were given were as follows:

On the following pages, you will find 24 lists of words. Your task is to simply decide how categorised each word list is in terms of the meaning of the words. You are asked to give *each list a score between 1 and 7. If you think that the words on the list are all highly related, then give that list a 7. If you think that the words are not at all related, then give that list a 1. Use numbers between 1 and 7 for lists that are not so obvious. You have as much time as you require to complete this booklet.*

Results:

Mean ratings across list types were obtained. Categorised lists (mean: 5.46; SD= .19) were rated as more semantically related than non-categorised lists (mean: 1.76; SD= .16). This difference was evident in all 10 participants, and was significant in a repeated measures t-test design, using means across list types as data points (paired t = 29.16, df = 9, p<.001).

Conclusions:

This experiment created two types of word list, covering two variables, word frequency and semantic categorisation. Furthermore, these stimuli were controlled in terms of the effect of a number of confounding factors, namely, concreteness, age of acquisition, familiarity and imagability. Hence, it was hoped that these lists would prove a valuable tool in determining the processes involved in memory for linguistic items. The lists themselves are presented in the following tables.

Table A.4.1: Mixed-Frequency, Noncategorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
RULER	3	571	555	543	311
YUCCA	1	136	316	167	694
GARMENT	6	440	552	507	453
PRELUDE	5	383	364	340	572
DEVICE	55	500	444	391	472
CONTROL	223	559	311	347	411
FELLOW	63	475	502	435	378
EXTRA	50	570	262	337	331
SHUTTER	5	398	562	533	389
FORFEIT	3	380	295	310	533
ARMOUR	2	406	591	536	400
CANOE	7	441	623	602	394
MINUTE	53	621	361	473	264
CHILDHOOD	50	515	335	489	372
MURDER	75	528	445	549	381
PLEASURE	62	583	302	511	394
BROCADE	3	298	540	416	592
BERRY	9	470	573	551	289
SELLER	6	459	444	427	381
MICA	1	253	474	397	626
PERIOD	265	573	358	429	458
NOVEL	59	530	529	547	475
ESCAPE	65	524	341	459	353
PLATFORM	72	498	547	529	386
ALGAE	7	317	545	424	631
MONSOON	3	336	508	498	508
JUDGEMENT	1	506	260	333	497
TREMOR	2	401	487	491	511
LEADER	74	559	487	502	353
ISLAND	167	507	596	643	289
COLUMN	71	519	520	491	436
PURPOSE	149	572	280	280	428

Table A.4.1 (continued): Mixed-Frequency, Noncategorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
EXPANSE	5	433	388	497	547
WOODLAND	2	423	585	608	381
FALSEHOOD	2	348	260	326	531
ECHO	10	478	415	556	356
COUNTRY	324	592	465	539	329
FASHION	69	548	356	474	467
SYSTEM	416	588	356	340	461
MILLION	204	519	364	440	419
TUNIC	1	373	563	508	406
HYGIENE	3	493	349	459	481
OUTHOUSE	1	378	573	494	461
TOPIC	9	539	366	364	483
THEORY	129	534	287	317	557
CONCERN	98	519	509	353	444
PROBLEM	313	596	360	411	367
PUBLIC	438	602	356	448	400
ANNEX	1	390	464	435	536
SORROW	9	486	282	429	394
SEQUEL	1	338	353	323	556
PLAZA	2	228	479	351	642
SIGNAL	63	507	464	513	367
BODY	276	610	568	614	267
PEOPLE	847	628	540	548	281
SERIES	130	525	373	398	461
LIAR	3	534	409	425	308
ELBOW	10	564	607	602	237
AERIAL	8	481	517	567	428
LINEN	6	515	581	551	386
PERMIT	77	516	399	388	458
UNCLE	57	557	580	574	192
PICTURE	162	597	579	581	219
OPENING	83	542	455	462	336

Table A.4.1 (continued): Mixed-Frequency, Noncategorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
FOOTSTEP	3	517	470	526	317
MUZZLE	10	369	585	513	461
FLORA	1	338	557	472	569
BOATSWAIN	3	230	543	343	578
WINTER	83	615	499	621	236
TROUBLE	134	590	310	395	322
REPORT	174	590	417	411	386
BACKGROUND	67	540	383	427	400
CIDER	2	507	626	626	431
SYNOD	1	215	399	229	683
OFFAL	1	276	583	402	633
PUPPY	2	522	623	635	203
INSTANCE	82	501	284	250	471
RIFLE	63	477	606	581	322
EXCHANGE	70	504	356	394	443
SECOND	373	598	344	371	289
MENU	5	550	555	613	433
HERRING	2	425	617	524	397
WAREHOUSE	4	449	578	502	492
FAIRY	4	471	433	536	242
PRESENCE	76	514	326	339	456
AREA	323	560	384	394	392
WATER	442	641	616	632	153
MARKET	155	518	551	583	328
NOZZLE	4	412	555	513	411
AGUE	1	196	365	224	661
COSTUME	10	456	544	538	392
MISTRESS	5	472	530	535	517
METHOD	142	556	303	304	481
FREEDOM	128	568	277	437	425
MATTER	308	563	439	298	411
MOMENT	246	560	301	334	350

Table A.4.2: Mixed-Frequency, categorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
SEGMENT	10	451	485	480	428
MARGIN	10	499	472	494	403
RULER	3	571	555	543	311
TALLY	4	374	331	308	536
AREA	323	560	384	394	392
BACKGROUND	67	540	383	427	400
CIRCLE	60	581	515	591	214
DESIGN	114	538	444	407	403

CHASSIS	1	327	561	386	589
EXHAUST	7	510	467	520	429
LABOUR	4	559	406	424	506
PEDAL	4	512	602	556	306
METAL	61	559	582	541	308
MOTOR	56	545	565	521	344
PROBLEM	313	596	360	411	367
TROUBLE	134	590	310	395	322

MONARCH	3	428	525	572	456
DEFENCE	1	479	346	413	447
TRAVELLER	3	494	492	491	364
KEEPER	3	464	459	421	392
MEMBER	137	573	455	399	392
NATION	139	508	415	436	425
COUNTRY	324	592	465	539	329
PEOPLE	847	628	540	548	281

MONSOON	3	336	508	498	508
SUNSHINE	8	627	527	655	206
TWILIGHT	4	423	467	588	411
SKYLARK	1	350	614	548	394
EVENING	133	630	439	559	303
SEASON	105	565	445	495	328
SUMMER	134	612	439	618	253
WEATHER	69	623	439	537	292

Table A.4.2 (continued): Mixed-Frequency, categorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
AGUE	1	196	365	224	661
SHIVER	4	517	455	578	308
SPASM	3	422	439	486	558
ULCER	5	423	558	516	517
PATIENT	86	538	487	526	366
SILENCE	52	570	352	470	333
TREATMENT	127	529	343	408	433
PRESENCE	76	514	326	339	456

MAIDEN	2	374	545	554	429
DAMSEL	1	292	544	551	406
FOREARM	3	474	608	529	436
BOSOM	8	425	552	593	489
BEAUTY	71	561	336	513	344
BODY	276	610	568	614	267
LADY	80	573	564	571	231
SHOULDER	61	553	589	577	264

CATFISH	2	416	614	544	411
CANAL	3	464	598	588	350
DELTA	7	359	494	499	554
BOATSWAIN	3	230	543	343	578
SUPPLY	102	534	368	340	417
WATER	442	641	616	632	153
SURFACE	200	573	447	453	378
SYSTEM	416	588	356	340	461

ATHLETE	9	482	545	591	428
DREAMER	2	517	442	507	350
ANKLE	8	543	608	613	264
FIELDER	3	365	509	472	458
ATTEMPT	95	558	313	302	392
EFFORT	145	585	296	367	411
RESULT	244	623	318	324	406
SUCCESS	93	568	295	443	411

Table A.4.2 (continued): Mixed-Frequency, categorised words used in Chapter Two.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
SAGA	7	388	366	345	506
SEQUEL	1	338	353	323	556
PRELUDE	5	383	364	340	572
IMPRINT	1	439	421	332	542
NOVEL	59	530	529	547	475
SUBJECT	161	590	406	418	417
WRITING	117	630	467	540	256
PAPER	157	635	599	590	229
CREEPER	1	390	555	526	431
FLORA	1	338	557	472	569
THICKET	1	361	571	511	469
LILY	1	437	609	541	317
COUNTRY	324	592	465	539	329
FOREST	66	513	609	633	297
NATURE	191	535	414	513	342
VALLEY	73	515	575	600	339
ALLY	9	410	485	453	446
COMRADE	4	397	497	515	492
DEFENCE	1	479	346	413	447
ONSLAUGHT	4	338	346	374	600
ARMY	132	555	543	578	317
CONFLICT	52	494	305	432	497
DANGER	70	557	338	505	300
ESCAPE	65	524	341	459	353
AUDIT	4	326	371	253	642
BANKER	5	524	547	565	392
RECEIPT	4	498	474	432	517
BEQUEST	5	316	349	294	600
MONEY	265	631	574	604	247
BALANCE	90	545	366	429	367
STATEMENT	141	505	379	386	481
INCOME	109	521	429	475	506

Appendix Five: Pilot Study for Creation of Word Lists Used in Chapter Three

Pilot Study 3:

Aim: To create four types of word list, for use as stimuli in the experiments described in chapter two. The four different types were represented by lists containing purely low-frequency words in semantically uncategorised lists (LFN), purely low-frequency words in semantically categorised lists (LFC), purely high-frequency words in semantically uncategorised lists (HFN) and purely high-frequency words in semantically categorised lists (HFC).

Introduction: Each of the four word list types were to be created using two distinct procedures. Firstly, the word frequency aspects of list creation were to be resolved using the Oxford Psycholinguistic database (Coltheart, 1981). This allowed the experimenter to download all words with a low frequency of occurrence for which information on other, confounding variables was available (concreteness, imagability, familiarity and age of acquisition). Similarly, all high-frequency words with similar available information was obtained. These two source lists then formed the pools from which the experimenter created semantically categorised and uncategorised lists. The assumption that semantic categorisation was valid was tested by asking individuals to rate lists on the basis of the semantic relatedness of the words in each list.

Participants: Thirteen participants were drawn from the Staff and postgraduate student corpus at Aston University. None of the participants were used in any other study or pilot study.

Design: There was only one variable of interest in this study, namely categorisation status of the word lists. This had two levels: categorised or uncategorised. The experiment followed a repeated measures design.

Materials: Two pools of 144 bisyllabic words were obtained from the Oxford Psycholinguistic Database (Coltheart, 1981). All words with a frequency of occurrence less than 8 per million were obtained, with the proviso that information about each word existed on the following criteria: concreteness, age of acquisition, familiarity and imagability. This formed the low-frequency word pool. Similarly, words with a frequency of occurrence of 50 or more per million were gathered, also subject to the availability of information on the four factors of concreteness, age of acquisition, familiarity and imagability. This formed the high-frequency word pool.

Procedure: Within each pool, the experimenter manipulated these words into eight-word lists which were subjectively thought to contain words semantically related to each other (for example, *science* and *knowledge* were thought to be related, as were *nature* and *country*). No formal method was applied in determining semantic relatedness at this stage. Two further groups of lists, LFN and HFN, were created by purposely choosing words which appeared subjectively to have no obvious semantic relationship with each other. No word from a categorised list was ever repeated in a non-categorised list. Across all four list types, care was taken to ensure that the mean concreteness, age of acquisition and imagability scores for each list did not greatly differ from any other list.. Each of these factors are scored on the Oxford Psycholinguistic Database on a scale from 0 to 700. The means and standard deviations for mean scores across all 32 lists for these factors were: concreteness, mean rating = 454.1 (SD= 65.4); age of acquisition, mean rating = 408.7 (SD= 58.1); familiarity, mean rating = 491.3 (SD= 76.0) and imagability mean rating = 470.2 (SD= 53.7). Despite an attempt to ensure that all scores for all of the confounding factors were within 100 rating points of the mean values across all lists, four lists of the 32 did not conform to this standard. Two LFC lists had mean concreteness scores outside this margin, as did two HFC lists. In addition, one LFC list and one HFC list from these four outliers also had an imagability rating outside the 100 point margin. However, for each of these four lists, all values were within 120 rating points of the grand mean. All lists created can be found in tables A5.1- A5.4, below.

The next step of the process was to assess how effective the subjective semantic categorisation had been for each of the lists created. Participants were each given a booklet containing the 32 word lists in a random order, four lists to a page. Each booklet used a different randomised order for the word lists and for words within lists. Participants were asked to rate each list on the basis of semantic relatedness alone. The exact instructions they were given were as follows:

On the following pages, you will find 32 lists of words. Your task is to simply decide how categorised each word list is in terms of the meaning of the words. You are asked to give each list a score between 1 and 7. If you think that the words on the list are all highly related, then give that list a 7. If you think that the words are not at all related, then give that list a 1. Use numbers between 1 and 7 for lists that are not so obvious. You have as much time as you require to complete this booklet.

Results:

Mean ratings across list types were obtained. Categorised lists (mean: 5.90; SD= .45) were rated as more semantically related than non-categorised lists (mean:1.89; SD= .47). All 13 participants rated categorised lists as being more semantically related than uncategorised lists, and this difference was significant in a repeated measures t-test design, using means across list types as data points (*paired* t = 26.29, df = 12, p<.001).

Conclusions:

This experiment created four types of word list, covering two variables, word frequency and semantic categorisation. Furthermore, these stimuli were controlled in terms of the effect of a number of confounding factors, namely, concreteness, age of acquisition and imagability. Hence, it was hoped that these lists would prove a valuable tool in determining the processes involved in memory for linguistic items. The lists themselves are presented in the following tables.

Table A.5.1: Low-Frequency, Categorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
CANAL	3	464	598	588	350
CANOE	7	441	623	602	394
BOATSWAIN	3	230	543	343	578
CORAL	5	425	572	561	434
DELTA	7	359	494	499	554
HADDOCK	1	475	636	532	389
MERMAID	1	391	494	578	322
MACKEREL	2	398	636	540	436

AUDIT	4	326	371	253	642
BANKER	5	524	547	565	392
BARGAIN	7	552	399	505	394
BEQUEST	5	316	349	294	600
FRANCHISE	5	313	360	309	650
MARGIN	10	499	472	494	403
RECEIPT	4	498	474	432	517
TALLY	4	374	331	308	536

GODHEAD	1	272	358	324	614
GRIEVANCE	3	434	295	340	542
MORTAL	10	454	406	402	508
OMEN	2	394	319	413	558
STEEPLE	9	405	561	559	361
SYNOD	1	215	399	229	683
VIGIL	1	370	353	426	550
BLESSING	10	483	277	422	392

ALLY	9	410	485	453	446
HONOUR	2	523	258	363	439
ARMOUR	2	406	591	536	400
COMRADE	4	397	497	515	492
DEFENCE	1	479	346	413	447
ONSLAUGHT	4	338	346	374	600
OUTPOST	3	368	462	378	481
DUEL	5	439	456	494	408

Table A.5.1 (continued): Low-Frequency, Categorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
BARON	2	339	513	498	472
DESPOT	2	313	364	297	650
GENTRY	1	309	452	462	556
HENCHMAN	1	273	503	447	578
PEASANT	7	422	550	540	419
MONARCH	3	428	525	572	456
RULER	3	571	555	543	311
SQUIRE	5	323	502	459	525
MENU	5	550	555	613	433
HERRING	2	425	617	524	397
OVEN	7	577	593	599	236
APPLE	9	598	620	637	211
NAPKIN	3	495	585	582	342
OFFAL	1	276	583	402	633
GRAVY	4	522	606	594	269
SALAD	9	554	595	623	342
BRISTLE	3	461	558	562	383
BUTTON	10	573	613	580	192
COSTUME	10	456	544	538	392
GARMENT	6	440	552	507	453
GAUNTLET	2	269	538	450	511
HALTER	1	374	550	453	511
TAILOR	2	417	535	499	378
TUNIC	1	373	563	508	406
ALGAE	7	317	545	424	631
BERRY	9	470	573	551	289
CEDAR	1	381	608	516	425
CREEPER	1	390	555	526	431
FLORA	1	338	557	472	569
THICKET	1	361	571	511	469
WILLOW	9	425	589	565	386
WOODLAND	2	423	585	608	381

Table A.5.2: Low-Frequency, Noncategorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
UNREST	5	444	324	379	481
PIPING	5	451	538	491	436
KEEPER	3	464	459	421	392
PALETTE	5	301	565	437	542
ROBIN	2	487	637	615	233
ASSENT	4	325	311	345	575
BROCADE	3	298	540	416	592
STOIC	3	302	305	340	639
HOBBY	4	518	449	494	361
HARNESS	10	421	563	513	458
FIELDER	3	365	509	472	458
CIDER	2	507	626	626	431
FIDDLE	2	465	582	555	367
MICA	1	253	474	397	626
ABYSS	4	293	450	453	597
JUDGEMENT	1	506	260	333	497
CRAVEN	2	288	295	262	608
CANTEEN	2	490	587	540	436
MONSOON	3	336	508	498	508
DISCORD	1	381	298	343	611
SEQUEL	1	338	353	323	556
SORROW	9	486	282	429	394
SELLER	6	459	444	427	381
ECHO	10	478	415	556	356
RATTLE	5	448	549	554	261
BOSOM	8	425	552	593	489
IMPRINT	1	439	421	332	542
SKYLARK	1	350	614	548	394
TORMENT	4	398	288	386	472
REGRET	9	529	260	359	428
TURNER	6	368	462	345	517
DAMSEL	1	292	544	551	406

Table A.5.2 (continued): Low-Frequency, Noncategorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
FOOTSTEP	3	517	470	526	317
POLKA	1	281	477	475	571
SUNSHINE	8	627	527	655	206
TREMOR	2	401	487	491	511
PARDON	8	532	307	355	342
CHERRY	6	514	611	582	317
EXPANSE	5	433	388	497	547
MISTRESS	5	472	530	535	517
WARRIOR	5	368	525	553	374
HUMOUR	1	555	309	462	417
MADMAN	2	407	470	545	417
DECREE	3	376	385	341	583
PEDAL	4	512	602	556	306
PIGEON	3	499	609	610	325
ANKLE	8	543	608	613	264
BEGGAR	2	435	533	593	364
BAGPIPE	1	397	601	594	377
AERIAL	8	481	517	567	428
FAIRY	4	471	433	536	242
RIOT	7	490	414	548	456
LABOUR	4	559	406	424	506
SEGMENT	10	451	485	480	428
TRAVELLER	3	494	492	491	364
CLEARANCE	4	418	328	319	547
KETTLE	3	551	602	594	274
TOKEN	10	473	467	416	450
ATHLETE	9	482	545	591	428
FRENZY	6	409	303	450	506
JEWEL	1	519	594	621	292
MAIDEN	2	374	545	554	429
VETO	10	361	326	336	636
ACRE	9	427	462	464	411

Table A.5.3: High-Frequency, Noncategorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
SUCCESS	93	568	295	443	411
PLEASURE	62	583	302	511	394
DEVICE	55	500	444	391	472
HERO	52	510	428	483	361
SURPRISE	51	583	326	451	322
COVER	88	597	502	443	289
MURDER	75	528	445	549	381
SHOULDER	61	553	589	577	264
BODY	276	610	568	614	267
REPORT	174	590	417	411	386
TREATMENT	127	529	343	408	433
BREAKFAST	53	657	576	586	233
PATTERN	113	555	472	453	319
FAILURE	89	542	282	437	439
JUSTICE	114	522	307	379	500
OBJECT	65	586	487	408	339
ENTRANCE	57	555	484	493	372
MOTOR	56	545	565	521	344
PERMIT	77	516	399	388	458
EXTRA	50	570	262	337	331
MINUTE	53	621	361	473	264
PRESENCE	76	514	326	339	456
WOMAN	224	623	580	626	258
LEADER	74	559	487	502	353
CONTENT	53	553	300	391	389
SIGNAL	63	507	464	513	367
LADY	80	573	564	571	231
TRIAL	134	509	446	516	433
RELIEF	66	551	303	432	443
TITLE	77	523	384	413	375
PRESSURE	185	544	379	446	444
HOTEL	126	565	591	597	308

Table A.5.3 (continued): High-Frequency, Noncategorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
THEORY	129	534	287	317	557
PATIENT	86	538	487	526	366
FASHION	69	548	356	474	467
STATION	105	548	572	554	300
TROUBLE	134	590	310	395	322
PURPOSE	149	572	280	280	428
RESPECT	125	571	280	343	433
NATION	139	508	415	436	425
ADVICE	51	566	291	352	425
CONTROL	223	559	311	347	411
KNOWLEDGE	145	575	278	348	477
SILENCE	52	570	352	470	333
WELCOME	50	529	350	470	342
INCREASE	195	590	315	356	419
METAL	61	559	582	541	308
RESULT	244	623	318	324	406
OPENING	83	542	455	462	336
ANSWER	152	605	397	368	294
EFFECT	213	602	295	280	414
ISSUE	152	525	338	315	472
RECORD	137	609	558	591	322
DANGER	70	557	338	505	300
BEDROOM	52	646	615	629	206
ATTEMPT	95	558	313	302	392
ITEM	54	545	436	369	406
MEMBER	137	573	455	399	392
CAREER	67	589	373	418	514
MUSIC	216	599	512	549	272
SURFACE	200	573	447	453	378
ESCAPE	65	524	341	459	353
STANDARD	110	556	324	319	472
METHOD	142	556	303	304	481

Table A.5.4: High-Frequency, categorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
BROTHER	73	598	585	589	219
CHILDHOOD	50	515	335	489	372
COUSIN	51	515	502	478	278
FELLOW	63	475	502	435	378
JUNIOR	75	470	384	391	356
MARRIAGE	95	559	398	556	383
PEOPLE	847	628	540	548	281
UNCLE	57	557	580	574	192

DEGREE	125	574	406	521	508
DISTANCE	108	594	353	432	344
DIFFERENCE	148	592	270	293	378
EXTENT	110	510	267	248	489
MEASURE	91	555	366	379	344
NUMBER	472	599	395	489	239
PORTION	62	507	384	399	411
FACTOR	71	499	328	269	474

LETTER	145	610	577	595	256
NOTICE	59	634	479	467	369
NOVEL	59	530	529	547	475
PAPER	157	635	599	590	229
STUDENT	131	632	549	603	481
SUBJECT	161	590	406	418	417
TEACHER	80	599	569	575	247
WRITING	117	630	467	540	256

COUNTRY	324	592	465	539	329
FOREST	66	513	609	633	297
NATURE	191	535	414	513	342
PRODUCE	82	534	432	396	431
CATTLE	97	511	600	619	261
VALLEY	73	515	575	600	339
VILLAGE	72	524	576	578	317
ISLAND	167	507	596	643	289

Table A.5.4 (continued): High-Frequency, categorised words used in Chapter Three.

WORD	FREQUENCY	FAMILIARITY	CONCRETENESS	IMAGABILITY	AGE OF ACQUISITION
FUNCTION	113	518	343	294	511
ORDER	376	570	344	352	344
SYSTEM	416	588	356	340	461
SERIES	130	525	373	398	461
AREA	323	560	384	394	392
BACKGROUND	67	540	383	427	400
CIRCLE	60	581	515	591	214
DESIGN	114	538	444	407	403
EVENING	133	630	439	559	303
FUTURE	227	612	311	413	414
SEASON	105	565	445	495	328
SUMMER	134	612	439	618	253
WEATHER	69	623	439	537	292
WINTER	83	615	499	621	236
WATER	442	641	616	632	153
EXTREME	62	557	265	332	458
BALANCE	90	545	366	429	367
BUDGET	59	517	366	394	497
INTEREST	330	572	305	359	411
INCOME	109	521	429	475	506
MARKET	155	518	551	583	328
MONEY	265	631	574	604	247
PAYMENT	53	527	432	472	439
STATEMENT	141	505	379	386	481
APPEAL	62	514	333	402	431
AUDIENCE	115	511	515	555	425
CONGRESS	152	389	384	356	575
COUNCIL	103	508	435	405	464
MEETING	159	575	403	451	356
PLATFORM	72	498	547	529	386
PUBLIC	438	602	356	448	400
QUESTION	257	588	387	398	314

Appendix Six: Instructions given to participants in Chapter Three

EXPERIMENT SIX INSTRUCTIONS

Speaking task instructions:

In each trial, you will be shown eight words, one at a time. Your task is to remember these words. Throughout the whole time, you are asked to say the words '1, 2, 3, 4' out loud.

Once you have seen the words, you will be asked to write down all the words you can remember on the sheets provided. Use a new sheet for every trial. Throughout recall, you are asked to continue saying '1, 2, 3, 4'. You will be informed when you are allowed to move on to the next trial by a message on screen.

There will be 16 trials in this condition. You may take a break in between trials for as long as you require. If you have any questions, now is the time to ask them.

START YOUR SPEAKING TASK (1-2, 3-4, 1-2, 3-4, ...) AND THEN PRESS ANY KEY WHEN YOU ARE READY TO BEGIN

Tapping task instructions:

In each trial, you will be shown eight words, one at a time. Your task is to remember these words. Throughout the whole time, you are asked to tap a beat with your non-writing hand.

Once you have seen the words, you will be asked to write down all the words you can remember on the sheets provided. Use a new sheet for every trial. Throughout recall, you are asked to tap the beat. You will be informed when you are allowed to move on to the next trial by a message on screen.

There will be 16 trials in this condition. You may take a break in between trials for as long as you require. If you have any questions, now is the time to ask them.

START YOUR TAPPING TASK (1-2, 3-4, 1-2, 3-4...) AND THEN PRESS ANY KEY WHEN YOU ARE READY TO BEGIN

EXPERIMENT SEVEN INSTRUCTIONS:

Instructions: In this experiment, you are asked to remember lists of words that will appear, one at a time on the screen. Do not speak whilst the words are being shown.

After every eight words, five stars will appear on the screen (*****). This is the signal for you to begin to recall the eight words you have seen.

Try to remember all eight words. Say each word loudly and clearly into the microphone as you remember it.

If you have any questions, now is the time to ask the experimenter. Otherwise, the experiment will begin when you press any key.

Appendix Seven: Statistical Summaries for Chapter Three

EXPERIMENT SIX

Table A.6.1. Experiment Six: 2x2x2 Repeated Measures ANOVA on Free Recall means

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
ARTIC	31.250	1	31.250	68.361	.000
Error(ARTIC)	17.828	39	.457		
FREQ	9.800	1	9.800	14.256	.001
Error(FREQ)	26.809	39	.687		
CAT	50.007	1	50.007	120.883	.000
Error(CAT)	16.134	39	.414		
ARTIC * FREQ	2.113	1	2.113	3.359	.074
Error(ARTIC*FREQ)	24.528	39	.629		
ARTIC * CAT	1.445	1	1.445	3.368	.074
Error(ARTIC*CAT)	16.727	39	.429		
FREQ * CAT	.851	1	.851	2.508	.121
Error(FREQ*CAT)	13.227	39	.339		
ARTIC * FREQ * CAT	9.453E-02	1	9.453E-02	.228	.635
Error(ARTIC*FREQ*CA	16.140	39	.414		

Table A.6.2. Experiment Six: 2x2x2 Repeated Measures ANOVA on Asch-Ebenholtz means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
ARTIC	.868	1	.868	35.772	.000
Error(ARTIC)	.946	39	2.427E-02		
FREQ	.115	1	.115	5.827	.021
Error(FREQ)	.767	39	1.966E-02		
CAT	.218	1	.218	11.800	.001
Error(CAT)	.720	39	1.846E-02		
ARTIC * FREQ	6.375E-03	1	6.375E-03	.477	.494
Error(ARTIC*FREQ)	.521	39	1.336E-02		
ARTIC * CAT	1.985E-02	1	1.985E-02	1.513	.226
Error(ARTIC*CAT)	.512	39	1.312E-02		
FREQ * CAT	2.371E-02	1	2.371E-02	1.366	.250
Error(FREQ*CAT)	.677	39	1.735E-02		
ARTIC * FREQ * CAT	1.031E-04	1	1.031E-04	.008	.930
Error(ARTIC*FREQ*CAT	.511	39	1.309E-02		

EXPERIMENT SEVEN

Table A.6.3. Experiment Six: 2x2 Repeated Measures ANOVA on Free Recall means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	3.896	1	3.896	31.129	.000
Error(FREQ)	2.879	23	.125		
CAT	16.154	1	16.154	111.374	.000
Error(CAT)	3.336	23	.145		
FREQ * CAT	.570	1	.570	4.837	.038
Error(FREQ*CAT)	2.712	23	.118		

Table A.6.4. Experiment Six: 2x2 Repeated Measures ANOVA on Asch-Ebenholtz means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	.165	1	.165	11.643	.002
Error(FREQ)	.326	23	1.417E-02		
CAT	1.042E-02	1	1.042E-02	.742	.398
Error(CAT)	.323	23	1.404E-02		
FREQ * CAT	2.600E-02	1	2.600E-02	3.712	.066
Error(FREQ*CAT)	.161	23	7.006E-03		

Table A.6.5. Experiment Six: 2x2 Repeated Measures ANOVA on Speech Rate Means

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	1441.733	1	1441.733	1.540	.227
Error(FREQ)	21526.967	23	935.955		
CAT	2086.842	1	2086.842	7.731	.011
Error(CAT)	6208.738	23	269.945		
FREQ * CAT	4261.202	1	4261.202	33.128	.000
Error(FREQ*CAT)	2958.436	23	128.628		

Table A.6.6.a. Experiment Six: 2x2 Repeated Measures ANCOVA SummaryTable with Speech Rate as a covariate on Free Recall means- Repeated Measures effects

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	.106	1	.106	.816	.376
FREQ * SPTRUEAV	1.234E-02	1	1.234E-02	.095	.761
Error(FREQ)	2.866	22	.130		
CAT	.826	1	.826	5.837	.024
CAT * SPTRUEAV	.223	1	.223	1.579	.222
Error(CAT)	3.113	22	.141		
FREQ * CAT	.166	1	.166	1.405	.248
FREQ * CAT * SPTRUEAV	.107	1	.107	.905	.352
Error(FREQ*CAT)	2.605	22	.118		

Table A.6.6.b. Experiment Six: 2x2 Repeated Measures ANCOVA SummaryTable with Speech Rate as a covariate on Free Recall means- Covariate Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	14.085	1	14.085	18.068	.000
SPTRUEAV	.619	1	.619	.793	.383
Error	17.150	22	.780		

Table A.6.7.a. Experiment Six: 2x2 Repeated Measures ANCOVA SummaryTable with Speech Rate as a covariate on Asch-Ebenholtz means- Repeated Measures effects

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
FREQ	2.549E-03	1	2.549E-03	.177	.678
FREQ * SPTRUEAV	9.105E-03	1	9.105E-03	.632	.435
Error(FREQ)	.317	22	1.440E-02		
CAT	6.204E-03	1	6.204E-03	.434	.517
CAT * SPTRUEAV	8.182E-03	1	8.182E-03	.572	.458
Error(CAT)	.315	22	1.431E-02		
FREQ * CAT	6.736E-04	1	6.736E-04	.093	.763
FREQ * CAT * SPTRUEAV	1.921E-03	1	1.921E-03	.265	.612
Error(FREQ*CAT)	.159	22	7.238E-03		

Table A.6.7.b. Experiment Six: 2x2 Repeated Measures ANCOVA SummaryTable with Speech Rate as a covariate on Asch-Ebenholtz means- Covariate Effects

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	.331	1	.331	4.228	.052
SPTRUEAV	2.127E-03	1	2.127E-03	.027	.871
Error	1.723	22	7.832E-02		