D7470087

IV

373

D 74700/87

KEATING G.C.

373

CITY OF LONDON POLY (CNAA).

THE EFFECTS OF WORD CHARACTERISTICS ON CHILDREN'S READING

Geraldine Corriene Keating

Department of Psychology
City of London Polytechnic

Thesis submitted in partial fulfilment for the degree of
Doctor of Philosophy to the
Council for National Academic Awards
March 1987

CONTENTS

List of Tables in Text

List of Tables in Text

List of Figures

Acknowledgements

I am especially grateful to my supervisors Dr Veronika Coltheart and Dr Veronica Laxon for the encouragement and constructive advice they have given me throughout every phase of this research.

My thanks also go to the Economic and Social Science Research Council who supported me financially during the three years of my research studentship.

Finally, I am indebted to the teachers and children of Sir John Cass School, Wellington School, Ben Jonson School, Christchurch School and Thomas Buxton School for their interest and enthusiastic cooperation.

THE EFFECTS OF WORD CHARACTERISTICS ON CHILDREN'S READING

G.C. KEATING

ABSTRACT

The object of the research reported in this thesis was to investigate the effects of word characteristics on children's reading performance.

The experiments investigating word imagery and age of acquisition showed that imagery was a highly significant word characteristic for less skilled readers. There was an age of acquisition effect which was inversely correlated with reading ability.

Probabilistic measures of orthographic regularity (such as Initial Bigram Frequency and Versatility and First Order approximation to English) were shown to be significant predictors of reading for good and poor readers and lexical decision performance for average readers. It also appeared that as reading ability improved, word properties such as the Orthographic Neighbour Ratio, which takes into account neighourhood size and frequency affected reading accuracy in the good and average reader in the lexical decision task.

Other measures of orthographic regularity-orthographic neighbourhood size and body type were also shown to affect reading accuracy although effects appeared less marked for skilled readers. The regularity effect was seen to be dependent upon hostility and frequency of word neighbours, and the frequency of the target word itself, rather than due to a regularity-irregularity dichotomy.

C H A P T E R   O N E

READING RESEARCH - A GENERAL INTRODUCTION

1.1 INTRODUCTION

Contemporary reading research is a rich mixture of influences from cognitive and physiological psychology, linguistics, computer sciences, social psychology and educational practice. The influences range from the most abstract, theoretical points of view to the most practical, applied situations. Current research in reading consequently has at least two clear, definable thrusts. First, research is aimed at understanding the basic nature of the reading process. Attempts to do this include the generation of models and theories of the reading process. The second thrust is a renewed and increasingly intense search for better methods of teaching with the ultimate goal of improving education and reducing illiteracy.

Kerlinger (1977) has described the difficulties of "applied" research. He states, rather pessimistically that it is rare that researchers from either end of the basic-applied continuum interact to generate mutually interacting research problems.

Reading practitioners have often been impatient with a perceived lack of consensus among reading researchers, leading to a lack of consistency in recommendations for improving classroom practice. Researchers have similarly been loath to tackle the "messy" problems of reading instructions without having precise details of the parameters of the reading and teaching process used. The hope for reading research would seem to lie in increased dialogue between researchers and practitioners.

The split between practitioners and theorists has produced a predictable difference in type of reading research efforts. On the practical side, there are many evaluations of individual programs, methods or techniques. The question in each case is whether one program or method produces "better" learning than another. Theoreticians, on the other hand tend to generate programs of research in which the outcome of each study bears on the design of the next - the selection of variables and methodologies. Usually the outcomes of studies are used as feedback to revise the framework or model governing the research program.

Reading research is, and has been, undergoing changes that seem similar to what Kuhn (1962) has termed a paradigm shift. Kamil (1984) suggested that the reasons for the changes have been, firstly, an emphasis on the reader as an active information processor; secondly, the development of comprehensive systems of discourse analysis that could be applied to reading (e.g. Halliday & Hasan, 1976; Kintsch & Van Dijk, 1978) and thirdly, an increased interdisciplinary interest in precisely translating research into practice.

There would seem to be several distinctly identifiable goals for current reading research - research to generate a theory of reading, research of a more limited kind - to collect data, and research to make instructional decisions. Not all reading research necessarily yields explanatory information, e.g. word lists or frequency counts are the results of descriptive research. They are helpful in that they provide measurable and stable differences in variables which can then be manipulated experimentally.

1.2 THE READING PROCESS

Among the many skills in the repertoire of the average adult, reading is one of the most complex. The journey taken by words from their written form on the page to the eventual activation of their meaning involves several stages of information processing. For the fluent reader, this processing takes only a fraction of a second. The acquisition of reading skills can take years, and there are many who do not succeed in becoming fluent readers, even though they may have quickly and easily mastered the skill of undertanding speech. It has been assumed that the normal reader can read aloud by transforming the orthographic representation of letter string into a phonological form without the involvement of semantic processing, and also that certain acquired dyslexic patients read aloud using basically the same procedure (Coltheart, 1978; Morton & Patterson, 1980; Shallice & Warrington, 1980). The various theories of word recognition have not yet reached agreement in their description of the reading process. On the other hand, as researchers have begun to appreciate the range and complexityof phenomenon to beaccounted for, and as theorists have responded with increased complexity in their accounts, the differences between them do not seem so great. This point is illustrated very well in the recent review of the literature by Humphreys and Evett (1985). One of the aims of this thesis, therefore, is to show that the various models have more in common than previously thought.

1.3 NEUROLOGICAL STUDIES OF READING

A great deal of research has been undertaken in terms of impairment to components of a model of normal reading and isolation of what cannot be done under those circumstances (e.g. Coltheart, 1980 Patterson, 1981, Dérousné and Beauvois, 1979). While this approach has yielded much of value to our knowledge of reading, the growing research on acquired and developmental dyslexias will not be dealt with in this thesis, except to state that much of our knowledge of errors in nonword reading comes from studies of phonological and surface dyslexias. (Patterson, 1982; Marshall & Newcombe, 1973).

As Patterson and Morton (1985) succinctly pointed out, in the 1970's "the only major variables which seemed germane to research on assembling phonology from print were the distinction between words and nonwords and, for words, the distinction between regular and irregular (or 'exceptional') spelling-to-sound correspondences" (p.5, 1985).

The experimental work presented in this thesis aims to show that the variables which affect the reading process are slightly more complicated than that.

The intention of this thesis is to assess the effects of a variety of word characteristics on the psychological processes involved in young children's reading and to draw some conclusions from the results as to which theories best describe it. Attempts are made to trace the development of orthographic knowledge and other word variables and to demonstrate the part played by these in reading, so as to arrive at certain conclusions as to the nature of reading and of the difficulties in learning to read.

C H A P T E R T W O

METHODOLOGICAL ISSUES

2.1 INTRODUCTION

The two different measures of children's reading used in the experimental work reported in this thesis, are the lexical decision task and the reading task. They have been used to evaluate the effects of various structural characteristics of words (e.g. orthographic regularity) on reading. Therefore this chapter provides a general description of the lexical decision task, and the assumptions made in the lexical decision task. A review of recent experiments suggests that the results of lexical decision tasks should be treated cautiously because lexical decision is affected by semantic variables (Whaley, 1978; Balota and Chumbley, 1984, 1985) and proportion of words to nonwords (Seidenberg, Waters, Sanders and Langer, 1984). It is argued that the lexical decision task remains a useful methodological tool however, because various theories predict differing effects of, for example, regularity on lexical decision and reading tasks.

The chapter describes the reading task; how it differs from the lexical decision task and why some researchers prefer it e.g. Henderson, 1984.

More general issues, such as the applicability of experiments using single word presentation, methodological considerations when using the lexical decision task, and how reaction time data is usually treated in reading research are then discussed. It should be noted that all the lexical decision and reading task experiments described in this chapter involved skilled readers.

2.2 GENERAL DESCRIPTION OF THE LEXICAL DECISION TASK AND ITS ASSUMPTIONS

In studying variables that affect the speed of lexical access, researchers have relied heavily upon the lexical decision task (LDT), where the subject determined whether a letter string was or was not a word, either by using a key press or by making a verbal response, YES or NO. Latency of response to a word was seen to reflect the time taken to access that word from the lexicon and then to decide that the accessed lexical entry was the correct one. Latency of response to a nonword was seen to reflect the time taken to decide that there were no lexical representations that sufficiently matched with the stimulus item.

When the manipulation of a variable causes a corresponding variation in response latency in the LDT, it has usually been assumed that the variable is having an effect on the ease of extracting sufficient information from a letter string to recognise it as a word, i.e. to access its lexical representation. The assumption is that lexical access is the only process in the LDT being affected by the manipulated variable. Balota and Chumbley (1984) presented experiments which may lead to the validity of this assumption being questioned. They argued that the demand characteristics of the decision process in the LDT may result in an exaggerated role of word frequency in lexical access. Their experiments suggest that the familiarity/meaningfulness dimension is very important in the lexical decision task.

Chumbley and Balota (1984) showed that a word's meaning affected the decision in a lexical decision task in a series of four experiments. Subjects in two experiments made lexical decision judgements, those in a third experiment read the words used in the lexical decision task, and those in a fourth experiment produced speeded word associations to the words. Differences in lexical access time for the words were measured in the reading task, and differences in meaning were assessed with the association task. Multiple regression analyses of lexical decision reaction time (RT) were conducted using associative RT, pronunciation RT, and other target word properties (frequency, length, instance dominance, and number of dictionary meanings) as predictor variables. They found a relationship between lexical decision RT and associative RT after the effects of the other variables had been partialled out. In addition, word frequency continued to have a significant relationship to lexical decision RT beyond that shared with pronunciation RT and the other variables. They considered that these results indicated that some of the effects of word meaning and word frequency in lexical decision was attributable to a decision stage following lexical access.

Balota and Chumbley (1984) again showed that semantic variables affected the decision stage in a lexical decision task. They first used a simple category - exemplar verification task in which a category name (e.g. bird) was presented and then followed 800 ms later by an exemplar from that category (e.g. robin) or from a different category (e.g. sofa). The subject's task was to make a YES-NO judgement about the validity of the category exemplar relationship being presented. They investigated the impact of five lexical variables (instance dominance, category dominance, word frequency, word length in letters, and word length in syllables) in a full multiple regression analysis. The full analysis was used instead of a stepwise analysis because they had defined in advance which variables were of theoretical importance, and they were not primarily interested in comparing the relative importance of the variables within a task. Although the results of their study yielded an interesting pattern of effects of instance dominance (likelihood of producing the exemplar given the category name) and category dominance (availability of the category name given the exemplar) on verification time, there was no influence of word frequency on either trials in which a YES response was correct, or on trials in which a NO response was correct. Balota and Chumbley considered that the word frequency effect found in previous studies was actually an effect of other variables, such as Instance Dominance. They then conducted a lexical decision experiment and a reading experiment with the same set of predictor variables. Results of the lexical decision experiment yielded large effects of frequency over and above the effects of the other variables. It was also found that both Category Dominance and Instance Dominance significantly predicted lexical decision performance. This finding is important because it had previously been argued that semantic information, because it was not logically necessary to the task, became available only after lexical access (although the semantic priming data of Meyer et al 1973, 1975, demonstrated the role of semantic variables in the LDT). The results indicate that either this argument is incorrect or that some other component of lexical decision is sensitive to semantic variables. Thus, there is evidence that the LDT may not be the perfect tool for studying a lexical access process that is presumed to be unaffected by semantic variables.

A number of theorists (West and Stanovich, 1982) had suggested that the LDT involves postrecognition processing that may influence performance. They suggested that the pronunciation task may be a better reflection of pure lexical access. Balota & Chumbley (1984) found a large effect of frequency on lexical decision, a moderate effect on pronunciation, and a very small effect on category verification. These results clearly indicate that word frequency has dramatically different effects, depending upon the task that is used to assess lexical access. One possible reason that the LDT produces such large frequency effects is that the task places a premium on frequency information at the decision stage of the task. Balota & Chumbley point to the fact that lexical decisions are typically more than twice as long (500-600ms) as normal reading rates (250 ms/word) in adults, and suggest that there is much more involved in the LDT than simple lexical access. It is necessary to consider more closely the task faced by the subject when making a lexical decision. The subject is asked to discriminate meaningful stimuli from nonword letter strings that have never been seen before. The two most obvious pieces of information the subject could use to make this discrimination are the frequency with which the stimulus has been seen before and its meaningfulness. Balota and Chumbley found useful a basic notion that words and nonwords differ on a familiarity/meaningfulness dimension. A particular letter string's value on this FM dimension is based primarily on its orthographic and phonological similarity to actual words. The word and nonword distributions on the FM dimension are separated and overlap.

The subject can use this fact in following the LDT instructions to both maximise speed and minimize errors. Because some word targets are relatively much more discriminable from the nonword distractors (and vice versa), the subject can set two criteria that will allow rapid decisions for at least some of the stimuli being presented (see Figure 2.1). A low criterion could be set so that very few words will have FM values that would fall below this criterion. Similarly, a high criterion could be set so that very few nonwords will have FM values that would fall above this criterion. If the FM value falls between the upper and lower criteria, then the subject needs more information before the decision can be made. The necessary information is obtained by performing a more analytic evaluation of the letter string. For example, the subject may actually need to check the spelling of the letter string against the spelling of a word in the subject's lexicon. This extra analysis requires additional time - thus longer latencies.

Balota and Chumbley consider that the demand on the subject in the category verification task is completely different. A discrimination between words and nonwords is not

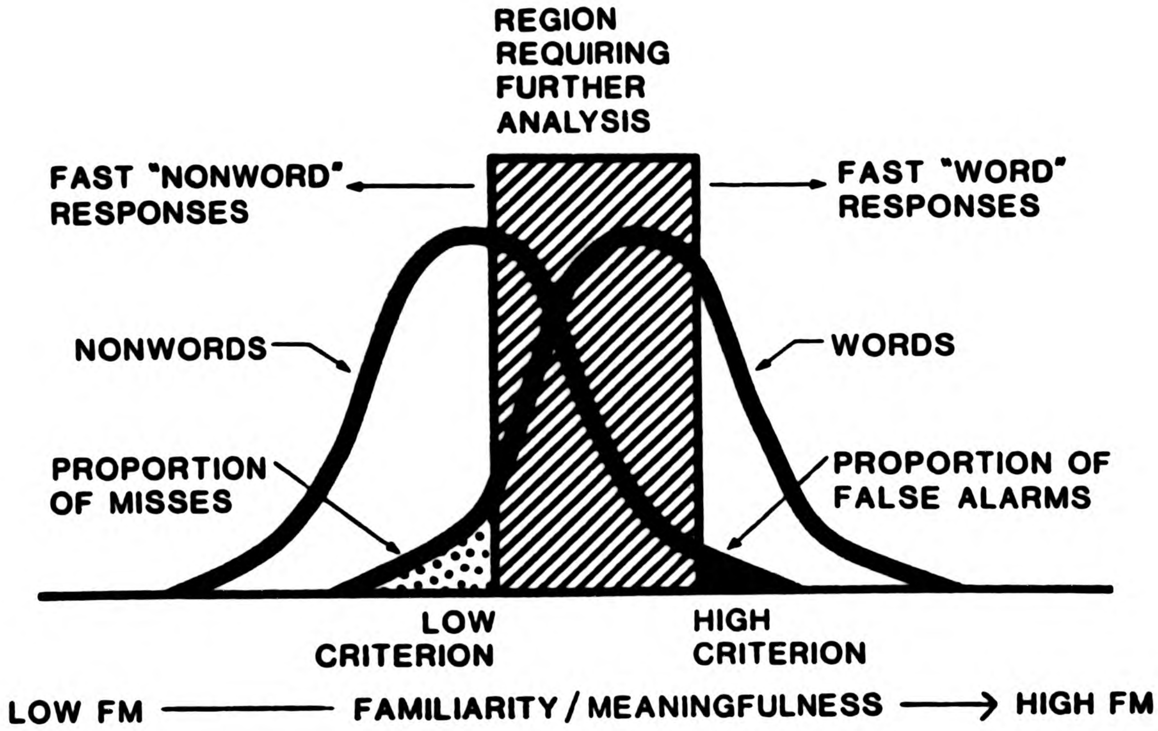


FIG 2.1
WORD AND NONWORD DISTRIBUTIONS ALONG THE FAMILIARITY/MEANINGFULNESS DIMENSION
BALOTA & CHUMBLEY (1984)

required - meaningfulness is more relevant. Word frequency could affect the time to determine what is known about a word (lexical access), but this encoding process may be a very small part of the overall judgement process and there would only be very small frequency effects. Similarly in reading, where the primary task is the extraction of meaning, frequency of occurrence is much less important than it is for the lexical decision task. Thus when using lexical decision tasks to ascertain the importance of certain variables in lexical access, one must take into account that word frequency may well swamp that effect.

It is interesting to compare Balota and Chumbley's meaningfulness dimension with Patterson and Morton's (1985) predictions for naming latencies in their modified standard model (described in Chapter 3). In their model, the mean time for the lexical routes (visual and semantic) is faster than for the nonlexical route (OPC, GPC) but the distributions have considerable overlap, as is the case with Balota and Chumbley's word and nonword distribution.

As Henderson (1982) emphasised, the data on naming latencies for most words and nonwords indicated such overlap, though the overlap may be negligible in the case of high frequency words (Seidenberg et al, 1983).

Whaley (1978) demonstrated that a cluster of semantic variables (concreteness, meaningfulness, imagery and age of acquisition) significantly predicted lexical decision performance above and beyond obvious lexical variables such as word frequency and word length. This finding is important because it implies that semantic characteristics of the word being perceived in isolation are affecting the ease of lexical access and the lexical decision task involves components other than lexical access that are sensitive to the semantic characteristics of the word being judged.

The lexical decision task has been used to evaluate the effects of various structural characteristics of words (e.g. orthographic regularity) on recognition. These effects in lexical decision vary greatly depending on experiment-specific factors. Waters and Seidenberg (1983) showed that effects of spelling-sound irregularity in lexical decision depend upon the composition of the stimuli i.e. a mixed pool of words, but this is not the case in reading performance. The different lexical decision results follow from the fact that performance depends on the discriminability of words and nonwords, and subject biases influenced by the proportion of items from different conditions. The variability in lexical decision performance suggests that although lexical decision latencies were found to be typically longer than naming latencies (Forster & Chambers, 1973; Frederiksen & Kroll 1976) this is not a necessary outcome - it merely reflects the conditions that are typical of word recognition experiments. Seidenberg, Waters, Sanders and Langer suggest that the lexical decision task should be used with extreme caution in experiments that introduce fewer extraneous factors. Seidenberg (1985) reviewed the effects of literal and virtual context in word recognition. He suggested that what was important was the reader's knowledge of the structural relationships that hold among words. This information - which is in the reader's head rather than the actual text - represents the virtual context of occurrence, and it is the focus of current theories of the recognition process.

He found task-related differences in the access of phonology. The occurrence of the exception effect in the lexical decision task depended on the composition of the stimuli. Where a mixed pool of words was used and there was no time limit for a response, a phonological effect was found. In reading the effect occurred regardless of the composition of the stimuli. These results follow from the different demands of the two tasks.

West and Stanovich (1982) suggested that conclusions from lexical decision experiments about the role of sentential context in word recognition were suspect, since the effect could be occurring at either a lexical access or a decision stage. More recently Seidenberg, Waters, Sanders, and Langer (1984) compared priming effects in lexical decision and pronunciation tasks and concluded that semantic and associative priming effects were present in both tasks but that other contextual effects found with lexical decision tasks were due to postlexical processes. They used both lexical decision and pronunciation tasks because previous research had suggested that they were differentially sensitive to post-lexical processing. Pre-lexical processing was defined as the decoding processes that resulted in identification of the signal as a particular word. Post-lexical processes occur after recognition is completed. These post-lexical processes involve the selection, elaboration and integration of lexical information for the purpose of comprehending a text or utterance.

Their first experiment examined the syntactic priming effect using both lexical decision and naming.

STIMULI CONDITIONS	PRIME		TARGET
1) SYNTACTICALLY RELATED PAIRS	men	-	swear
	whose	-	planet
2) UNRELATED PAIRS	men	-	planet
	whose	-	swear
3) PSEUDOWORD PAIRS			

Subjects were instructed to read the prime word silently, and then to respond to the target. Results showed that synactically related pairs were facilitated in lexical decision, in naming there was only a small amount of facilitation. They were consistent with the hypothesis that associative priming and syntactic priming result from different processes; the former due to spreading activation through lexical memory, the latter to subjects' post-lexical decision that stimuli are grammatical.

Seidenberg, Waters, Sanders and Langer (1984) examined the effect of the proportion of related stimuli on the magnitude of the associative priming in both tasks, the proportion of related stimuli only influenced lexical decision. These results were interpreted as being consistent with the view that lexical priming derives from two sources. First there is an associative priming effect that appears in both tasks. This effect is an automatic consequence of spreading activation. Secondly, subjects' expectation concerning the relatedness of stimulus pairs can produce facilitation; and because these effects are post-lexical they appear in lexical decision, but not naming.

Seidenberg et al considered the priming effect due to backward association observed by Koriat (1981). They hypothesized that forward priming was due to spreading activation and that backward priming was due to a post-lexical process. Forward priming should occur with both tasks, while backward priming should only occur in lexical decision. The stimuli consisted of pairs of words that were symmetrical associates (e.g. sheep-wool) and pairs of words that were asymmetrical associates (e.g. fruit-fly) highly related in the forward direction, unrelated in the backward direction. Results showed that forward priming occurred in both tasks, while backward priming only occurred in lexical decision.

The results of these studies firstly emphasize the importance of considering the loci of contextual effects on word recognition, secondly they show that lexical decisions are influenced by post-recognition judgements of prime-target relations more than naming. An explanation of the different pattern of results follows from an understanding of the requirements of each task.

2.3 THE READING TASK

As Glushko noted "Reading aloud is a valuable skill, and most people learn to read by first learning to read aloud." (Glushko, 1979, p.1). The plausibility and success of reading aloud as a criterion measure in reading makes it natural to adopt the naming task to study the information - processing components they share with word recognition and reading. Thus in some studies "word naming" i.e. reading aloud is tested - the ability to read aloud lists of isolated words, without much regard to their meaning. Tests of oral reading must be administered individually, and are thus too time consuming for large-scale surveys. It is customary in these to employ tests which depend upon the comprehension of sentences, as demonstrated by the capacity to fill in missing words or answer questions on context. We shall be more concerned in this study with more circumscribed studies related to the processes involved in reading than with large-scale surveys and no detailed investigation of reading tests is intended.

Henderson (1984) focused more on the reading task, which he termed "the speeded naming task", than the lexical decision task. He considered that experimenters turned to the lexical decision task in hope of determining whether phonological recoding played a role in lexical access, whereas the reading task seemed to offer a more direct means of investigating the mechanisms of phonological recoding.

He saw the virtue of the reading task as compelling phonological recoding by some means or other. Its limitations were that the latencies may include a component due to articulatory programming (Sternberg, Monsell, Knoll and Wright 1978), though such effects were likely to be small. Secondly there was a logical possibility that subjects could stop the clock with the initial articulation before assembling the whole of a word's phonology; and thirdly the inferences linking latency effects to particular processing mechanisms are highly indirect.

Balota and Chumbley (1985) also investigated how useful the reading task was for studying word recognition. They pointed out that one cannot assume that a variable affects lexical access simply because it affects reading latencies, as it is not known whether the lexical access component or the production component is affected. If the reading task is a useful method for investigating how frequency influences lexical access, then frequency effects should be limited only to the short delay intervals (at most 400 milliseconds) and should disappear after the subject has had time to recognise the stimulus. On the other hand, if there are production-frequency effects then frequency effects would be expected even at the longer delay intervals.

Balota and Chumbley used a delayed reading task in which a word was presented and, after some delay, a cue was presented to pronounce the word aloud. Presumably, subjects should access the word's lexical representation during the delay interval and any effect of frequency can be attributed to its influence on production rather than on lexical access.

In their first two experiments frequency still had a significant effect with a delay interval of 900 milliseconds. A significant word-frequency effect at a delay interval of 2900 milliseconds was found in the third experiment where any rehearsal during the delay interval was disrupted by asking subjects to whisper the alphabet while waiting for the cue to respond. The results of these experiments show that there are quite large frequency effects in the reading task which are traceable to the output of the stimulus word rather than only to its encoding i.e. there are frequency effects at the production stage.

The question remains as to why the lexical decision task should be used in conjunction with the reading task. Andrews showed that different theoretical predictions can be made for the lexical decision task and the reading task. She attempted to distinguish between the dual-access and the activation and synthesis models of word recognition. She did this by manipulating word regularity and consistency, looking at the effects of these manipulations in the lexical decision and reading task. It was necessary to look at both these tasks because of the predictions made by the two models. Glushko's (1979) model proposed that consistency effects arise because of the difficulty of synthesising a pronunciation from inconsistent activated information. He predicted that consistency effects would be evident in the word naming task but he would not expect effects of consistency in the lexical decision task because pronunciation was not required.

The dual-access model predicted that regularity effects would arise whenever the phonological route completed before the visual route. Coltheart (1978) suggested that this is rarely the case in the lexical decision task, but may be more likely in the naming task.

The results for the lexical decision task showed that there were effects of consistency and essentially no effects of regularity. The data seem to support Glushko's (1979) activation and synthesis model. However, the implication of Glushko's discussion of his model was that consistency effects arise in the course of synthesising a pronunciation. In order to account for the fact that such effects occurred in the lexical decision task, in which a pronunciation is not required, it would be necessary to assume that inconsistent information causes general problems in deciding which particular lexical entry matches the stimulus, as opposed to specific problems in generating a pronunciation.

In the reading task, the existence of consistency effects again supported Glushko's model, but the presence of a regularity effect for low frequency words in the naming-words-only task support the dual-access model. If there is no phonological route, as Glushko suggests, then there should be no effects of regularity. Andrews' data therefore causes problems for both the one-route and two-route models. She considered that, in general, the more detailed specification of the activation procedure described by McClelland the Rumelhart (1981) in their interactive activation model seemed more capable of explaining the data.

The Andrews paper illustrates that different theoretical predictions can be made for the lexical decision task and the reading task, and since many of the following experiments reported in later chapters explore further the concept of consistency, it was decided that both reading and lexical decision data would be collected, bearing in mind the fact that the lexical decisions can be influenced by semantic variables.

In lexical decision the performance depends upon the discriminability of words and nonwords, and the subject's response criteria, both of which may vary across experiments and individuals. The reading task is different in at least two respects. Firstly, the subject does not have to discriminate between words and nonwords; secondly, subject's responses are constrained by the requirement that they pronounce stimuli correctly.

The basic difference between the tasks is not that lexical decisions are influenced by post-lexical processing, while naming is uninfluenced. Rather, because of the signal detection character of the task, lexical decisions are more likely to be influenced by such processes. These differences between the tasks are of great relevance to studies of single word recognition.

2.4 EXPERIMENTAL DESIGNS IN READING RESEARCH

Several different experimental designs have been used in studies of the reading process. One is the reading-age match. This simply means that the two groups being compared are at the same level of reading. In contrast there is the design - the mental-age match, because it compares backward readers with other children of the same age and intelligence. This design is considered ambiguous by Bryant and Bradley (1985), because differing results may be due to a difference in reading level.

The longitudinal method deals with the same children over a period of time. The most interesting period to choose in longitudinal studies of learning to read goes from the time before children go to school and begin to learn to read through to several years after they have arrived at school. Although longitudinal studies are very valuable, a discovery that one variable precedes and is related to another is not on its own enough to establish a causal link.

Psychological researchers have long recognised the difficulties inherent in using group data in analyses (Deese & Hulse, 1967). Some researchers have consistently advocated studying single subjects across an extended range of time. Perhaps the most prevalent paradigm in reading research using the single/small N methodology has been that of miscue analysis (e.g. Goodman, 1969). Most miscue studies utilise anlaysis of extensive data collected in individual sessions with, at most a few readers. Because only a few subjects are involved, the collected data can be analysed intensively. There is also a similar growing emphasis in psychological research on analysis of errors. A comprehensive review of oral reading error analysis can be found in Leu's (1982) review.

However, the risk involved in doing research with single/small N samples is the potential for studying non-representative individuals. Conclusions based on a nonrepresentative sample are less appropriate for extended applications. The smaller the sample, the greater the risk of sampling error. This paradigm necessitates repeated measures, as do the experiments reported in this thesis. Both the benefits and disadvantages of repeated measure designs are involved. Repeated measures allow each individual to serve as his or her own control, reducing the variance and increasing the precision of the statistical tests. When repeated measures are used, however, contrast effects may arise. That is, subjects may react differentially to the different treatments, only because they realise the treatments are different. In addition, repeated measures are subject to practice and fatigue effects, so this has to be taken into consideration, especially with children.

Reading research often can be characterised by the number of factors manipulated. Much reading research has involved manipulating only a single variable. It is becoming increasingly common for researchers to use designs that manipulate several variables simultaneously. This reflects a realisation that clusters of variables have to be studied to arrive at a thorough description or explanation of the reading process. Greater reliance on computer-aided analysis will make the use of these designs more manageable and more likely to be used.

2.5 GENERAL ISSUES IN THE LEXICAL DECISION AND READING TASK

The discussion of the decoding of words in lexical access and pronunciation has concentrated on experiments using single-word presentations. It could be argued that what has been said has little, if any relevance to the reading of words in text. However, recent experimental evidence suggests that context does not qualitatively alter the lexical access process; that is, word recognition in text is the same as word recognition out of text (except perhaps when the word is highly predictable from the context). The few studies on morphological processing that have presented words in context (the letter cancellation work of Drewnowski and Healey, 1980, and Smith and Sterling, 1982) have produced results that are consistent with the results of single-word studies. Thus it seems fair to say that single-word paradigms are a valid way of examining the reading process.

Santee and Egeth (1982) argued that reaction time and accuracy were not necessarily interchangeable measures of the same underlying process, and in support of their argument showed that the pattern of a set of results depended upon how performance was measured. Under data-limited viewing conditions (the short exposure durations of the typical tachistoscopic task), response accuracy was sensitive to early perceptual interference between target and noise items, whereas reaction time was more sensitive to later processes involved in response interference.

Under resource-limited viewing conditions (the long exposure durations of the typical reaction time task), both accuracy and reaction time were affected by processes occurring in the later rather than the earlier stages of processing.

They do not suggest that the two dependent measures will reflect different processes, but that the two dependent measures are not necessarily interchangeable. The difference between accuracy and latency measures in children is more straightforward. Accuracy is a more fundamental measure than naming latency, and because of the large accuracy differences in our data it was decided that the analysis of accuracy data would be more appropriate than analysis of naming latencies. In the case of adult data collected, naming latencies were analysed where possible.

The selection of materials for the lexical decision experiment was seen by Shoben (1982) as being the most important aspect of the entire design. He also stressed the importance of controls. Word length is almost always controlled in lexical decision tasks. Although some investigators have matched length in terms of number of syllables (Schuberth and Eimas, 1977), it is more common to match in number of letters. For the most part, this distinction does not matter as it is hard to imagine a circumstance in which these two measures would not be extremely highly correlatled.

The other factor that must be controlled is word frequency, which has a clear influence in lexical decision tasks. As one would expect, more frequent items are more readily identified as words than infrequent items.

Although the issue of which variables to control in these types of experiments is not always an easy one, in some ways other questions are more difficult. In particular, the questions of which kind of nonwords to use is a difficult problem. The nature of the nonwords can clearly influence the results one obtains from the positive trials. Specifically, the use of orthographically irregular nonwords can eliminate some of the most robust effects. For example, James (1975) found that if the nonword distractors were not pronounceable (e.g. bneo), then word-nonword judgements were very easy and common effects, such as the word frequency effect, disappeared under these circumstances.

There is also some conflict as to whether the nonwords should be just orthographically regular, or whether they should be additionally constrained to be mis-spellings of target words. e.g. using a nonword target ROBEN along with targets such as ROBIN. However, it appears that there are drawbacks to choosing this method of constructing the nonwords. First, it seems that this method changes the lexical decision into a spelling test; one is no longer deciding if a string has meaning, but whether a string is properly spelled. Moreover some mis-spellings will be more visually and phonologically similar to the target word.

Probably the safest thing to do therefore would be to use pronounceable nonwords that obey the orthographic rules of English or pronounceable anagrams of the target word.

2.6 COLLECTION AND ANALYSIS OF REACTION TIME DATA

Although the actual collection of data in this field may seem straightforward, there are a number of precautions that can be taken to minimise the variability of the reaction times collected. If possible practice trials should be used. It is generally true that reaction time declines quite rapidly over the first few trials and more slowly after that. If this practice effect were the only consideration, then one would simply run a large number of practice trials. Unfortunately, large numbers of practice trials have at least two costs associated with them. First, practice trials take time, and a subject's time is usually limited. Second, large amounts of practice may change the process by which the subject performs. In the experiments reported in this study, usually a list of 10 practice items was used before the list proper, to eliminate a large part of the practice effect. It was essential to do this with young subjects especially, to get them used to the equipment and procedure.

The first question that arises in connection with the reaction times collected is whether they should be transformed. Many classical statistical books will argue that one should perform a log transformation when there is a dependancy between the mean and variance. Most, if not all reaction time distributions have this dependency. Although such a transformation may be good statistical practice, reaction times are often viewed as a kind of naturalistic measure. Unlike percent correct, which is determined to a great extent by the experiment, one can argue that reaction times reflect the time required to perform a particular cognitive operation. Current practice seems to permit raw reaction times to be used, provided extremely long reaction times have been excluded, and lacking any evidence that this procedure leads to statistical errors, it seems very reasonable to do so.

For long reaction times, there is no consensus. For a start, there is no fixed definition of what constitutes a long reaction time. Often, an arbitrary cutoff is selected. In other cases, long reaction times are defined as exceeding some number of standard deviations above the mean. Still other papers have used no cutoff at all.

If one accepts the fact that a reaction time that is six times as long as the average does not reflect the standard processing for that item, then it would seem that one would want to exclude it. These long reaction times are often the result of a subject's inattention or unfamiliarity with the item or equipment malfunction. Although excluding long reaction times seems desirable, one should not exclude too much data. If long reaction times constituted more than 2% of the data, one might reasonably be suspicious of the items which make up the list. In most studies, the excluded data constitute abut 1% of the total.

If one is willing to establish some criterion for determining long times, then the next decision is what to do with these excluded items. There have been two common procedures in the literature. First, one can treat the long times as errors. This treatment has some intuitive appeal, in that it is consistent with the idea that some kind of non-normal processing is occurring on these trials. The undesirable effect of this procedure is that it tends to reduce the mean reaction time of difficult items, because the longest times are excluded before the mean is calculated. This drawback has resulted in the practice of truncating the excluded reaction times to the criterion value. Thus, if long reactions times are defined as those that exceed 2.5 seconds, then all times over this value are entered into the analysis as 2.5 seconds. Although this procedure avoids the drawback of the count-as-error method, it would seem that it has no a priori justification. If long reaction times reflect erroneous processing, then we would not include them in the mean for the same reasons that error rection times are nearly always excluded from the determination of mean reaction time. This procedure can also lead to misleading results. Consequently, excessively long reaction times are counted as errors and not included in our analysis of correct reaction times.

There are problems of collecting reaction time data inherent in the reading task. Words with stop consonants as beginning phonemes e.g. P,T,K,C, might be produced more quickly or might activate the voice key more quickly than words which have a beginning phoneme H,CH,SH (i.e. fricatives) because of their acoustic properties.

Balota & Chumbley (1985) in a series of regression analyses showed there was no consistent pattern of significant regression coefficients with the mean response latency for each word across subjects using beginning stop consonants and fricatives categories as predictor variables. However, as a precaution, it seems sensible not to have a predominance of beginning stop consonants and fricatives in a stimulus list.

2.7 CONCLUSION

The consensus from research described in this chapter illustrates the fact that results from a lexical decision task may be less easy to interpret than results from a reading task. This is because the lexical decision task is more likely to be affected by semantic variables, frequency and stimulus composition. This does not mean that the lexical decision task is worthless and should never be used, but it indicates that it should be used with greater care - and that it should never be considered synonymous with reading. The lexical decision task is worth using because different theories of word recognition predict that lexical decision and reading will be differentially affected by regularity, consistency (Andrews, 1982), and because of the Seidenberg, Waters & Tanehaus finding that there were task-related differences in the access of phonology i.e. the exception effect always occurring in the reading task regardless of the composition of the stimuli but the exception effect depending upon a mixed pool of words in lexical decision.

C H A P T E R   T H R E E

THEORIES OF WORD RECOGNITION

3.1 INTRODUCTION

This chapter reviews the main theories which have been used to describe the reading process. There has been such a proliferation of theories that this review, must be, of necessity, a selective one.

Firstly theories of word recognition based on experiments with skilled adult readers are reviewed, because they may provide an insight into how the reading process develops in children and also because relevant data on children is sparse. Interesting questions and problems raised by these theories are then tackled in the experimental chapters which follow. Although there has been much research done on the development of reading in children, there has been less success in providing a developmental model of children's reading, except for theories that exist in the light of adult data.

Secondly a developmental model of reading is described, so that its applicability to the data collected in this research can be evaluated.

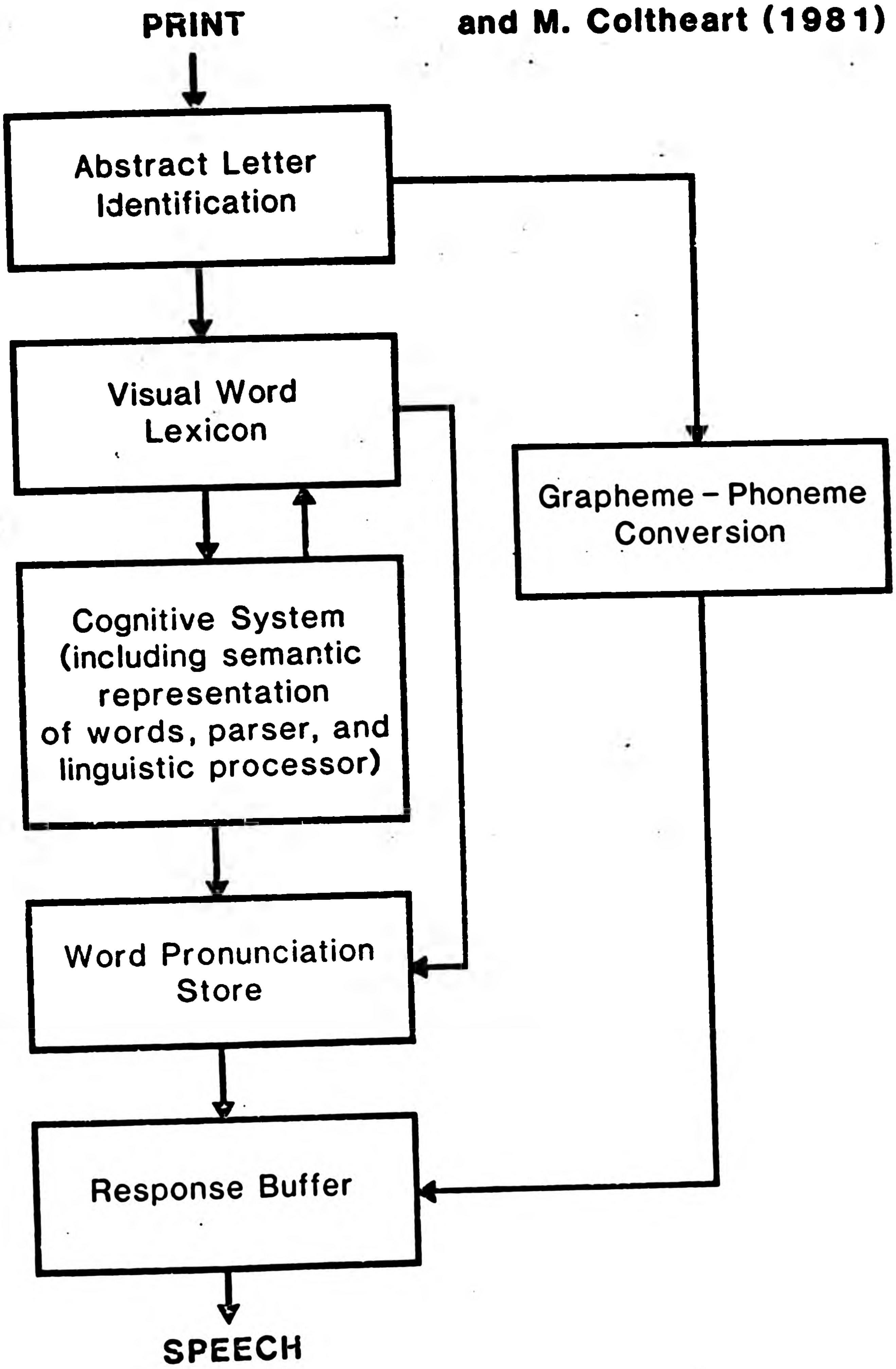
3.2 MODELS OF WORD RECOGNITION BASED ON WORK WITH ADULT READERS

It is useful to begin by considering several related models of the recognition process that have been proposed over the past fifteen years. These constitute the first information processing models of recognition, in that they attempted to represent the flow of information through a sequence of stages, beginning with sensory input and ending with the correct identification of meaning. The first proposal of this type was Morton's (1969) Logogen Model. This model has undergone successive elaborations by Morton and others (e.g. Coltheart, 1978, 1980; Morton and Patterson, 1980; Newcombe & Marshall, 1981). According to this conceptualisation a word could be pronounced either on the basis of non-lexical phonology using spelling-to-sound rules (non-lexical because the GPC rules describing mappings between orthography and phonology are not specific to particular words) or through lexical phonology - recognition through the visual pathway giving access to a stored representation of its pronunciation which can simply be read out of memory storage. In this case recognition is not phonologically mediated. (See Figure 3.1).

The dual route theory has been most explicitly applied to lexical decisions (e.g. Coltheart, Davelaar, Jonassen & Besner 1977), where it has been assumed that stimuli are automatically entered into both routes and the faster route will trigger the response decision. The model has been generalised to the naming task by Morton & Patterson (1980)

One source of evidence for the dual-process account has been the double-dissociation between the lexical and nonlexical routes in the acquired dyslexias. Considerable dispute attaches to the question of how clearly this dissociation has been demonstrated and how relevant such a demonstration might be to our understanding of normal reading.

Figure 3.1 A model of reading adapted from Morton & Patterson (1980) and M. Coltheart (1981)



As Henderson (1982) notes, many theorists have gone beyond the dual route position to the assumption that there exists parallel, independent and self-sufficient mechanisms i.e. they propose "horse-race" models e.g. Forster & Chambers (1973). A number of tests of the independent route, horse-race model are possible on normal word naming data. The fact that words can be named faster on average than pseudowords leads to the conclusion that the lexical route is faster at least on some occasions. On the other hand, effects of spelling - sound predictability on reading accuracy must lead to the conclusion that the nonlexical route is sometimes faster, or that both routes work together. It follows that the two routes must have overlapping response time distributions. Since exception words can supposedly only be read by the lexical route and pseudowords by the nonlexical GPC route, the overlap of the naming latency distributions for these two classes of item should allow us to estimate the proportion of trials on which the nonlexical route was the faster. However, when the nonlexical route is faster for an exception word, and presuming that the routes are independent, the response must be a regularisation error. It should be possible, therefore, to predict the error rate from the latency distribution overlap, provided that the exception and pseudowords are matched carefully on frequency and word length.

Questions about the nature of the non-lexical system have generally been ignored, with most theorists content simply to label it as a system of "grapheme-phoneme correspondence rules". One often-quoted exception was the attempt by Coltheart (1978) to provide a more detailed GPC model of non-lexical reading. Two principle stages are distinguished in Coltheart's model. During an initial segmentation stage, functional spelling units (letters or groups of letters) which correspond with individual phonemes are marked. At the second stage of phonemic translation, phonemic values are accorded to segmented functional units with the aid of context-sensitive rules. Such rules are assumed to be based on the formal linguistic rules of English pronunciation such as those developed by Wijk (1966) and Venezky (1970).

The corpora of spelling-sound correspondence rules described by Wijk and Venezky are based upon a TYPE count of the number of words in which each correspondence occurs in a particular orthographic context. This reflects a general linguistic tendency to determine the regularity of particular alternations by counting TYPES (the number of occurrences of each type of pattern under stated conditions), rather than TOKENS (the frequency of occurrence of each type of pattern in the language). Thus, the short pronunciation of a as /æ/) is derived from a count of the number of monosyllabic words BACK, MAP, ALP, versus ALL and does not take into account the frequency of occurrence of a particular type - so although ALL is the most frequent occurrence it is still considered to be irregular.

According to the theoretical underpinnings of the model, GPC rules are assigned in an all or none manner. The normal correspondence is always applied, assuming that the

GPC look up table is complete, so if the system is faced with an existing word with minor correspondences an incorrect rule-governed pronunciation is assigned.

It is important to note that translation rules which operate at the level of functional spelling unit and phoneme do not take into account the regularity or predictability which may exist between larger clusters of letters and corresponding sound patterns. If we take TALK as an example - it is irregular at the GPC level because the commonest pronunciation of the vowel in this context is /æ/. When other words ending in -ALK are considered, it is clear that they are all pronounced in the same way - and the correspondence is therefore entirely predictable and consistent.

Thus, the GPC system, as described by Coltheart (1978) can successfully construct rule-governed pronunciations for unfamiliar letter strings. It can supply correct pronunciations for the majority of English words which obey GPC rules. On the other hand, exception words to these general rules e.g.STEAK where the major pronunciation differs from that given by GPC's, are treated as if they were regular and therefore mispronounced i.e. regularised.

There is considerable experimental evidence for the involvement of a GPC system in word naming. In the first study to investigate speeded naming of regular and exception words with adults, Baron and Strawson (1976) found a large latency difference between the two word groups of 165 milliseconds a word. However their chosen reaction time measure was the average time to read through each list, so response times for wrong pronunciations were also included. This makes it very difficult to compare their results with other reaction time studies in which mean correct reaction time is usually analysed. Baron and Strawson suggested that their results showed that exception words were more difficult, because a GPC procedure would yield incorrect regularised pronunciations.

Stanovich and Bauer (1978) also showed a significant reaction time advantage of 18 milliseconds for regular words and further showed there were no differences in the production latencies between the regular and exception words. Other replications of this small, but robust effect followed.

Glushko (1979) reported a significant difference between regular and exception words of 20 milliseconds; for Seidenberg et al (1984) this difference was 36 milliseconds, although they observed that the exception effect disappeared altogether with high-frequency stimuli.

Although higher error rates for exception words were also observed, the nature of these errors had generally not been reported. Glushko (1979), however, showed that many exception word errors were regularisations which appears to indicate the involvement of a GPC system.

Within the two-process framework, a number of different explanations of the exception effect in reading aloud have been advanced. They all share the basic assumption that the effect is a consequence of the fact that for exception words, GPC procedures will yield incorrect regularised pronunciations. The two-process framework advocates that the effect is due to differences in processing time between lexical and non-lexical systems which race against each other to produce a pronunciation. In the "race horse" system, the product of the faster process is selected and any subsequent candidate is inhibited. In this way, any conflict which might arise between alternative specifications is avoided.

However, it is clear that an assumption that either the lexical process or the GPC process is invariably faster cannot account for mean latency differences between regular and exception words. If lexical processing were always faster, as has been postulated by Forster and Chambers, 1973; Coltheart, Besner, Jonasson and Davelaar, 1979), then exception words should be named as quickly as regular words and never given regularised pronunciations.

The simplest course for two-process theorists is perhaps to retreat to the claim that lexical and GPC systems have more or less overlapping distributions, with neither process being predominantly faster. Regular words would be pronounced faster than exception words on average, since they would be correctly pronounced no matter which process wins. However, if the distributions of processing time for the two systems were assumed to overlap completely, it becomes difficult to account for the latency advantage for words compared with nonword pronunciations.

An alternative type of account holds that all candidate pronunciations of the written stimulus are considered at a decision stage common to both processes.

Clearly, for exception words, lexical and GPC pronunciations will conflict, and it is the resolution of conflict which is responsible for the latency delay.

One suggestion as to how the conflict is resolved is that a lexical check is carried out which matches with spelling and sound. A second suggestion is that GPC rules are assigned not in an all-or-none fashion, but recursively, so that less frequent correspondences are applied when conflict arises. However both views beg the question of why conflict should have arisen in the first place, since lexically supplied pronunciations are always correct for both word types.

For this reason, these versions of the conflict hypothesis seem improbable. However, if the notion of conflict is entertained as a credible explanation of the exception effect, then one is forced to concede that the two systems are not totally independent.

In spite of the evidence to support Coltheart's account of non-lexical processing, it has had problems in explaining the results of Glushko (1979) in a series of naming tasks. In his first experiment, Glushko used regular and exception words and nonwords. An exception effect was again recorded. There was also an exception effect with nonwords e.g. the exception nonword HEAF took longer to pronounce than regular nonwords HEAN and there was a greater tendency to regularise the exception nonwords (84% of all errors). In his third experiment Glushko used a three-way classification: words are not regular or exception in themselves, but only in the context of the other words that are activated in the course of reading them. This implies that words can be exceptions, regular and consistent or regular and inconsistent. Glushko found that the two classes of words that produce inconsistent activation of orthographic and phonological structure took longer to pronounce than the regular and consistent words. Moreover, pronunciation latencies of exception/inconsistent and regular/inconsistent words did not differ significantly (although there was a much higher error rate for exception/inconsistent words). So, when compared with regular/consistent words regular/inconsistent words behaved in much the same way as did exception words - despite the fact that they conformed to GPC rules. Like nonwords then, real words are apparently affected by the extent to which the pronunciations of visually similar words either agree or conflict. The salient factor in Glushko's experiments therefore appears not to be GPC regularity, but phonological consistency of lexical analogues. That both words and nonwords might tap knowledge of consistency through common access to visually similar words in a single system, was the theoretical alternative put forward by Glushko, that the two-process models had to contend with. Glushko's (1979) activation - synthesis model will be described later in this chapter.

Glushko's findings may be explained in terms of a two-process framework as long as certain modifications are made. Thus, one could elaborate the GPC system to include stored abstract rules about larger units such as three letter endings. Inconsistent spelling patterns would be distinguished by the existence of rules governing the assignment of alternative phonological correspondences.

Thus Glushko's experiments provided problems for the dual route model and it was viewed in a more critical light.

CRITIQUE OF THE DUAL ROUTE MODEL

Kay (1985) noted that a basic problem for the dual route model is that it has been difficult to provide a coherent formulation of the rules governing spelling-to-sound correspondences in English. GPC rules ignore the morphophonemic character of English orthography (e.g. SIGN, SIGNATURE).

In the dual route model pronouncing an irregular or exception word requires the reconciliation of an incorrect pronunciation based on the use of GPC rules with a correct pronunciation accessed post-lexically. Thus even for regular words naming always requires the use of both pathways. This is paradoxical - if word recognition ultimately requires access to post-lexical phonology, why does the reader attempt to generate a pronunciation on the basis of GPC knowledge? Possibly because the GPC route works for most words i.e. regular words, so although post-lexical phonology is necessary for naming, pre-lexical phonology speeds access to naming. The second possibility is that the use of non-lexical phonology is a reflection of the manner in which reading skills were acquired i.e. strong phonics teaching. The third possibility is that the non-lexical GPC route exists because of its role in reading unfamiliar words and nonwords, but an analogy theory can cope with this. It may be that redundancy is peculiar to the naming task and subjects will use the GPC route to try and obtain the pronunciation of a word quickly. However Seidenberg, Waters, Barnes & Tanenhaus (1984) present evidence that faster reaction times are associated with smaller effects of irregular pronunciation, so it cannot be the case that forcing subjects to name words quickly yields greater reliance on pre-lexical phonology.

The idea that subjects can control initial decoding processes - that they use different decoding 'strategies' or suppress the use of phonology has not been supported by much empirical evidence. Yet this is how the dual route theorists account for why the phonological effect is stronger for reading than for lexical decision. Coltheart et al (1979) and Seidenberg (1985) established that the exception effect is larger in reading aloud than in lexical decision; often the lexical decision yields no effect of irregular pronunciation at all e.g Shulman, Hornak and Sanders (1978) found that the phonological effect on lexical decision latencies depended on the type of nonword stimuli. If there were pronounceable nonwords then there would be phonological access, if the nonwords were unpronunceable then no phonological access was observed. The fact that access of phonology appears to depend on the type of reading task and stimulus material led Coltheart and others to conclude that its use is under the tacit control of the individual. This conflicts with another common assumption of reading researchers, namely that initial decoding processes are automatised among skilled readers (e.g. LaBerge and Samuels 1974). Perhaps the most compelling evidence is that subjects access phonological information even under conditions where it has a negative effect on performance e.g. Doctor & Coltheart (1980).

There has also been criticism of the flow-chart formalism employed by the dual route model. Seidenberg (1985) pointed out two basic problems. The boxes and arrows are arbitrary graphical conventions which are difficult to interpret theoretically. Several negative consequences follow from this practice. Firstly models cannot be evaluated in terms of parsimony - it is difficult to judge a model with four large boxes as being better than one with six small ones. Secondly arrows are used in a variety of ways. They are sometimes intended to represent the temporal sequence of processing events; at other times they merely indicate possible pathways without regard to sequence - double ended arrows.

The second basic problem is the ad hoc character of model building within this framework - more boxes and arrows are added to help explain new data.

The dual route model (as summarised by Morton & Patterson, 1980) underwent successive elaborations and modifications over the years with the introduction of a third recognition pathway and various input and output buffers. This was largely in response to neurolinguistic data indicating specific patterns of impairment in cases of acquired reading disability. However, Seidenberg still considers that the explanatory value of these models is limited because new pathways and other mechanisms can be stipulated as necessary. He raises questions as to whether any such model could be falsified, or have much predictive value.

Where the dual route model has been useful is in providing a sound framework for research on acquired dyslexias. Functional impairments of the dyslexic patient are attributed to malfunctioning of one of the routes described in the model. Hence the deep dyslexic is affected by the degree to which a word is imageable (easily imageable words are easier to read), because of some impairment in the semantic part of the model.

ACTIVATION - SYNTHESIS MODELS

Glushko (1979) preferred a more radical explanation which entirely dispensed with the notion of two routes to the pronunciation of a word. His alternative was pronunciation by analogy. Suppose a reader encounters the nonword VATE and the response is /VEIT/ so as to rhyme with RATE. Although this pronunciation could be produced by abstract grapheme-phoneme rules, it might instead have been based on a more specific rule using the multiletter correspondence of the familiar -ATE pattern, or by analogy with a word like GATE. Therefore, much of the evidence cited in support of abstract rules in reading aloud is equally consistent with the idea that readers use larger and more specific units of orthographic and phonological structure.

Glushko suggested that because English spelling is roughly phonemic does not mean that a reader who knows the relationship between print and sound will necessarily use the alphabetic principle in assigning pronunciations to letter strings.

In Glushko's simpler framework, he proposes that:-

> "words and pseudowords are pronounced through the integration of orthographic and phonological information from a number of sources that are activated in parallel, much as readers comprehend sentences by integrating lexical, syntactic and contextual information. As letter strings are identified there is parallel activation of orthographic and phonological knowledge from a number of sources in memory. This knowledge may include the stored pronunciation of the letter strings, pronunciations of words that share features with the letter string, and information about the spelling-to-sound correspondence of various subparts of the letter strings. A pronunciation is generated using procedures for determining how to modify the activated information in order to synthesize the desired articulatory program."
>
> Glushko, p.678,(1979)

Glushko's activation and synthesis proposal does not make a sharp distinction between the lexical and orthographic knowledge bases and does not assume that knowledge of spelling-to-sound regularity is organised and stored as abstract rules. In an activation framework the difference between the pronunciation of words and nonwords is only quantitative; words are generally pronounced using larger units (up to the entire letter string) than nonwords, which might be parsed into smaller units to activate analogies or specific spelling-to-sound correspondence. Rather, a word is consistent or inconsistent with the orthographic and phonological structure that it activates. Regular and exception words are pronounced using the same kinds of knowledge. Exception words are simply those words whose phonological structures are likely to conflict most strongly with other activated information. So according to Glushko "exceptions are not that exceptional".

Thus, according to Glushko's (1979) activation-and-synthesis approach, words in lexical storage are inter-connected according to orthographic and phonological similarity (where similarity is defined as all words that share the terminal vowel-consonant segment with the item in question). As a written letter string is analysed, it activates such orthographic and phonological neighbours, so that a word like SEAM, for example would activate orthographic representations and corresponding phonological forms of BEAM, TEAM, etc., as well as SEAM itself. (On a broader-based activation model, initial-position neighbours like SEAL and SEAT, as well as more remote neighbours like SENT, BEAN and FOAM, or even SORT and FARM, would be accessed; see Glushko, 1979, p.684). Prior to pronunciation, activated phonological information is somehow synthesised to produce a complete phonological response. In explaining his findings Glushko uses the concept of conflict between competing pronunciations. Thus, it is assumed that the orthography of activated words in the neighbourhood of inconsistent stimuli - words and nonwords - will also activate conflicting phonology and will therefore be slower in reaching a final consensus on pronunciation than orthographic and phonological neighbourhoods of consistent stimuli. As Glushko himself notes, however, this characterisation of processes involved in oral reading is "intentionally vague and evasive".

Glushko's activation and synthesis approach has led several researchers to consider whether a process of analogy can better account for the ability of readers to deal with unfamiliar words.

An orthographic mechanism that uses analogies with existing words need not always predict the same pronunciations as abstract rules. Ohala (1974) had people generate novel pronunciations for which analogies and rules made conflicting predictions to determine the generality with which people use orthographic and phonological regularity. Smith and Baker (1975), Baker and Smith (1976) and Steinberg and Krohn (1975) also used this conflict technique to compare analogical phonological rules with the abstract rules of Chomsky and Halle's (1968) generative phonology, where the grapheme string is used to access the abstract lexical representation. The comprehension process then assigns a surface syntactic structure to the lexical representation. To this structured representation the phonological rules can be applied and a phonetic representation derived. In all these experiments, skilled readers often used exception words as analogies to generate novel pronunciations that were irregular.

Baron (1977a, 1977b, 1979) also presented demonstrations that skilled readers can also use analogies in reading aloud. Baron defined analogy as a <u>conscious</u> strategy of recalling a similar word and then modifying its pronunciation. He found (Baron 1977b) that adult readers reported the conscious use of analogies in "giving the best pronunciation" to a pseudoword, and that subjects who volunteered an analogy strategy did slightly better than those who did not. In addition, subjects became more successful at pronouncing pseudowords when they were given explicit analogy instructions. This improvement with analogy instructions also occurred with elementary school children (Baron 1979). However, since adults can accurately name nonwords in approximately 500-600 milliseconds, the role of consciously applied strategies remains questionable.

Brooks (1977, 1978) has proposed a rather different form of analogy procedure for pronouncing novel words from Baron, Brook's analogical mechanism operates implicitly rather than explicitly; since words that look alike tend to sound alike, readers might pronounce novel words by generalisation from existing words without any awareness of the spelling-to-sound correspondences in either letter string.

Glushko (1979) proposed that the previously obtained regularity effects were the result of confounding regularity and consistency and that the presumed phonological effects were not the result of rules applied either prelexically or postlexically, but that they arose within the lexicon because of the activation of inconsistent information from visually similar words. His findings showed that when regular/ consistent, regular/inconsistent and exception words were compared, there were effects of consistency but not regularity. However Andrews (1982) considered that it was impossible to conclude this from the data because there was no irregular inconsistent condition used, and the exception words used were of a higher frequency than the regular words, which might have masked effects of regularity. She therefore attempted to distinguish between the dual access and the activation and synthesis models using a full factorial combination of regularity and consistency. Regularity was defined as conformity to GPC rules, such as those of Venezky (1970). Consistency was defined as meaning that all other words that differed from the stimulus in only the initial consonant or consonant cluster, and were of approximately the same length as the stimulus, were pronounced in the same way.

Andrews studied performance on reading aloud and a lexical decision task. Glushko's (1979) model proposed that there would be consistency effects in the reading task because consistency effects arise from the difficulty of synthesising a pronunciation from inconsistent activated pronunciation. He would not expect consistency effects in lexical decision because pronunciation is not a task requirement.

The dual access model predicted that regularity effects occur whenever the phonological route completes before the visual route. Coltheart (1978) suggested that this is rarely the case in the lexical decision task, but may be more likely in the naming task.

Andrews' results showed that there were effects of consistency, but no effects of regularity in the lexical decision task thus supporting Glushko's activation-synthesis model. She also manipulated the degree of reliance on a phonological code by comparing a condition in which half the nonword stimuli were pseudohomophones with a condition in which no homophones were present. She found that there were differences between the pseudohomophone -present and -absent conditions. This would not be predicted by Glushko's (1979) model - as there is no phonological route, there can be no mechanism by which phonological effects can appear and disappear according to task demands.

However this difference is not consistent with the dual access model either. Inclusion of pseudohomophones makes the task easier. Davelaar, Coltheart, Besner and Jonasson (1978) expected that inclusion of pseudohomophones would decrease reliance on the phonological effects, as subjects realised that if they used the phonological route then they would make lexical decision errors on pseudohomophones.

The lexical decision data are equivocal. They support Glushko's (1979) model because there are effects of consistency but not regularity. The improvement in performance resulting from adding pseudohomophones however, is difficult to interpret without assuming that there is decreased reliance on phonological information, thus providing evidence for the GPC route.

As in the lexical decision tasks, the existence of consistency effects in the reading task supports Glushko's (1979) model, but the presence of a regularity effect for low frequency words support the dual-access model.

An important aspect of Glusko's activation and synthesis model is the proposal that during the course of identifying a word, the lexical entries for other words possessing similar features are also activated. The fact that

Andrews found consistency effects in both the lexical decision task and the reading task suggests that activated inconsistent information affects a more general word identification process rather than the process involved in synthesising a pronunciation. However she considered that Glushko's implicit rejection of the GPC route may be premature.

To summarise the three main points of Glushko's analogy theory; firstly his theory does not contain a strictly nonlexical mechanism. Secondly the units that matter are those that determine similarity with neighbours. These are clusters of graphemes such as -EAST. Thirdly it is not regularity per se but the correlation of regularity with consistency of pronunciation among a group of visually similar words that is responsible for latency effects.

What remains implicit in Glushko's activation-synthesis model is the importance of orthographic units and the orthographic neighbourhood, which obviously underlie his theory. These variables are investigated in Experiment 1 and Experiment 4.

Marcel (1980) has proposed a model of oral reading which is similar to that described by Glushko (1979) in that lexical knowledge is used in assigning pronunciations to known as well as to new words. In his model, orthographic entries corresponding to known words or morphemes exist in segmentable form in a visual input lexicon. That is, constituent letters are abstractly coded in their spatial positions within each lexical address. Similarly, corresponding phonological addresses in an output lexicon are also segmentable.

Written letter strings are assumed to undergo a process of segmentation prior to lexical access. The segmentation procedures first segment the initial letter of the string, and as each subsequent letter is encountered, it is bracketed with any previous letters as an orthographic segment. Thus, given the word UNCLE, the left-to-right parse would produce the following bracketed segmentations:

U...UN...UNC...UNCL...UNCLE

As each new segmentation is yielded, previous bracketings are preserved, so that in the above example, the five alternative segmentations would all exist on completion of parsing. Marcel suggests that once a letter or sequence of letters is bracketed, the remaining letters in the string correspondingly form a potential segment:

e.g. U + NCLE, UN + CLE etc.

Marcel suggests that previous segmentations of the string may be over-ridden as "more satisfactory" larger candidate segmentations are produced.

As each segmentation is yielded and analysed by matching orthographic segments of lexical addresses, it will automatically activate corresponding phonological segments in an output lexicon.

Regular and consistent words like SEAM will generate fewer conflicting phonological correspondences than either regular and inconsistent or exception and inconsistent words. However, in the case of known words, any phonological conflict between competing segments will be resolved as the segmentation corresponding to the complete letter string is yielded, over-riding prior orthographic segmentations and producing a phonological specification for the whole string.

Nonwords would not have complete lexical specifications which would over-ride competing orthographic segmentation. Marcel (1980) suggested that their pronunciation will be a function of two factors. Firstly that segmentation most economically accounted for by known letter combinations and morphemes, or that segmentation found in most words of that consonant and vowel structure and secondly the pronunciation found for each segment in the largest number of lexical cases.

Parkin (1983) argues that analogy models could be criticised because the consistency effect could not be reliably replicated. Parkin (1983) reported a series of experiments concerned with regularity and consistency effects in the pronunciation of single words in adults. His first experiment was a straightforward replication of Glushko's (1979) findings that both exception and inconsistent words produced longer pronunciation latencies than consistent words. He found that exception words e.g. PINT produced longer pronunciation latencies than consistent words e.g. PILL. However he did <u>not</u> find any difference in the pronunciation latencies for inconsistent words e.g. GLOVE and consistent words.

His second experiment attempted to replicate Glushko's finding that nonwords based on letter patterns that are pronounced in more than one way produce longer pronunciation latencies than nonwords based on consistently pronounced letter patterns. In addition the experiment also examined how the extent to which a given letter pattern was pronounced in different ways influenced performance. Parkin notes that within Glushko's original definition of 'exception' nonwords no account was taken of the fact that the frequency with which any given letter pattern is mispronounced can vary enormously. For example DASTE was defined as an exception because of the relatively uncommon word CASTE. However BREAT has three possible correspondences, the majority exhibiting the regular correspondence /i:t/, followed by two exhibiting /et/ and finally one, the word GREAT, in which -EAT = /eit/. In terms of <u>types</u> CASTE and GREAT are similar since both are unique in their pronunciation. However, in terms of <u>tokens</u> they are very different. CASTE accounts for 9% of the -ASTE tokens in the Kucera and Francis count whereas GREAT accounts for 64% of occurrences of -EAT in the count. Thus in terms of the frequency with which a given mispronunciation occurs the latter is much more commonly encountered.

Parkin therefore subdivided inconsistent nonwords into high, medium and low token inconsistent nonwords matched with high, medium and low token consistent nonwords for initial letter and letter length. The results of his second experiment did not show any difference in the pronunciation latencies for inconsistent and consistent words, thus failing to replicate Glushko's results. However Parkin's manipulation of high and low token inconsistent items also failed to affect pronunciation latencies. He suggested that if there were an inconsistency effect then it would be independent of the number of tokens carrying a given mispronunciation. Therefore his new hypothesis suggested that the inconsistency effects with nonwords was determined by the number of types exhibiting the various pronunciations of any given letter pattern rather than the number of tokens carrying each occurrence.

In Parkin's third experiment, this possibility was explored. Pronunciation latencies were collected for inconsistent letter patterns which were commonly pronounced in more than one way. e.g. BROVE, GROVE (regular word equivalent) and for exception letter patterns which were commonly pronounced in more than one way, but with one dominant pronunciation. e.g. YINT, PINT (irregular word equivalent). Pronunciation latencies were also collected for matching consistent words and nonwords. Parkin found a consistency effect across subjects but not across materials. This suggested that there might be a consistency effect but that it is smaller than the regularity effect and therefore less likely to replicate over a small sample of inconsistent words (n=18). When pronunciation latencies were carried out for a larger sample of inconsistent words (n=32) there was a consistency effect across subjects but not materials.

Failure to find a strong consistency effect led Parkin to reject one process accounts of pronunciation espoused by Glushko (1979) and Marcel (1980). Instead he argued that a modified dual route model might provide the best account of the data from the single word pronunciation task. Thus

Parkin's results show that Glushko's model requires further experimental evidence as the consistency effect is unreliable.

McClelland and Rumelhart (1981) have provided a considerably more detailed working model of letter and word recognition from an activation-and-synthesis approach. Their model of word recognition makes use of the concept of distributed parallel processing in which complex phenomena may result from interactions among a large network of simple processing elements. Elements become activated and the pattern of activation through the network changes by means of excitatory or inhibitory spread of activation.

Strictly speaking, theirs is not a model of word recognition per se; it was developed to account for word superiority effects - that letters were perceived more rapidly within a word than in isolation.

Their model is largely concerned with ways in which structural relationships among words could influence recognition. This is one step forward from Glushko (1979) where a word is exceptional or strange only by virtue of its relationship to other words, its neighbours.

Rumelhart (1977) proposed an interactive model of reading to account for how perceptual processes were affected by context and familiarity. The central feature of this model was that the processing of information in reading was assumed to consist of a series of levels. Information could flow in both directions at once - from lower to higher levels and from higher to lower levels. The proposal that information from a higher level could feed back and affect the processing at a lower level could explain how knowledge of a higher level unit i.e. a word, may affect the processing of a lower level unit i.e. a letter.

Rumelhart and McClelland (1982) combined the fundamental features of the interactive model with the flow-of-activation assumptions of McClelland's (1979) cascade model to build the interactive-activation model (see Fig.3.2). Each level consists of a set of units or nodes; associated with each node at any moment in time is a momentary activation. The degree of activation corresponds to the strength of the hypothesis that the input contains the unit. A node whose activation level exceeds a threshold excites other nodes with which it is consistent (e.g. an initial T will excite the node for the word TAKE), and inhibit the other nodes with which it is not consistent.

Rumelhart and McClelland assume that inhibition and excitation sum algebraically, and the net effect of the input of a node is modulated by the prior activation of that node. In this way the letter nodes with the most active feature nodes receive the most net excitation. In their model, reading is treated as an interactive process in which contextual input is almost as important as direct evidence in the processing of stimulus material, so that as activation grows for one letter in a word it serves to facilitate the perception of the surrounding letters. According to the interactive-activation model, the pronounceability of a letter string does not determine how accurately the letters in the string will be perceived. What matters is how strongly the particular arrangement of letters produces partial word

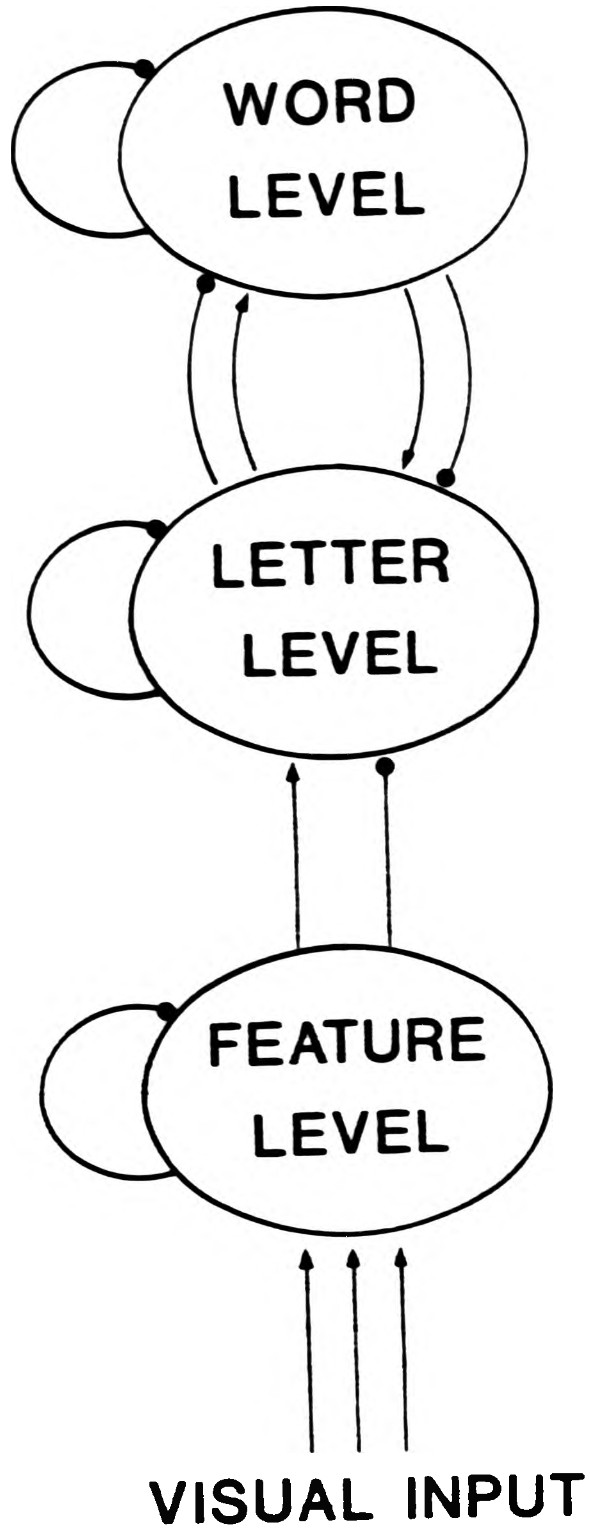


FIGURE 3.2

THE INTERACTIVE ACTIVATION MODEL
RUMELHART & McCLELLAND (1982)

The various levels of processing considered in the interactive activation model and their interconnections

——— EXCITATORY EFFECTS

——— INHIBITORY INTERACTIONS

activations that feed back and reinforce activations at the letter level. The model suggests that there would be unpronounceable nonwords that would produce as much facilitation of perception of the letters in them as comparable pronounceable nonwords would (e.g. considering the target letter P in the string SPCT, which occurs in three words having three letters in common with this display - SPAT, SPIT, SPOT. The nodes for these words should be activated by the letter string and produce feedback reinforcing the activation of the P node). Thus pronounceability and orthographic regularity per se should make little difference.

Rumelhart and McClelland consider their model quite consistent with some correlation of positional frequency and accuracy in words because positional frequency in words is strongly related to the number of words a letter might help activate. They also consider that their model is very much in keeping with Glushko's (1979) ideas of pronouncing words and nonwords. Glushko suggested that pronunciation is preceded by partial activation of the pronunciations of all the words in the neighbourhood of the target nonword, and followed by synthesis of a pronunciation from these partial activations. However their model would have to be extended to include stored information about the pronunciations, as well as the spellings, of words in English.

SEIDENBERG'S TIME-COURSE MODEL OF WORD RECOGNITION

Seidenberg (1985) used McClelland & Rumelhart's parallel interactive framework as the basis for his time-course model of word recognition. It was assumed that recognition is initiated with the extraction of visual information from the input, resulting in interactive processes of the type detailed by McClelland and Rumelhart (1981).

Seidenberg assumed that as orthographic units are identified they activate phonological representations that are also connected by excitatory or inhibitory links, depending on phonological similarity or dissimilarity. Under such an arrangement, access of phonological information lags behind access of orthographic information. Differences in the availability of these two types of information over time provide the basis for explaining the various word recognition phenomena.

The time-course model emphasises a single interactive process with differences in the availability of orthographic and phonological information over time. "Direct access" results when sufficient orthographic information is extracted from the input to permit recognition prior to the access of phonology, which is post lexical. "Mediated access" occurs if a word cannot be recognised prior to the activation of phonological information. Recognition then depends upon interactions among both orthographic and phonological units.

The time-course model predicted that only lower frequency, more slowly recognised words should be influenced by phonology. Seidenberg's (1985) experimental results showed that low frequency exception words yielded longer reaction times; and because they were decoded more slowly this allowed time for phonological information to be activated.

Henderson (1982) extended Glushko's ideas in a lexical pooling model which included classes of words that might be expected to have special properties in lexical pooling theory due to the nature of their orthographic neighbourhood. He described classes of HERMITS and HERETICS.

"A lexical hermit was described as a letter string with no close orthographic neighbourhood e.g. LYNX, since one cannot change a single letter in the word to transform it into a familiar new word. There are two possible stages of lexical pooling models at which hermits suffer a disadvantage. First is the activation stage. It takes longer for activation to spread to remote analogies (MANX, MINX, LYMPH etc. in the case of LYNX). Secondly is the assembly stage. If a word has no close neighbours, the analogy mechanism depends upon the assembly of pronunciations from many small segments. In the limiting case, these segments, shared between the word and its activated analogies, will be as analytical as those postulated in the GPC procedure. However, even in this boundary case the theories differ, since the analogy theory does not require that the frequency of various correspondences have been codified in advance and stored as a "rule".

Lexical heretics eg. COMB, SWORD and HAVE are maximally inconsistent. These are at odds with almost all their orthographic neighbours. Some, like HAVE, possess many close neighbours who are united in a hostile orthodoxy (CAVE, SAVE, RAVE, etc): in others, like COMB, neighbours are few and disparate (TOMB, BOMB, etc.)" Henderson, 1982, p.159.

Henderson described a mechanism involving facilitative interaction amongst members of an orthographic neighbourhood. Thus a pseudoword would have a perceptual advantage over a random string because it has orthographic overlap with real words, perhaps especially high frequency words. The two types of stimuli i.e. words and pseudowords differ merely in the fact that the region of orthographic space excited by their display contains, in the case of a real word, a representation of the word itself.

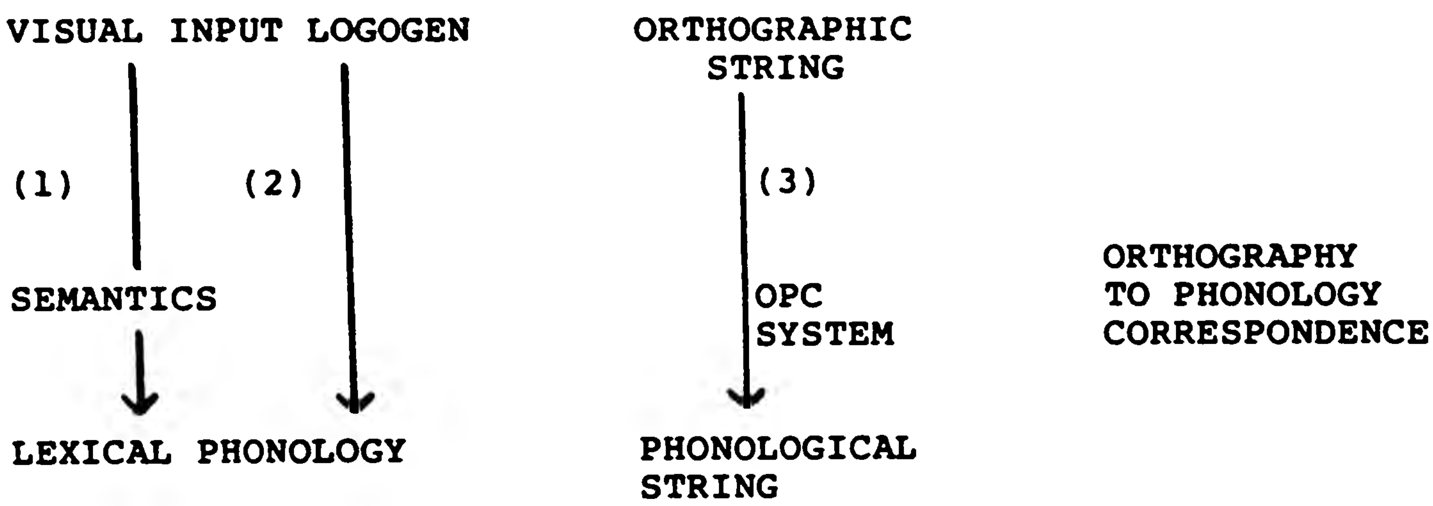
PATTERSON & MORTON'S MODIFIED STANDARD MODEL

To summarise the review of the literature so far:- one of the most comprehensive theories of word recognition is the dual route model (Morton, 1969; Morton & Patterson, 1980) in which there is a direct visual route to lexical access and a non-lexical GPC route for assembling the pronunciation of a word. The major alternative to a non-lexical route for assembling phonology is the theory that pronunciations are assigned by analogy with and by specific reference to known lexical items (Baron, 1977; Glushko, 1979; Marcel, 1980; Kay and Marcel, 1981; Henderson, 1982).

Deciding between these alternative approaches is far from straightforward. Patterson and Morton (1984) proposed a modified standard model which goes some way to incorporating the best aspects of both approaches. Their model has three routines for word pronunciation, two of which are lexical and one which is non-lexical (see fig.3.3).

FIGURE 3.3

MODIFIED STANDARD MODEL - PATTERSON & MORTON (1985)



The lexical routines both require that the visual input logogen system is used. In route 1, the output from this logogen system accesses the semantics of the item, from which the lexical phonology can be addressed. In route 2, the visual input logogen addresses the lexical phonology directly. Route 3 is the non-lexical route, its central procedure can be described as a set of mapping rules from orthographic strings to phonological strings, called the OPC system (Orthography-to-Phonology Correspondences).

Patterson and Morton's OPC system differs from Coltheart's (1978) GPC system in two ways. Firstly the OPC system deals with two different sizes of orthographic unit:

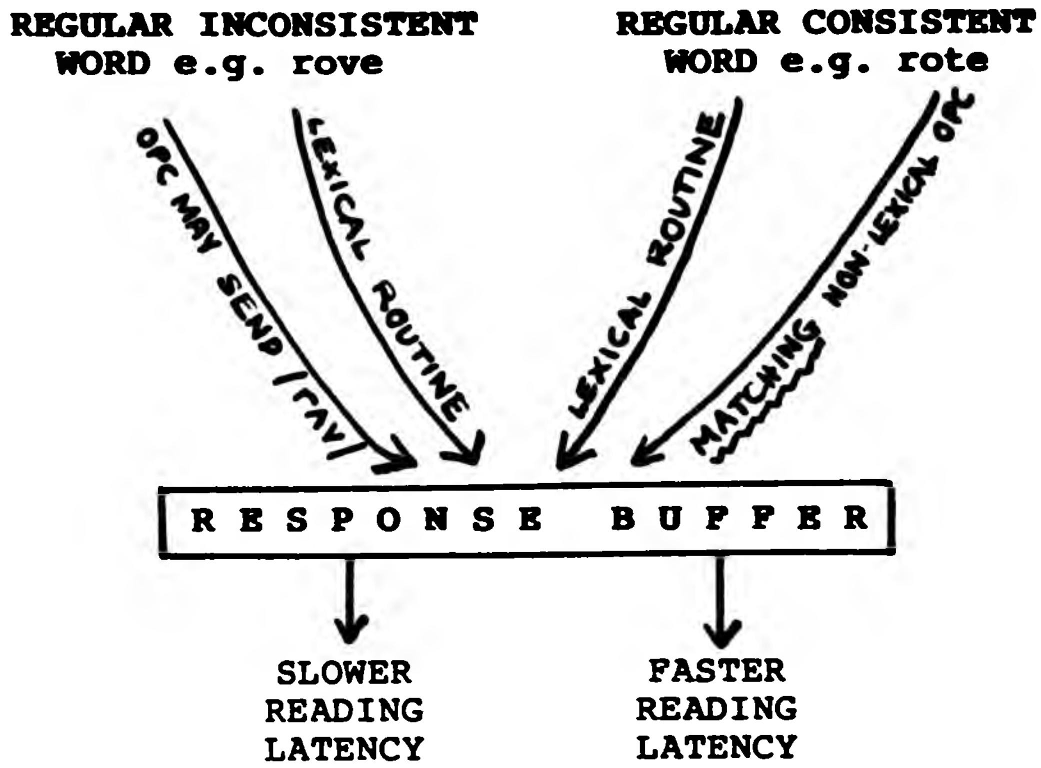
graphemes letters which correspond to single phonemes and bodies the vowel-plus-terminal consonant clusters. Secondly, mappings at the grapheme level are assumed to be simple one-to-one translations, but the mapping rules for bodies are more complex and sometimes need one-to-several translations.

The modified standard route model can cope better with experimental observations which have proved troublesome for the dual route model in the past.

Inconsistent nonwords e.g. POVE with an ambiguous body are sometimes read aloud with an irregular pronunciation - Glushko (1979) - this may be because the OPC mapping rules for this body include -OVE /ʌV/ which is irregular e.g. love as well as -OVE /OUV/ regular e.g. cove.

Glushko (1979) found that reading latencies may be significantly longer for inconsistent nonwords like HEAF than consistent ones like HEAN (a 22 millisecond difference). The modified standard model suggests that this can be explained in terms of a time penalty for inconsistent nonwords where there is a choice between alternative pronunciations.

The third result that proves difficult for the dual route model is that the reading latencies may be longer for regular inconsistent words e.g. ROVE than regular consistent words e.g. ROTE (Glushko, 1979; Andrews, 1982). The small latency disadvantage suffered can be explained in terms of interference from the OPC route.



The phonological codes produced by the two routines go to a response buffer - a system whose function is to transform the code into a form suitable for production. If nothing else is received then the transformed code is pronounced. If another code is received, it is compared with the first. If they agree, then the transformation process continues. If they disagree, then there will be a time penalty due to the interruption and a decision between the alternatives will require a lexical check.

Kay and Marcel (1981) found that an inconsistent pseudoword could be significantly shifted towards irregularity - YEAD would be pronounced irregularly (/jed/) if there was prior presentation of the appropriate irregular word HEAD.

ORDINARY CONDITION		**IRREGULAR BIAS CONDITION**	
yead		yead	
↙	↘	↙	↘
/ji:d/	/jed/	/ji:d/	/jed/
REGULAR PRONUNCIATION	IRREGULAR PRONUNCIATION	REGULAR PRONUNCIATION	IRREGULAR PRONUNCIATION
body r = 0.50		favoured selection irregular	

body-r= average proportion of unbiased regular pronunciations

In the absence of a biasing word, Patterson and Morton propose that the OPC body sub-system selects at random between /ji:d/ and /jed/ as pronunciations for YEAD, producing a body-r = .50. If HEAD has just been pronounced, this random selection might abruptly shift to favour selection of the /jed/ alternative, and then gradually drift back to its usual lack of preference.

Finally, Rosson (1983) showed that the pronunciation assigned to an inconsistent nonword can be shifted towards irregularity by prior presentation of a word semantically related to a word which would have produced the Kay and Marcel biasing effect.

STIMULUS	sofa (semantically related to COUCH-regular)	feel (semantically related to TOUCH-irregular)
STIMULUS	louch	louch
	↓	↓
	/lautʃ/	/lʌtʃ/
	MORE REGULAR PRONUNCIATION	FEWER REGULAR PRONUNCIATION

Patterson and Morton's model deals with Rosson's finding in the following way. The logogen for TOUCH is affected by the stimulus LOUCH, together with feedback from the cognitive system following prior presentation of FEEL, resulting in a shift towards an irregular response.

Patterson and Morton describe five types of bodies requiring differentiation within the OPC system.

CONSISTENT BODIES like -aze where the body sub-system offers a single mapping.

CONSENSUS/HERETIC BODIES like -int where the overwhelming consensus is /Int/ (e.g. mint, lint, hint etc.) with the single heretic pint.

Bodies of the GANG type are the opposite from consensus and consistent types.

A GANG WITH A HERO (e.g. -OOK spook is the hero of the irregular gang: book, cook, look, hook etc).

A GANG WITHOUT A HERO (e.g. -old where every word belong to the gang: cold, hold, bold, fold). This matches the consistent type, but the gang pronunciation is irregular rather than regular with respect to GPC rules.

Patterson and Morton's contention is that representations in the body sub-system are established on the basis of experience with words, so that only the majority mapping is represented.

AMBIGUOUS bodies are the only type for which it seems clear that at least two OPC mappings must be represented. Thus -EAF will have both /i:f/ and /ef/. Parkin (1983) showed that reading latencies were significantly slower for ambiguous nonwords.

Thus in Patterson and Morton's present model, the GPC routine will provide 'regular' pronunciations for all letter strings; the body routine will provide a) regular pronunciations for both consistent and consensus/heretic bodies, b) irregular pronunciations for gang bodies and c) an equivalent number of each alternative pronunciation for ambiguous bodies.

The development of a body sub-system is directly investigated in Experiments 7 and 8, reported in Chapter 7.

Recent work on acquired dyslexia has provided a useful framework within which Patterson and Morton's body types could operate. Shallice, Warrington and McCarthy (1983) described the reading performance of a surface dyslexic patient (HTR) who suffered from impairment of the non-lexical GPC route. He showed an effect of 'degree' of regularity in terms of accuracy of reading response which Shallice et al described as being due to 'typicality of divergence', a concept described at greater length later on in this chapter.

The pattern of impairments of the acquired reading disorder, surface dyslexia, first drawn together by Marshall and Newcombe (1973), has been attributed to defective lexical processing, with oral reading relying on non-lexical GPC translations from print to sound. However the pattern of deficits shown by the patients investigated by Marshall and

Newcombe deviated in several ways from what one might expect from a simple two-process account of impaired lexical processing and reliance on a GPC rule-based system.

Firstly the bulk of error responses could not be attributed to regularisations of irregular words. Both regular and irregular words were often mispronounced. The most common errors were partial failures of grapheme-phoneme conversions e.g. RECENT "rikunt".

Secondly, as Marcel (1980) observed, such an account failed to explain the large number of error responses that corresponded to real words (lexicalisation) e.g. INCENSE read as "increase". According to Marcel, the tendency towards word errors indicates that the GPC route is not independent of the lexicon. He went on to reject the two-process view of reading and suggested that lexical phonological information about words (stored in segmentable rather than whole word form) could be used in constructing pronunciations of nonword strings in normal reading. Shallice, Warrington and McCarthy (1983) argued that lexicalisation of responses reflected guessing strategies used to compensate for a faulty phonological process. They considered such strategies were not part of the reading process.

SHALLICE, WARRINGTON & McCARTHY'S VISUAL WORD-FORM SYSTEM

Shallice and Warrington (1980), Shallice (1981) and Shallice, Warrington and McCarthy (1983) put forward a visual word-form system or lexical route. This system assumes that lexically stored spelling-to-sound correspondences are not limited to graphemes, but exist for various sizes of orthographic units - graphemes, consonant clusters, sub-syllabic units (i.e. (C)CV and VC(C) units), syllables and morphemes. In this multiple-levels model, all of these units are separately represented in the visual word-form system.

e.g. orthographic components of the word HEAD are

h- + -ea + -d

he- + -ad

hea-

-ead

head

The system is capable of operating on orthographic units of a number of different sizes and transmitting information about them to corresponding units in the phonological system.

Thus, it is implicit in the multiple-levels model that measures of orthography such as initial bigram frequency, initial bigram versatility, terminal bigram frequency, terminal bigram versatility, number of neighbours and body measures are important in the reading process. The importance of some of these measures of orthography will be investigated in Experiment 1.

In addition Shallice et al assume that information about different units of the same size cannot be processed simultaneously. Thus the spelling-to-sound correspondences for DR and NK in DRINK are achieved relatively serially. Thus the larger the size of the word-form unit on which the translation process is based, the faster is information about the pronunciation of a given letter string obtained.

Shallice et al explored further the dimension of spelling-sound regularity and how it affected reading. According to the GPC approach to spelling-sound translation (e.g. Coltheart, 1978) any word containing a GPC that is not the most frequent is classified as an exception and differences in the number of correct pronunciations of mildly and very irregular words are not predicted. Shallice et al have further classified irregular words into typically divergent and atypically divergent words, (see Fig.3.5).

The word HEAD is classified as irregular according to standard GPC theory (because the regular phonemic correspondence of the vowel digraph ea is /i/; ea /ɛ/ is divergent). According to Shallice and Warrington's account whereby the orthographic unit -EAD is taken into account, the most frequent correspondence is /ɛd/, since the majority of -EAD words thyme with HEAD. Thus mildly irregular or divergent words like HEAD were classified by Shallice et al as "typically divergent". The term "atypically divergent" was applied to words such as BOWL which contain less frequent and divergent correspondences. They predicted that typically divergent words would be easier to read than atypically divergent words.

They were therefore able to explain why their patient HTR was able to read many mildly irregular words, although their non lexical GPC system was impaired.

FIGURE 3.4

SHALLICE AND WARRINGTON'S CONCEPT OF TYPICALITY OF DIVERGENCE (1983)

TYPICALLY DIVERGENT WORDS e.g. head

ORTHOGRAPHIC UNIT -EAD

	PRONUNCIATION		
dead	/ɛd/	(irregular)	
head	/ɛd/		
lead	/ɛd/	/id/	
read	/ɛd/	/id/	
dread	/ɛd/		
bread	/ɛd/		
bead	/id/	(regular)	
mead	/id/		
knead	/id/		

ATYPICALLY DIVERGENT WORDS e.g. bowl

ORTHOGRAPHIC UNIT -OWL

	PRONOUNCIATION	
bowl	/ou/	(irregular)
cowl	/aʊl/	(regular)
fowl	/aʊl/	
howl	/aʊl/	
jowl	/aʊl/	

N.B. It is considered that Shallice and Warrington's typically divergent example would be more appropriately catagorised midway in the typicality of divergence continuum - that is, reclassified as an ambiguous body type.

Shallice et al (1983) made the further assumption that the spelling-to-sound translation process was subject to progressive impairment, the larger size orthographic units being affected first. This was an ad hoc hypothesis based on observation of their patient.

Their multiple-levels model is similar to that of Glushko (1979) in that both theories assume that spelling-to-sound correspondences exist not only for graphemes and morphemes, but also for units of intermediate size, particularly for sub-syllabic units. However Glushko's position is different in that he also allows for reading by analogy. Shallice et al did not consider that this was needed in order to explain the results that Baron and Glushko obtained i.e. the longer naming latencies of irregular over regular words (Baron and Strawson, 1976) and of inconsistent words over consistent words (Glushko, 1979).

According to Shallice et al, at least two possible sets of sub-syllabic units in the phonological system could be activated. Those sub-syllabic units which lead to the correct response will be activated by using morphemic orthographic-to-sound correspondences, and those sub-syllabic units which lead to regularisation will be activated by using sub-syllabic correspondences. On the theory, the former will on average be activated more quickly. However the similarity in naming latency between exception words and regular nonwords shown by Glushko (1979) suggested to Shallice et al that the activation of the two competing sets of phonological sub-syllabic units overlap in time and so interference can occur. Such interference would not occur for regular words, because use of the different types of correspondences produces the same results. What both models leave open is how segmentation of the letter strings is achieved and how the translation processes operating on a number of different levels is achieved.

Their model fits neatly with the evidence reported by Santa, Santa and Smith (1977) who found that a sub-syllabic unit such as AST in BLAST was detected faster than a non-sub-syllabic unit LAS. Seymour and May (1981) found that manipulations of length and format distortion had similar effects on words and nonwords - implicating a common visual orthographic processing stage.

To summarise the two approaches - alternative correspondences are represented at all levels of the phonological system. The choice between the alternatives is based on the strength of each correspondence i.e. the strongest is equivalent to the most frequent pronunciation of a particular letter group (in terms of number of word tokens). Each correspondence is weighted - /ɛd/ has a greater weighting than /id/, so the reader would choose /ɛd/, which explains the good performance on typically divergent words.

Alternatively the GPC method of spelling-sound translation accounts for good performance on typically divergent words by assuming that there are probabilistic weightings which are associated with alternative phonemic correspondences of individual graphemes, e.g. the correspondence o → /ʌ/ only occurs in a few words - LOVE, DONE, and is assigned a low weighting compared with the more usual pronunciation /aʊ/ in HOWL. However, the correspondence ea /ɛ/, though divergent, occurs in many words - HEAD, DEATH, SWEAT and is assigned a higher weighting. So in comparing a typically divergent word HEAD with an atypically divergent word LOVE, HEAD would be easier to read because its correspondence has been assigned a high probabilistic weighting.

Shallice and McCarthy (1985) provide more detail on their Multiple Levels Approach. They consider that it differs from the standard position on phonological reading in two ways. Firstly, it is claimed that different levels of unit are used, particularly subsyllabic units and also syllables. The second difference is that the multiple-levels position holds that different levels of the process operate in an integrated way in parallel.

To show that the multiple-levels theory could work, one needs to show for each level how a set of these visual word-form units could be selected that do not overlap and yet "cover" the word completely. The question remains as to how the relation between different levels would operate on this system. If compatible units on different levels are made mutually facilitatory and possibly also incompatible ones on different levels mutually inhibitory, then because of their simple hierarchical relation it would seem possible for all levels of unit to be activated together and for the different levels of parsing to reinforce each other in the overall process.

Shallice and McCarthy summarised the two important characteristics of this procedure. First, the parsing constraints occur between all levels of orthographic units with the exception of the morphemic level. Therefore, in the orthographic "part" of the spelling-to-sound translation process parsings can take place on all orthographic levels at the same time, and incompatible parsings on different levels will not occur. Secondly, the units that are detected are ones with immediate phonological correspondences. Thus the orthographic parsing minimises any supplementary reading specific spelling-to-sound parsing process. Individual orthographic units, when activated, can in turn activate a corresponding phonological unit.

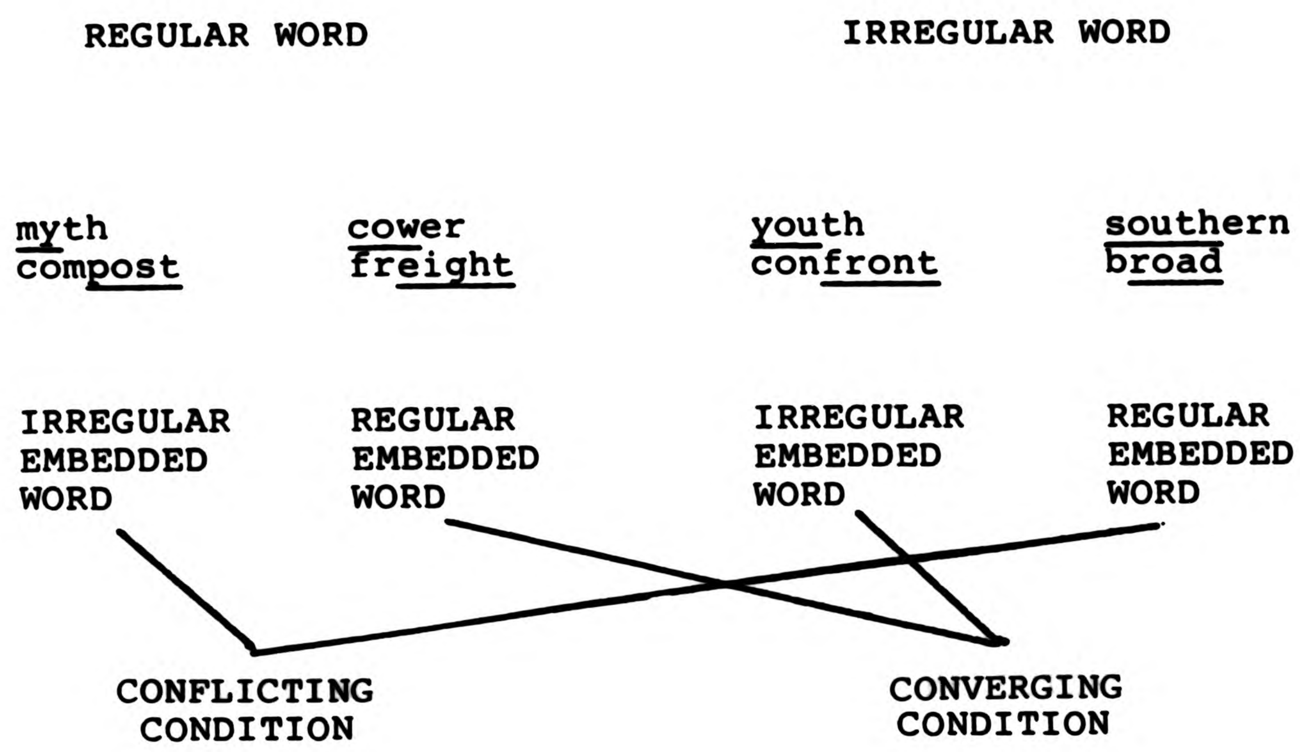
Kay and Lesser (1985) compared the reading performance of a neurological patient (PT) to that of the surface dyslexic patient (HTR) reported by Shallice, Warrington and McCarthy (1983). Both patients showed an effect of 'degree' of regularity in terms of accuracy of reading response. PT did not, however, show an effect of 'typicality of divergence' or the existence of the sub-syllable as a reading unit. He appeared to rely on grapheme-phoneme correspondences that were assigned in a probabilistic way.

PT was given three word lists for regularity - the Coltheart word list (1979), the Levels of Regularity List (Shallice, Warrington, McCarthy 1983) and the Vowel Pronunciation List (Shallice et al 1983). A regularity effect was shown both in reading aloud and in reading speed. Of the irregular word set, errors of regularisation were the most common type of mispronunciation. The most frequent type of error in the regular word set (and second most frequent in the irregular word set) was the stress error. Like errors of regularisation, stress errors would be expected in non-lexical spelling to sound procedures because stress assignment is assumed to fall under lexical control. In contrast to Shallice et al's patient HTR, almost all of TP's mispronunciations were errors in individual GPC transformations and several errors were partial failures of GPC conversion, indicating the independent operation of a relatively unimpaired non-lexical GPC process.

Kay and Lesser found no trend for typically divergent words to be read better than atypically divergent words. In fact, in the Vowel Pronunciation List, the trend was in the opposite direction to that hypothesised by Shallice et al. PT tended to assign incorrectly regular and less frequent correspondences at the VC(C) level. Thus it appeared that for Kay and Lesser's patient, typicality of divergence at the sub-syllabic level was not responsible for the effect of degree of regularity in single word reading.

In a final effort to show how influential higher order spelling-sound correspondences could be, Kay and Lesser (1985) selected thirty regular and thirty irregular words. All of the words contained an embedded orthographic segment corresponding to an existing word (see overleaf).

It embedded orthographic segments corresponding to real words exert an influence in determining pronunciation, then Kay and Lesser predicted that for regular words, those segments corresponding to irregular words would interfere in the production of the whole word response, compared with words in which the pronunciation of the embedded segment and the whole word converge. They found that both regular and


REGULAR WORD
IRREGULAR WORD
myth
compost
cower
freight
youth
confront
southern
broad
IRREGULAR EMBEDDED WORD
REGULAR EMBEDDED WORD
IRREGULAR EMBEDDED WORD
REGULAR EMBEDDED WORD
CONFLICTING CONDITION
CONVERGING CONDITION

irregular words containing orthographic segments with converging pronunciations were read more easily than those with conflicting pronunciations of embedded segments.

Therefore, it seems that the theory of spelling-sound translation based only on the simple grapheme-phoneme level does not fully account for the evidence brought together by Kay and Lesser.

Although PT showed a superior performance on regular words, he was also more successful in reading irregular words composed of divergent GPC correspondences than those composed of exceptional GPC correspondences. Thus he was not simply influenced by whether a word was regular or irregular, but also by degree of regularity of a word. However he failed to show more successful pronunciation of typically than atypically divergent words, as proposed by Shallice et al (1983). Shallice et al hypothesised that the spelling-to-sound translation process was subject to progressive impairment, the larger size orthographic units being affected first. Kay and Lesser therefore considered that the differences in reading performance between PT and HTR were because they were at different stages of impairment - HTR's deficit corresponding with an intermediate stage of dissolution in which GPC and sub-syllable levels are relatively intact.

Thus there are three theoretical approaches which address the question of how readers perform on words differing in degree of regularity -

a) The dual route model with its non-lexical GPC route (Coltheart 1978)

b) Patterson and Morton's modified standard model with its OPC system. This can be incorporated with Shallice and Warrington's multiple lelvels model where sub-syllabic units are used by skilled readers i.e. body types can be placed along the typically/atypically divergent continuum.

c) Analogy theory - which rejects the idea of a GPC route altogether, where the number of consistent or inconsistent orthographic neighbours is of importance.

CONCLUSION

The Standard Dual Route Model (e.g. Coltheart, 1982; Morton & Patterson, 1980), the Multiple-Levels position (Shallice, Warrington & McCarthy, 1983) and the Lexical Analogy Position (Henderson, 1982; Marcel, 1980) are not sharply conceptually distinct. Thus by elaborating the non-lexical process the first theory becomes quite similar to the second. In addition the second and third theories can make many similar predictions. However they differ on whether lexical activation of similar words is crucially involved in phonological reading. Also the Multiple Levels position, but not the Lexical Analogy position, predicts that lexical and non-lexical processing can be dissociated.

The three models, although very similar, can be criticised for different reasons.

In their version of the lexical analogy theory Kay and Marcel (1981) say: "A printed letter string is segmented in all possible ways...Each segment automatically accesses matching segments in the orthographic lexical input addresses of all words which contain those segments in equivalent positions". This means that for any word the pronunciation of all possible segments must be stored. For a seven-letter word there are 28 of these (seven of length one, six of length two, and so on) and even if ones that disrupt functionl spelling units are excluded, this approach to phonological reading requires that far more information be stored and accessed than does simple morphemic reading. In addition how the mass of information accessed is recombined is unclear.

Henderson's (1982) development of Glushko's (1979) lexical pooling model avoids the complications of Kay and Marcel's (1981) approach. Henderson assumes that each letter string activates all visually similar words by an amount proportional to the degree of relatedness and that each entry in a visual lexicon in twin activates its phonological counterpart proportionally. The "most heavily weighted" candidate is chosen. Shallice and McCarthy criticises this lexical analogy approach by considering the attempt to pronounce EHELF. H almost never occurs in the second position in English words other than graphemes CH, PH, SH and TH. How are these pronunciations over-ridden when one produces say /shelf/, or alternatively, why is SHELF not a visually similar word?

The main problem with Shallice & McCarthy's multiple levels approach is that there can be multiple representations of the same element in a letter string. Secondly, the details as to how the morphemic levels interact with the others still have to be worked out. However, this approach is a useful one as it emphasises the importance of orthographic units in the reading process.

It appears from consideration of alternative theoretical approaches, (Humphreys & Evett, 1985; Carr & Pollatsek, 1985; Patterson & V. Coltheart, 1986) that what were once considered genuinely contrasting opposing views are becoming difficult to distinguish.

3.3 DEVELOPMENTAL MODELS OF READING

Marsh, Friedman, Welch and Desberg (1981) have described a sequence of stages which appear to be part of the developmental procedure for most normal readers. It is not a very detailed theory and as yet there has been little evidence that children do go through the developmental stages that they describe. Since the current study will be examining children of different reading ages, it will be of interest to examine the theory in the light of developmental data collected. Marsh et al proposed a developmental trend from visual to visual and phonological reading via four stages.

A basic assumption of their theory was that any cognitive process goes through a number of stages which change qualitatively with development. They emphasised the development of strategies in reading unknown words because they considered it a very challenging aspect of reading acquisition and the one which was most likely to engage high-level cognitive processing strategies. It was therefore one in which they expected to find qualitative developmental differences, as the ability to read aloud unknown words gives an indication of childrens' knowledge of orthographic structure and its relation to the language system they already possess.

PROPOSED STAGES IN THE DEVELOPMENT OF READING

Stage One - Linguistic Guessing

In learning to read, the young child in the first stage approaches the task with a strategy of simple rote association. The rote association is between an unsynthesised visual stimulus and an unanalysed oral response. "The child typically centres on the first letter and associates that with the oral response". Thus one might expect to see that measures of orthography such as initial letter frequency and versatility affecting ease of reading.

They cannot perform phonemic segmentation tasks. Their natural strategy is congruent with the "whole word" approach to teaching reading. When faced with an unknown word in a sentence context the child responds with a syntactically and semantically appropriate word.

Stage Two - Discrimination Net Guessing

In the second stage the child typically responds to an unknown isolated word on the basis of its shared graphemic features with a known word. The number of graphemic features a young child can process is limited initially to the first letter, and it is only later that additional features such as word length, final letter etc are added. Thus one might expect to see measures of initial and terminal letter frequency and versitility affecting the ease of reading. i.e. words which have high initial and terminal letter frequencies and versatilities will correlate highly with how easy they are to read.

The child at this stage appears to be operating according to a "discrimination net" mechanism in which graphemic cues are processed only to the extent necessary to discriminate one printed word from another. Such a "discrimination net" strategy is typically found in learning tasks in which children are required to learn novel material by rote. At Stage Two the child almost always limits his response to previously learned printed words.

Stage Three - Sequential Decoding

Stage Three is characterised by the use of combinational rules which allow the reader to "decode" new words. Two factors are seen to be involved in the child's switching from the earlier strategies to strategies of Stage Three:- firstly, the increase in the number of items in the print vocabulary results in an increase in the memory load and a rote learning and partial graphemic cues are no longer appropriate; secondly, there is an increase in cognitive processing capacity as the child moves into the stage of concrete operations. Older children (approximately 7 years old) may be able to pay attention to a word's sound as well as its meaning and to process the order of a series of letters and to coordinate this series with a series of sounds.

At Stage Three, the child treats the alphabetic principle as a simple invariant code where each letter represents a single sound. Children at Stage Three can assemble a correct pronunciation for new words if they are regular, and invariant word patterns such as the CVC pattern where the vowel has a short pronunciation. They are unable to deal with words which require conditional and other higher order rules e.g. silent e. Thus regularity effects may be observed in Stage Three, but only for a limited set of words e.g., PINT, SWORD.

Stage Four - Hierarchical Decoding

It is not until Stage Four that the child has the ability to deal with the conditional and higher order rules. An analogy strategy also appears to be available by this stage. Experimental evidence for the developmental trend from visual to visual and phonological reading was provided by Bradley and Bryant (1979), Frith (1981) and Campbell (1983) who suggested that simple GPC rules may not reflect a basic skill, but a high level of literacy and abstraction.

ANALOGY STRATEGIES

A number of investigators have pointed out that an unknown word may be read by analogy to a known word rather than by use of GPC/combinational rules. Barron (1977) provided evidence for the use of the analogy strategy by adults, but the evidence was based primarily on the subject's post hoc report on strategies used. Baker & Smith (1976) used a conflict technique where reading by rule and reading by analogy would produce different pronunciations and found evidence for both rule and analogy processes in adults' placement of stress on polysyllabic words. The cognitive-developmental approach presented by Marsh et al (1981) would predict that older children and adults would have a number of different strategies available in reading and the use of various strategies would depend on the specific task factors. They suggest that as reading skill increases, there is a shift from word recognition based on access to phonemic form to word recognition based on access to meaning.

Marsh et al (1981) conducted a study investigating the use of analogy strategies in skilled and less skilled readers, subjects were twenty 7 year olds, sixty 10 year olds and 40 college students. The reading materials consisted of two short paragraphs adapted from the Gray Oral Reading Test with a number of nonwords inserted in noun positions. Some of these nonwords assessed the subject's use of analogy strategy:-

e.g. faugh < faw - by decoding strategy
pronunciation by analogy to 'laugh'

They found that 10 year olds used analogy as did adults. In contrast to a previous study (Marsh et al 1977) the subjects did not produce a preponderance of analogy responses. There was no significant increase in use of analogy between 10 year olds and adults. The percentage of analogy responses of 10 year olds was similar to the percentage in the previous paper, but the adults in this study produced only thirty percent analogy responses as compared to over seventy percent in the previous study. The pattern of results suggests that the analogy strategy is an option strategy for adults - its use strongly determined by task factors. Also in the previous study, all the nonwords had analogies while in the present study only some of the words had obvious analogies.

One of the predictions of the cognitive developmental approach is that there will be a qualitative shift in the stimulus factors which are most important at each stage. In Stage Two the most important factor will be the visual familiarity of the words the child reads. In Stage Three and Four the important factor will be the phonemic regularity of the word patterns. Evidence for such a qualitative shift has been reported by Pick (1978) who found that young children made word similarity judgements in a matching to sample task based on the words' visual similarity, while older children chose on the basis of phonemic similarity. Similarly Barron

(1980) reported that poor readers' judgements in a lexical decision task were primarily affected by visual factors while good readers' judgements were affected primarily by phonemic factors.

Marsh et al (1981) designed an experiment to investigate the relative role of visual familiarity and phonemic regularity in reading children of different ages and ability levels. An additional question of interest was the availability of the analogy strategy in Stage Three. According to Sternberg (1977), 7 year olds are capable of analogical reasoning and Baron (1979) reported that 7 year olds were capable of producing analogy responses after having been given examples. Marsh et al's subjects were twenty four 7 year olds, twenty one 9 year olds, and twenty four reading disabled children. They were asked to read two twenty word lists. The first list contained twenty high frequency real words, one half of which were regular, the other half irregular. The second list contained a transformation of each of the words in the first list into nonsense words. Results showed that phonemic regularity was the major factor in the performance of normal readers. The reading disabled subjects showed better performance on visually familiar real words. Analogy responses were shown by all groups. Thus Marsh et al appeared to show that there was a developmental trend in reading from visual to visual and phonological reading.

Frith (1985) has adapted Marsh, Friedman, Welch and Desberg's (1981) four stage developmental model of reading acquisition, collapsing Stages 1 and 2 and making explicit the relationship between reading development and spelling development, so as to provide links to current models of skilled reading. She divided the development of reading into three phases identified with three strategies - logographic, alphabetic and orthographic.

Logographic skills refer to the instant recognition of familiar words. Salient graphic features may act as important cues in this process. Letter order and phonological factors are not relevent i.e. the child pronounces the word after he or she has recognised it. If the child does not know the word, he or she will refuse to respond or the child may use contextual cues when prepared to guess.

Alphabetic skills refer to knowledge and use of individual phonemes, graphemes and their correspondences. It is an analytic skill involving a systematic approach - decoding grapheme by grapheme. Letter order and phonological factors play a crucial role. This strategy enables the reader to pronounce novel and nonsense words, though not necessarily correctly.

Orthographic skills refer to the instant analysis of words into orthographic units without phonological conversion. The orthographic units ideally coincide with morphemes. They are internally represented as abstract lettler-by-letter strings. These units make up a limited set - in loose analogy to a syllabary - can be used to create by recombination a large number of words. The orthographic strategy is different from the logographic one, in that it is analytic in a systematic way and is non-visual. It is distinguished from the alphabetic one by operating in bigger units and by being non-phonological. The ability to deal with sub-word segments ties in theoretically with the levels of orthographic structure outlined by Shallice, Warrington and McCarthy (1983) in their visual word-form system for skilled readers, where there are various sizes of orthographic units. Thus, the three stages, as defined by Frith (1985), can readily be related to components in current models of skilled reading (Morton and Patterson, 1980; Shallice, Warrington and McCarthy, 1983). For example, word-form analysers of skilled readers might be derived from early logographic skills; grapheme-to-phoneme skills would need to have been constructed out of alphabetic knowledge. In considering how a child might move from one step to the next, and what provides the impetus for the adoption of the various strategies, Frith further elaborated her three-phase model to a six-step model. This hypothesis results in a noticeable modification of the acquisition model, as shown in Fig 3.5

Frith argues that, although both reading and spelling depend on the three stages, the progression is asynchronous. The division into steps allows a differentiation in terms of level of skill in a particular strategy, as symbolised by number subscripts. Level 1 would imply that the skill is present in very basic form only; Level 2, that it is more advanced, and so on. Thus she hypothesised that only when logographic skill had reached Level 2 in reading was it ready to be adopted for writing.

FIGURE 3.5: THE SIX-STEP MODEL OF SKILLS IN READING AND WRITING ACQUISITION - FRITH (1985)

Level	*Reading*	*Writing*
1a	*logographic*$_1$	(symbolic)
1b	logographic$_2$	logographic$_2$
2a	logographic$_3$	*alphabetic*$_1$
2b	alphabetic$_2$	alphabetic$_2$
3a	*orthographic*$_1$	alphabetic$_3$
3b	orthographic$_2$	orthographic$_2$

The alphabetic strategy is first adopted for writing, whereas the logographic strategy continues to be used for reading, perhaps at an even more advanced Level 3. Only when the alphabetic strategy reaches Level 2 will it be adopted for reading. The rationale for the antecedence of writing here is provided by the idea that the alphabet is tailor-made for writing rather than for reading (Frith and Frith, 1980). Thus it is easy to learn to write using the relatively small set of letters. At the same time the limited number of symbols creates ambiguities when translation into sound is required, and so, for reading the logographic strategy is still required.

Phase 3 shows the orthographic strategy appearing in the reading process, in step 3a. Orthographic knowledge at Level 1 is presumed to be weak - sufficient to be used in recognising words, but not to be used in guiding the writing of words. Level 2 would imply that orthographic representations are now precise enough to be useful for spelling. It is plausible to assume that they would then be "transferred" to the spelling output system. The highly skilled reader/speller requires internal representations that are exact in terms of letter- by-letter detail.

In summary, Frith's theory states that at each phase there is a first step involving a divergence between strategies used for reading and writing, then a step involving convergence. Developmental progress is envisaged as an alternating shift of balance between reading and writing. Reading is the pacemaker for the logographic strategy, writing for the alphabetic strategy, and reading again for the orthographic strategy.

Thus Frith's developmental model has attempted to explain how the various strategies that are mastered by the skilled reader come into being. Both Marsh et al (1981) and Frith's (1985) developmental model can account for differences in speed of acquisition, which range from sudden improvements to very slow changes and apparent plateaux. This is possible because they view the acquisition of skills as a sequence of steps. A step forward in the sequence is identified with the adoption of a new strategy; sudden improvement is allowed for by the more from one step to the next higher. A fall-off in performance may be due to the application of a new unpractised strategy in place of a well-practised old one, causing an initial drop in performance.

The developmental approach also implies special awareness of the interaction of constitutional and environmental factors. This is not a notable feature of structural approaches. It may be the case that environmental factors are less important when one is dealing with highly skilled readers. At that stage it is assumed that automatic processes are operating, but this is not the case with developing readers, where the influence of school and home can plainly be seen in what, why and how they read (Francis 1982). A difficulty in Marsh et al (1981) and Frith's (1985) scheme, pointed out by Seymour and MacGregor (1984), is that they are based on an assumption of homogeneity within the population of beginning readers, in terms of individual characteristics and method of teaching. Seymour and Macregor (1984) in their report of the reading performance of four developmental dyslexics, produced a much more specific and testable model in terms of information processing and components, than the general developmental processes described by Marsh et al (1981) and Frith (1985).

Seymour and MacGregor's model consists of a phonological processor and a semantic processor, which both come before reading and spelling acquisition, and a visual (graphemic) processor and orthographic output lexicon which must be constructed during acquisition. They considered that the visual (graphemic) processor is the principle new element which must be set up in order for the child to become a reader. It consists of three levels - the letter identity level, selection and transfer level and the recognition level. The recognition level subdivides into two lexicons - a logographic lexicon and an orthographic lexicon. The logographic lexicon deals with visual feature information and discriminates among words in a known vocabulary on the basis of salient features, and accesses phonology through semantics. The logographic lexicon might preferentially be used in fast silent reading. The orthographic lexicon requires pre-processed input - a conversion of visual features to abstract letter identities, followed by a search for an orthographically defined vowel cluster and segregation of the vowel from initial and terminal consonant groups. The processed array is then passed to the orthographic lexicon for recognition. It is assumed that the transfer of information from the letter identity level to the orthographic lexicon may involve the whole array, subsets of letters or single graphemes. Seymour and MacGregor suggest that the orthographic lexicon links directly with the phonological processor, so in principle any word or nonword which is located in the lexical space can be pronounced. A point accessed in orthographic space could be directly related to the corresponding point in phonological space, and the pronunciation associated with that point generated. Evidence of lexical influences on nonword reading (Glushko, 1979; Kay & Marcel, 1981) suggests that this mapping may not be strictly one-to-one.

Seymour and MacGregor's dual lexicon model is within the framework of the dual access model for adults but differs from other proposals in that firstly, two lexical recognition systems are proposed; secondly, lexical and non-lexical routes to phonology are combined within the orthographic lexicon (cf Shallice, Warrington & MacCarthy, 1983); and thirdly, the orthographic input lexicon is seen as a crtical basis for spelling development. The dual lexicon model is seen as providing "a general framework to help translation of experimental data into a description of the information processing capabilities underlying the reading performance of individual dyslexic subjects" (Seymour and MacGregor, 1984, p.76). This study provides a detailed information processing analysis using a more individual, case oriented approach rather than trying to compare a "dyslexic group" with a "control group". It is of interest because it provides a more detailed account of the developmental processes described by Marsh et al (1981) and Frith (1985), and it also bridges research on beginning reading and cognitive models of adult reading.

Seymour and Elder (1986) studied the reading development of the individual members of a class of new entrants to primary school (aged 4½-5½ years) over a period of a year. The teaching they received emphasised the formation of a "sight vocabulary". Instruction in letter-sound associations was restricted to spelling and writing. The children appeared to be "reading without phonology". These results were discussed in terms of the formation of a rudimentary word recognition system - the logographic lexicon (Seymour & MacGregor, 1984). Absence of phonological mediation was shown by the very low frequencies of correct readings of unfamiliar words, low overt demonstrations of sounding individual letters i.e. lip movements preceding the response, low frequencies of nonword responses and regularisations (WAS--WASS).

Their results fit easily within the framework of Frith's (1985) developmental model and provide support for the logographic/alphabetic distinction and for the proposal of a spelling-reading asynchrony. The children in the sample were logographic readers who developed the basis of an input logogen during the year in which they were studied, and their reading process was characteristic of word recognition by a logographic lexicon.

Seymour and MacGregor suggest that the logographic lexicon appears to recognise words by a process of feature discrimination. The features may include length, shapes of salient letters, and salient feature position. One basic assumption in their design of a methodology for investigating children's reading was "that the objective of establishing a relationship between beginning reading and adult reading requires that COGNITIVE/PSYCHOLINGUISTIC procedures capable of assessing properties of an information processing system should be adapted for use with small children." (Seymour and Elder, p.3, 1986.) Since the outcome of the developmental sequence in any developmental model is the skilled reader, structural models of skilled reading are therefore helpful when considering developmental models, and they suggest fruitful lines of investigation with children.

3.4 SUMMARY

The experimental work which follows tackles questions raised both by theories of reading based on skilled readers and development models of reading. The brief description of Glushko's (1979) paper shows that what remains implicit in his activation-synthesis model is the importance of orthographic units and the orthographic neighbourhood, which obviously underlie his theory. Rumelhart and McClelland's (1982) extension of the interactive-activation model stresses the importance of the orthographic variable - positional letter frequency. Seidenberg's (1985) Time Course model of word recognition also emphasises the importance of orthographic information. Shallice, Warrington and McCarthy's Visual Word-Form System (1983), described as a multiple levels model, also appears to be of direct relevance to Experiment 1, which investigates different measures and sizes of orthographic units.

The dual route model has been very useful in providing a framework for research on acquired dyslexias e.g. the phenomenon of deep dyslexic patients being affected by Imagery is explained in terms of some impairment in the semantic part of the model. An interesting extension of this research was provided by Jorm (1979a) who drew attention to similarities between developmental dyslexia and acquired deep dyslexia. He suggested that word imagery affected the ease with which a word could be read via the direct visual route. However Baddeley, Ellis, Miles and Lewis (1982) found that both developmental dyslexics and normal nine year old readers were less able to read low imagery words. They suggested that this imagery effect was mediated by differences in word age of acquisition. The pattern of reading performance obtained is similar to that of deep dyslexic patients (Coltheart, 1980). The question therefore remains as to whether the imagery effect is a specific feature of the performance of acquired and developmental deep dyslexics or whether it is a much more general phenomenon. Thus, Experiments 2 and 3 investigate imagery and age-of-acquisition effects in children and adults.

Henderson's lexical pooling model (1982) is based on the concept of orthographic neighbourhood, and it's importance is implicit in Glushko's activation-synthesis model, but more explicit in Shallice, Warrington and

McCarthy's Visual Word Form System (1983). Experiment 4, 5 and 6 investigates the effect of orthographic neighbourhood size on reading.

Patterson and Morton (1984) proposed a modified standard model which tried to incorporate the best aspects of the dual route theory and analogy theory. They described the orthography-to-phonology system which is comprised of two different sized units - graphemes and bodies. The development of a body sub-system is directly investigated in Experiments 7 and 8, reported in Chapter 7.

The age group studied in the experiments reported in this thesis, were children of eight and over. Seymour and Elder (1986) have reported an extensive study of the early stages of the reading process - the logographic stage; this study concentrated on the intermediate stages that follow. It is hoped that clear experimental evidence will be found for what remains, at present, convincing theoretical concepts.

C H A P T E R   F O U R

THE EFFECTS OF ORTHOGRAPHIC VARIABLES ON CHILDREN'S READING

4.1 INTRODUCTION

One of the models discussed in the last chapter was Shallice and Warrington's multiple levels models (1983, 1985). Their model assumes that orthographic units of different sizes - graphemes, consonant clusters, sub-syllabic units etc - are analysed, and information about them is passed on to corresponding units in the phonological system. This theoretical approach is of great relevance to the experiment reported in this chapter, because the experimental work aims to investigate whether orthographic units are an integral part of the reading process.

This introduction to Experiment 1 describes the developmental perspective of visual word recognition in children; defines orthographic regularity and describes the various measures of it; reviews relevant child and adult studies on this topic.

THE DEVELOPMENT OF VISUAL WORD RECOGNITION

Relatively little is known about how visual word recognition actually develops in children. There are several reasons for this lack of knowledge. Most of the research has focused upon the skills of the fluent reader - there has been comparatively little work on the emerging skills of the novice. Also the research that has been done with children has often not been motivated by the developmental perspectives on word recognition (either in terms of strategy or changes in skill) although there are some notable exceptions (e.g. Ehri, 1980, 1981; Francis, 1984; Seymour & Elder, 1985; Marsh, Friedman, Welch & Desberg, 1981). Instead, many investigators have been concerned with how well children perform on tasks designed to assess adult abilities and they have tried to interpret the children's performance with reference to models of fluent word recognition e.g. Barron (1981) interpreted children's lexical decision performance with reference to Glushko's activation-synthesis model. Despite the limited amount of work specifically motivated by developmental models, the work that has been carried out on children's word recognition does provide some information about what children know about words, how this knowledge is expressed in reading tasks, and how this knowledge changes with increasing age and reading experience.

The question of interest is how the use of this knowledge develops. It is still unclear when children first show that they can use orthographic structure in word recognition, whether this ability increases with age and reading experience, and how various tasks influence children's ability to use orthographic structure. In order to deal with these questions, the review of developmental experiments will be organised around the tasks which have been most commonly employed in assessing childrens use of orthographic structure in word recognition: visual search, tachistoscopic recognition, lexical decision and visual matching, after a brief definition of orthographic regularity and various ways of measuring it.

DEFINITION OF ORTHOGRAPHIC REGULARITY

The most important characteristics of orthographic regularity for studying word recognition are those that define the allowable patterns of letters within single words. Two different approaches have been taken in describing this regularity. Massaro, Venezky and Taylor provide a useful summary of these approaches in their 1979 paper.

i) **PROBABILISTIC APPROACH** This uses word tokens sampled from real texts to define probabilities of occurrence for single letters, bigrams and trigrams. From these data, various types of approximations to English words have been generated, i.e. First Order and Second Order approximations.

ii) **RULE-GOVERNED APPROACH** The rule-governed approach, in contrast, attempts to define orthographic structure in terms of the more general linguistic patterns of English spelling. A description of the major orthographic constraints is given in the studies of the English orthography by Venezky 1967, 1970. Rules define which letters or letter sequences are allowed (or not allowed) in which positions or graphemic contexts.

The earliest use of the probabilistic approach was the generating of English nonwords by Shannon (1948) and used by Miller, Bruner & Postman (1954) and by Wallach (1963) in studies of word recall. A zero-order approximation to English was generated by selecting each letter for a string randomly, giving equal weight to each letter. A first-order approximation resulted from the same procedure but with the letters weighted by their frequency of occurrence in English texts. Thus more highly regular nonwords were generated at the higher order approximations. i.e.

Zero order _ b o y _ p = $\frac{1}{27}$

as there are 27 characters (26 letters plus blank) and all are equally likely.

First order The five 1/27 figures are replaced by their actual probabilities of occurrence.

Second order The probabilities of drawing a blank, a 'b' given the blank, and 'o' given the 'b', a 'y' given the 'o', and a blank given the 'y' are multiplied. This reflects the relative frequency of bigrams.

In contrast to these sequential dependency schemes were the correlational approaches that used letter and letter-string frequency tables to produce nonwords with controlled bigram and trigram counts. Underwood & Schulz (1960) generated bigram and trigram frequencies from 2,080 words sampled from Thorndike & Lorge (1944) and weighted with respect to their frequency of occurrence. These counts were based on overall frequency of occurrence (tokens), rather than the number of different words contributing to the sample (types), and summed over all occurring word lengths and serial positions. Failure to account for word length provides obvious problems in describing orthographic regularity e.g. the trigram ING occurs relatively often summed over all word lengths but does not occur often in three - or four-letter words so word length must be controlled for when looking at orthographic variables. Summed bigram or trigram frequencies without taking account of serial position are also inadequate for a description of orthographic regularity. e.g. the bigram CK is legal at the end but not at the beginning of a word and serial position in assigning frequencies. These limitations are overcome in the Mayzner & Tresselt (1965) tables, which give bigram frequencies for each word position in words of from three to seven letters in length. However, it was found that both orthographically regular and irregular strings could be generated with either high or low bigram counts in the Mayzner & Tresselt tables i.e. there were instances of orthographically regular strings which had very low bigram counts, and orthographically irregular strings with very high bigram counts. Thus bigram frequency does not necessarily predict orthographic regularity.

Letter and letter-string frequency tables had previously been used to generate approximations to English without regard to summed bigram and trigram frequencies. Thus Hirata & Bryden (1971) generated ten-letter strings for orders of approximation to English from zero to four, using the Mayzner & Tresselt (1965) and Mayzner, Tresselt & Wolin (1965a, 1965b) tables of single letter, bigram and trigram frequencies. Lefton et al (1973) and others have used these lists in developmental studies of guessing missing letters in pseudowords.

Mason (1975) used the Mayzner & Tresselt (1965) tables to generate words with high and low positional frequencies. A word has a high positional frequency count if the letters for that word are in positions in which they are frequently found in words of the same length in texts. She found that positional frequency was a good predictor of letter search speeds in pseudowords. However, her test items confounded positional frequency and orthographic regularity - the pseudowords which had a high positional frequency tended to be orthographically regular and those with low positional frequency tended to be orthographically irregular - so the relative contributions of the two variables remained unclear. This is one question that Experiment 1 tries to address, by using new measures of orthographic regularity.

Rubin (1981) provided first- and second- order approximations to English for Paivio, Yuille and Madigan's (1968) 925 nouns. Solso, Juel and Rubin (1982) reviewed previous word and letter counts and compiled a comprehensive count of initial and terminal letters and bigrams based on Kucera and Francis (1967) corpus of English words. Their tables included frequency of occurrence and versatility i.e. the number of different words in which the bigram occurred. These tables were used to calculate initial and terminal letter and bigram frequencies and versatilities for the words used in Experiment 1.

There is evidence that beginning readers rely on the initial letter as their primary cue in recognising words (Francis, 1984), but there is evidence that older children increasingly focus on higher-order units, such as letter clusters, which are formed by orthographic structures and positional frequencies (Massaro & Taylor, 1979; Juel & Solso, 1980; 1981; Marsh, Friedman, Welch and Desberg, 1981). It is therefore predicted that children will find high frequency and versatility initial letter and bigram units easier to read than low frequency and versatility initial letter and bigram units, because they have already often encountered them.

Rubin (1981) also introduced a measure of orthographic regularity called the Orthographic Neighbour Ratio. It is the frequency of a word divided by the sum of the frequencies of all the words (neighbours) that can be generated by changing one of its letters. Thus the orthographic neighbour ratio provides an alternative orthographic measure where ease or difficulty of reading due to the number and frequency of neighbours can be directly tested.

The measures of orthographic regularity that are used in the first experiment are those of Rubin (1981) and Solso, Juel and Rubin (1982). These measures were chosen because they were the most recent probabilistic count of orthographic regularity to be found; they were based on a very comprehensive sample size of 983,400 words, and because they included an untested measure of orthographic regularity - the Orthographic Neighbour Ratio. These measures were used in Experiment 1, along with measures of Imagery, Frequency and Word Length to investigate the role of orthographic regularity in word recognition and reading in children. It is essential, however, to stress the difference between orthographic regularity, as defined by probabilitic measures, and regularity defined by the dual route model. The probabilistic measures include First Order Approximation and Orthographic Neighbour Ratio. Regularity, as defined by the dual route model, can be described by the spelling-to-sound correspondence rules elaborated by Venezky (1970).

4.2 Literature review - experimental work on orthography

In this section a review is presented of a number of experiments that involve the recognition of letters, nonwords and words and consider the role that orthographic regularity (as measured by the probabilistic measures described earlier) plays in the recognition process. Relatively little is known about how visual word recognition develops in children, so adult experimental work will also be cited because it may provide a useful framework for the developmental studies.

Developmental literature

The review of the developmental literature has been organised around the tasks which have been most commonly employed in assessing children's use of orthographic structure in word recognition: visual search, tachistoscopic recognition, lexical decision and visual matching.

Visual search has been a popular task for assessing the use of orthographic structure. Juola, Schadler, Chabot and McCaughey (1978) and McCaughey, Juola, Schadler and Ward (1980) had kindergarten, six, seven and nine year olds, plus college students search for a single letter in words, pseudowords (nonwords which do not break the rules of orthography), and random letter string displays made out of single items. The search times for the kindergarten and six year old children did not differ between the pseudowords and letter strings indicating that they were unable to use orthographic structure to facilitate their search. The seven and nine year olds were influenced by orthographic structure as they searched more rapidly through pseudowords than letter strings. If early reading experience results in the acquisition and use of orthographic information about commonly occurring spelling patterns, then advantages for both words and regularly spelled pseudowords over irregular nonwords in visual search should emerge at the same time because irregular nonwords have less commonly occurring spelling patterns. Alternatively, if information about specific words was used in recognition before more general orthographic rules, then they expected words to show an advantage over both pseudowords and nonwords. Differences in performance for pseudowords and nonwords should then appear later, after a significant amount of orthographic information had been acquired.

The conclusion drawn by McCaughey et al was that, as reading skills develop, children apparently come to process familiar words differently from other letter strings. Although adults and older children seem to use orthographic information to aid word and pseudoword perception, children under seven have not yet developed these rules from their knowledge of specific words and individual letters sufficiently to facilitate rapid word perception. It is as though children must first build up a substantial reading vocabulary before they acquire enough information about the regularities in standard English spelling to help rapid processing of regular but unfamiliar letter strings.

There are, however, several experiments in the literature which failed to show an orthographic effect in visual search tasks. Gibson, Tenney, Barron and Zaslow (1972) found no difference in search time for a letter between pseudowords and letter strings for either grade five or adult subjects. As these investigations pointed out, however the letter N was the target throughout the experiment, and the subjects may have learned to use its graphic features to locate the target in the search list.

Similarly, Stanovich, West and Pollack (1978), using a word search task with 8 year olds, 11 year olds and adult subjects, also failed to find an increasing effect of orthographic structure with age. In a word search task subjects search for words in sets of words, pseudowords or random letter strings. They assumed that the task would be easiest for the word passage where there is more orthographic regularity and the letters have high positional frequency.

On the basis of results from these developmental studies Stanovich and West (1979) had good and poor eight year old readers search for words through fields consisting of words, pseudowords (high bigram frequency) and nonwords (low bigram frequency). They found no effect of orthographic structure on the performance of better readers when analysing the search times. They concluded that beyond the earliest stages of reading acquisition, the increasing use of positional frequency is not responsible for increasing reading ability.

These results conflict with the findings of Mason (1975). She used positional frequency (a variable different from, but correlated with bigram frequency). The major difference between the two studies was that Stanovich and West used the word as the target for which the subjects searched, while Mason's subjects searched for a letter. Mason (1975) studied the contribution of one aspect of orthographic regularity - positional frequency. Good and poor sixth-grade readers (11 year olds) searched through six letter strings for the presence or absence of a target letter. Words and nonwords were used, and the nonwords differed in the degree of orthographic structure as defined by positional frequency. The positional frequency of a letter in a letter string is the frequency of occurrence of that letter in the same position in words of the same length sampled from common texts. Given this definition, a letter string was given a summed positional frequency that represented the sum of the positional frequency of all the letters in the string. Mason investigated whether search time for a letter was an inverse function of the summed positional frequency of the letter string. The implicit model of the letter search task was that the subjects must first recognise the letters in a string and then compare these letters to the target letter. Thus the children had to determine the presence or absence of a target letter in six-letter strings that were either words or nonwords. Two nonword anagrams were constructed for each word; one high positional nonword anagram (with a higher value than the word) and one low positional frequency nonword anagram.

Differences in the search times for a given target letter in different letter strings should reflect differences in the time needed to recognise the letters of the strings. Mason found that good readers were faster in determining the presence or absence of a target letter on strings with high positional frequencies than on strings with low positionl frequencies. Her poor readers showed no difference. The results supported the idea that the time needed to recognise the letters in a string is influenced by the likelihood of letters occurring in their most common positions, and that good readers were able to make use of this orthographic information whereas the poor readers were not.

Although summed positional frequency appeared to account for the recognition times in Mason's study, Venezky and Massaro (1977) questioned whether it was the critical variable that defined orthographic regularity.

They examined the nonword anagrams which were generated from the letters that made up the word PERSON, which has a summed positional frequency of 1,141, based on the Mayzner and Tresselt (1965) single-letter table. For example the string PORNES contains the same letters in different positions and has a count of 1,858. The string ENSPRO has a count of 383. However PORNES is spelled like an English word and therefore would be relatively easy to recognise, whereas ENSPRO violates what we know about English spelling and should therefore be relatively difficult to recognise. They suggested that it was orthographic regularity as defined by graphemic rules and not as defined by the summed positional frequencies which influenced recognition times when these variables were independently varied in a target search task.

Thus, in a subsequent experiment Mason (1978a) used words from Baron & Strawson (1976, Exp.2), which were exception words or regular words as defined by graphemic rules, so as to counter Venezky & Massaro's criticism. Highly skilled and less skilled college students had to read these words aloud. She provided no rationale for why she switched from the visual search task to a reading task. Results showed that less skilled readers made more errors and took longer to read both regular and exception words aloud than did highly skilled readers. The effect of the exception-versus-regular word variable was negligible for both groups of readers. The exception words necessitated a visual access of the lexicon for naming, because they could not be phonologically recoded and pronounced by application of spelling-to-sound correspondence rules elaborated by Venezky (1970). Since the orthographic variable of regularity of pronunciation made no difference either within or across reading groups she concluded that neither groups of readers went from print to sound via phonological recoding.

In another experiment, Mason (1978b) asked skilled and less skilled readers to read words and nonwords which varied in length (four and six letters) and positional frequency (high and low). Mason found a length effect in both highly skilled and less skilled readers, the effect being limited only to nonwords with highly skilled readers. The result of greater interest was the highly significant effect of single letter positional frequency. High positional frequency strings were named more quickly than were low positional frequency strings. This effect, however was limited to the nonword strings. This experiment therefore illustrates the use of orthographic knowledge, at least in the reading of nonwords, in adult readers.

Mason (1975, 1978b) was able to show an effect of orthographic regularity using one measure - positional frequency, but not conclusively using another measure - regularity as defined by graphemic rules. Her results may be queried for two reasons. Firstly blocking of regular and irregular words was employed, and this may have encouraged students to use strategies specific to each list of items which eliminated any differences between them. Secondly, Mason used irregular words which were of higher frequency than regular words. This difference alone would be sufficient to eliminate the difference between regular and irregular words. So, taking into consideration these two points, her results should be viewed with caution.

On the whole, the visual search results reported do not provide clear evidence of a developmental trend in the use of orthographic information. Some of the visual search results suggests that by seven years of age, children can use orthographic structure to facilitate word recognition and the reduction in response times does not change very much with increases in age and reading experience. However, conflicting results arising from differences in the tasks, target size, experimental designs, and measures used (accuracy versus reaction time) indicate that these conclusions should be regarded as being tentative.

Tachistoscopic recognition tasks differ from visual search tasks as the items are displayed very briefly and response accuracy rather than response latency is typically measured. In an early tachistoscopic recognition task strings of eight letters were used by Lefton and Spragins (1974). One-half of the items were first- and the other half fourth-order approximations to English and they were briefly presented for durations ranging from 50-300 milliseconds. Lefton and Spragins found that six year olds were no more accurate on the fourth- than on the first- order items, whereas the eight year olds, ten year olds and adults were more accurate on the fourth order items. A somewhat similar pattern of results was obtained by Lefton, Spragins and Byrnes (1973) in a task in which subjects were required to guess the missing letter in letter strings. They found that eight and ten year olds were more accurate at guessing letters in fourth- than first- order strings, particularly on the right side of the display (where there was more sequential constraint). The six year olds, on the other hand, showed no difference in guessing accuracy between the two types of strings.

Bishop (1976), cited in Barron (1981), also employed tachistoscopic presentation, but she used a probe task in which subjects were required to decide which one of two letters appeared in an item in addition to a task where they just reported what they saw. The advantage of the probe over the report task is that it reduces the contribution of memory and guessing factors to the accuracy of item recognition. In the report task, both the eight year olds and eleven year olds were more accurate on pseudowords than letter strings, but with the probe procedure, only the eleven year olds were significantly more accurate on the pseudowords than the letter strings. Again using letter string performance as a base line, the relative increase in percentage accuracy on the pseudowords declined over age in the report task (72.5% for eight year olds, 58% for eleven year olds) and increased over age in the probe task (5.2% for eight year olds, 8.1% for eleven year olds).

Taken together, the experiments of Lefton and Spragins (1974) and Bishop (1976) suggest that between the ages of seven and nine children begin to use orthographic structure to help word recognition in tachistoscopic tasks. Contrary to some of the visual search results, the tachistoscopic evidence suggests that the percentage increase in accuracy due to orthographic structure varies over age. This variability might be due to differences in tasks and materials, or it may reflect floor or ceiling effects upon performnce at various age levels.

Henderson and Chard (1980) used response time as a dependent variable in a lexical decision task in which subjects were required to decide whether or not an item was a word. In this task, the absence of orthographic structure would allow subjects to decide more rapidly that an item is not word. Using seven and nine year old subjects and six letter words and nonwords, they varied the single letter positional frequency values of nonwords independently of whether or not they contained vowels. Both groups of readers made faster nonword decisions on the items when they were low than high in positional frequency and when vowels were absent rather than present. In addition, vowels did not have any effect on the response times of the nine year old subjects to the low positional frequency nonwords, but did influence both the high and low positional frequency nonword decisions of the seven year olds. These results are also consistent with the findings that subjects as young as seven are affected by the characteristics of orthographic structure. They also suggest a developmental shift whereby the older subjects may be able to reject items with low positional frequency very early in the lexical search process, possibly on the basis of initial or final bigrams. The importance of initial and terminal bigram frequencies in children's word recognition was illustrated by Forster and Gartlan (1975) who performed a lexical decision experiment where the presentation of a word (e.g. STEADILY) was immediately preceded by the display of either its first and last two letters (ST LY), its middle four letters (EADI), its first four letters (STEA), or its last four letters (DILY). They found that only the first of these conditions led to priming of the word, when compared to a condition where there was no prior display. That is, only the presentation of the extremeties of the word allowed processing to get underway. Forster interpreted this in terms of the access file being divided into "bins" whereby words with the same access code (i.e., the same first and last letter groupings) are stored together. Thus prior presentation of the first and last letters allows the correct bin to be accessed in advance of the presentation of the whole word.

Underwood and Bargh (1982), in an investigation of the interaction between three sources of information about a printed word, required skilled, adult readers to perform a simple reading task. Naming latency was observed as a function of case of presentation, congruency of context supplied in a previously presented sentence, and orthographic regularity. Their definition of orthographic regularity was that of pronounceability. They predicted that words which could be pronounced using the rules specified by Venezky (1970) would be expected to gain faster responses in a reading task.

They did not find a main effect of orthographic regularity in their analysis, although it did interact with case of presentation - regular words were read faster than irregular words only when they were printed in upper case letters. They thought this was because lower case words, with available distinctive graphemic information encouraged a recognition strategy based on direct visual access. The interaction of all their variables - word shape, orthographic regularity and context led them to support models of reading suggested by Rumelhart and MacClelland (1982) where there is no linear passage of information, but the recognition process strategically uses any source of information which is available.

Considered together, the results from visual search, tachistoscopic, lexical decision and reading tasks provide an inconsistent picture of the development of the use of orthographic structure. Some of the studies provide some positive evidence that between seven and nine years of age children seem to be able to use some of the gross characteristics of orthographic structure in visual word recognition. There have been instances where younger children were seen to use orthographic structure, but it appears to be critically dependent on using three letter items with CVC structures. Finally, if orthographic effects appear, their relative size does not seem to change very much across age group in visual search. In tachistoscopic tasks, however, the change in the relative size of the orthographic effect appeared to be more marked.

In this review of experimental work with children we have seen that experimental tasks differ to the extent that they require differing degrees of processing by the subject. The nature of the task should be accounted for both in the analysis of the results and in the implications that are drawn for theory. However some researchers have appeared unconcerned with the processes involved in tasks such as searching for a target letter in a letter string, pronouncing a letter string, or determining whether a particular letter string is a word.

In each of the experiments reviewed orthographic structure was assumed to be an important variable; yet the definitions of this structure are widely disparate, varying from single-letter positional frequency (Mason 1975) to pronunceability (Gibson et al 1972). Experiment 1 will look at probabilistic measures of orthographic regularity, as defined by Massaro, Venezky & Taylor (1979).

The role of orthographic regularity in adults

Although there have been relatively few experiments on how word recognition develops in children, there is a rapidly expanding literature that demonstrates a central role for orthographic regularity in adult word recognition.

Zechmeister (1969) first put forward the idea that words differ in orthographic distinctiveness. He defined distinctive orthography as "those structural characteristics of a word which make it physically unusual, interesting or distinctive". e.g. When considering the two words GENIE and GNOME - both have approximately the same frequency of occurrence, the same initial letter and length, and both represent imaginary, fairylike creatures. Yet, GNOME is structurally more unusual, largely because of the infrequent first-position bigram, GN. He asked adult subjects to rate 150 very low frequency words for orthographic distinctiveness or pronunciability on a nine point scale. Orthographic distinctiveness was found to be inversely related to pronunciability (r = 0.61). Furthermore, words rated as having highly distinctive orthography had very low frequency bigrams in either the first or last position. Besides the relative frequency of contiguous letter combinations, the repetition of letters which were next to each other, was also seen as orthographically distinctive e.g. FLOOR and ALLEY. The implication from Zeichmeister's study was that distinctive orthographic features somehow facilitate recognition, because the more highly orthographically distinctive words were rated as more pronounceable.

Barron & Pittenger (1974) presented pairs of high frequency words, orthographically acceptable pseudowords and random nonword letter strings in a "same-different" task. They found that the mean reaction times for "same" judgements were ordered; words were faster than pseudowords and pseudowords were faster than nonwords. The reaction times for the "different" judgements showed no differences among the three types of words. They suggested that "same" judgements were based upon a comparison process which performed a self-terminating search of the graphemic information in words.

Travers and Olivier (1978) studied pronounceability and statistical "Englishness" as determinants of letter identification. They found that established "Englishness" of letter strings, assessed by a measure based on bigram and trigram frequencies, exerted a significant effect on report accuracy, independent of string pronounceability. Their experiment was in keeping with an impressive array of studies which demonstrated that "word-like" non-words of various kinds exhibit some of the perceptual or response characteristics of words (e.g. Gibson, Pick, Osser & Hammond, 1962; Baron & Thurston, 1973; McClelland & Johnston, 1977; Spoehr & Smith, 1975). The study of wordlike nonwords is of interest because it bears promise of revealing an important aspect of the word-perception mechanism, in particular, of showing what kind of knowledge about morphological and orthographic structure the skilled reader uses in recognising words.

Perhaps the simplest hypothesis about the skilled reader's knowledge is that they know which letter clusters occur frequently in their printed language. However, several studies have found little or no relationship between cluster frequencies within words or nonwords and perceptibility of those stimuli, as assessed by a variety of measures (e.g. Gibson, Shurcliff & Yonas, 1970; McClelland & Johnston, 1977; Spoehr & Smith, 1975). More recent papers have generally advanced nonstatistical conceptions of the psychologically relevant aspects of word structure, such as grapheme-phoneme correspondence (Gibson et al., 1962), orthographic regularity (Gibson et al., 1970; McClelland & Johnston (1977) or syllabic organization (Spoehr & Smith, 1975). Thus probabilistic measures of orthography still need to be proved useful experimentally.

One of the most extensive examination of the relative importance of cluster frequencies and alternative conceptions of structure has been conducted by Gibson and her colleagues. Gibson, Pick, Osser & Hammond (1962) showed that pronounceable nonwords (eg. GLURCK) were read more accurately than unpronounceable nonwords formed by reversal of initial and final consonant clusters of the pronounceable set (e.g. CKURGL). Gibson originally interpreted the correlation between pronounceability and reading accuracy as showing that letter clusters that map consistently into sounds become perceptual chunks. Later she amended this interpretation when she found "pronounceability" effects in perceptual reports of deaf subjects (Gibson et al, 1970) and suggested instead that sheer orthographic regularity could lead to perceptual chunking without the direct mediation of sound.

McClelland & Johnston (1977) contrasted the effects of cluster frequency with those of "orthographic regularity". Orthographic regularity was assessed by judgement of whether a letter sequence could be a word in Englilsh; thus their definition orthographic regularity was operationally, if not conceptually, equivalent to pronounceability of the sequence. Cluster frequency was assessed by summing bigram frequencies across the four-letter stimulus strings. It was found to exert minimal impact on single letter perceptibility for skilled readers.

Spoehr and Smith (1975) also showed that average bigram frequencies for their nonword strings were unrelated to how fast a letter could be perceived in a nonword string.

The studies cited above use summed or average bigram or trigram frequencies as overall measures of the statistical Englishness of nonword strings. None of the authors gave an explicit rationale for choosing this measure, presumably because it bears an obvious intuitive relation to "Englishness", conceived in terms of cluster frequencies.

Massaro, Venezky and Taylor (1979) evaluated previous research findings that familiarity with the orthographic strucure within a letter string can facilitate the processing of the component letters. A probabilistic description based on the frequency of occurrence of letters in each position and a rule-governed description (as defined by Venezky, 1970) were independently varied in the construction of six letter nonword strings. College students and eleven year olds had to search for a target letter in these strings. For both groups of readers, orthographic regularity and summed positional frequency were found to have only a small facilitative effect on reaction time. In contrast, reaction times to say 'NO' increased dramatically with increases in the number of letters in the nonword string that were physically similar to the target letter. They suggested that the feature detection strategy in target search did not encourage use of the orthographic structure.

HYPOTHESES - EXPERIMENT 1

Experiment 1 addresses the issue of comparing various probabilistic measures of orthographic structure, as defined by Massaro, Venezky & Taylor (1979) to see which one most adequately measures the sort orthographic knowledge being used by children. It investigates various measures of orthography in a lexical decision task and a reading aloud task with two classes of children (mean age 8:6 and 9:6).

Before further discussion of the experiment, one variable requires more detailed introduction - the Orthographic Neighbour Ratio (Rubin 1981). This is the ratio of the frequency of a word in the Kucera and Francis (1967) count divided by the sum of the frequency of all its orthographic neighbours - all the words that can be formed by changing one letter of original word. This ratio provides a variable which can be used to assess interactive-activation models such as Rumelhart & MacClelland (1982), and the multiple levels model of Shallice and Warrington (1983), which hypothesises that partial orthographic information about a word is obtained and a decision is made on the basis of the relative frequency of the possible response.

The ratio not only takes into account the number of orthographic neighbours that a particular word has, but also their frequencies. This is an alternative measure of orthographic structure.

FIGURE 4.1 THE ORTHOGRAPHIC NEIGHBOUR RATIO, RUBIN (1981)

ORTHOGRAPHIC NEIGHBOUR RATIO	STIMULUS ITEM
↑ HIGH ONR e.g. 1.000	HAS NO NEIGHBOURS
	HAS A FEW, LOW FREQUENCY NEIGHBOURS
	HAS MANY, LOW FREQUENCY NEIGHBOURS
	HAS A FEW, HIGH FREQUENCY NEIGHBOURS
LOW ONR e.g. 0.001	HAS MANY, HIGH FREQUENCY NEIGHBOURS

It is hypothesised that if orthographic structure is an important variable in word recognition then there will be a high correlation between measures of orthographic regularity and children's performance in lexical decisison and reading tasks. It is expected that words with low ONR values will be easier to read than words with high ONR values, because a) they have more neighbours, b) these neighbours are more frequent.

The question of whether there is a developmental trend in children's use of orthographic knowledge remains unanswered because the results of previous reseach have proved contradictory. Marsh, Friedman, Welch and Desberg's (1981) developmental model of reading acquisition predicts that young unskilled readers are unaffected by measures of orthographic regularity, but that as children get older they become more reliant on orthographic units of increasing size. For example at Stage Two variables such as initial letter frequency and versatility, word length, terminal letter frequency and versatility are used because they provide the orthographic information that children need to read a word. By Stage Four when children are able to use an analogy strategy, more global orthographic variables such as the Orthographic Neighbour Ratio may provide useful information for reading performance and the lexical decision task.

It is hoped that Experiment 1 will help validate the cognitive-developmental approach both as a descriptive theory and as an explanation for reading acquisition.

Experiment 1 aims to clarify whether or not the qualitative differences described by Marsh et al (1981) are observed in children's reading performance. Poor readers may show characteristics of Marsh et al's Stage 2 - Discrimination Net Guessing, where children are described as paying attention to orthographic features such as initial and terminal letters, so as to discriminate one word from another. The hypothesis is that there will be a positive correlation between a word's initial and terminal letter frequency and versatility and ease of reading i.e. the higher the value of the measure, the easier it is to read.

As the child increases in reading skill he may use larger orthographic units such as bigram frequencies and versatilities, and more global orthographic information such as first and second order approximations to English. Thus it is hypothesised that average readers will use these orthographic measures in reading.

Good readers will be at Stage 4 - that of hierarchical decoding. They are able to deal with conditional and higher order rules and are also able to use an analogy strategy. One might expect that more complex measures of orthographic structure, such as the Orthographic Neighbour Ratio, may provide useful information in the reading process.

Both Marsh et al (1981) developmental model and Shallice and Warrington's multiple levels model assume that orthography is of great importance in the reading process. The multiple levels model assumes that orthographic units of different sizes are used, and the bigger the unit, the faster the recognition process is. Experiment 1 aims to investigate whether orthographic information is used by children in reading aloud and in lexical decision. These two tasks were used because although the literature review of developmental and adult data has shown that probabilistic measures of orthography are <u>theoretically</u> considered to be of importance in reading and lexical decision, they still need to be demonstrated experimentally. There has been no indication of whether orthographic regularity is more salient in the lexical decision task than in the reading task for adults, although Henderson and Chard (1980) showed that high positional frequency of letters in a stimulus item made the lexical decision task easier for children. Thus a tentative hypothesis for the differential effects of orthographic regularity for the two tasks would be that orthographic information would be more salient in the lexical decision task than for the reading task.

The experiment reported in this chapter uses a multiple regression approach, in common with a number of previous studies in this general area (Cohen and Cohen, 1975; Gilhooly and Gilhooly, 1979). In all cases words were the units of analysis rather than subjects. This research strategy was adopted because the large number of potentially relevant variables that are correlated with reading makes it impossible to carry out factorial experiments in which confounded variables are balanced out or experimentally manipulated, while still retaining a reasonable number of words per condition.

4.3(a) METHOD

SUBJECTS The subjects were 38 children from the first and second year junior classes of an inner city school, categorised by their class teacher as either good, average or poor readers. The mean age for the first year class was 8yrs 6mths, for the second year class, (9 yrs 6 mths).

This classification of children was used to see if there was a detectable developmental trend, as outlined by Marsh, Friedman, Welch and Desberg (1981).

STIMULI The stimuli were 79 words selected from Rubin's (1981) list of 925 nouns, with letter length varying from 4 to 9 letters. There were 79 nonword controls - made up of anagrams of the 79 words. These were not used in the analysis, but were included to ensure that the children attempted lexical access. The stimulus items are presented in Appendix A4.1. The variables available for the word list were Frequency, Imagery, First-order-approximation to English, Second-order-approximation to English, Orthographic Neighbour Ratio, Initial Letter Frequency, Initial Letter Versatility, Initial Bigram Frequency, Initial Bigram Versatility, Terminal Bigram Versatility and Word Length. These measures are presented in Appendix A4.2.

PROCEDURE

Presentation of the stimuli and recording of latencies were controlled by a BBC microcomputer equipped with a real-time clock and a video terminal (Hantarex). The real-time clock operated on a millisecond time base. On each trial, subjects pressed the space bar to make the stimulus appear. In both the lexical decision task and the reading task the response was said aloud into a microphone connected to a voice-key interfaced to the computer, which was triggered by the onset of the subject's voice. On each trial the stimulus stayed on the screen until the subject made a response. Timed responses were stored along with information as to whether the response was correct or not. The order of presentation was randomised for each subject both in the lexical decision task and the pronunciation task. This basic procedure was used in all the studies reported.

4.3(b) **RESULTS**

An initial comparison of the lexical decision and reading performance of the two classes showed that combining their data for further item analysis was justified - their level of performance on the two tasks were similar to each other. The mean number of words correct in the lexical decision task (Class 1 - 2nd year juniors, (n=20) $\bar{x}$ = 63.45, Class 2 - 1st year juniors, (n=18) $\bar{x}$ = 62.89) was not significantly different (t = 0.19, p = 0.85). The mean number of words correctly read by the children (Class 1 (n=20) $\bar{x}$ = 54.60, Class 2 (n=18) $\bar{x}$ = 56.61) were not significantly different (t = 0.31, p = 0.76)). It can be seen that both classes found the lexical decision task easier than the reading task.

The thirty eight subjects were therefore subdivided into good, average and poor reader groups according to class teacher ratings. Although this subdivision can be criticised because it assumes homogeneity within reader groups, it was done to see if the pattern of results of the three reader groups could provide any confirmation of Marsh, Friedman, Welch and Desberg's (1981) developmental model described in Chapter 3. Mean chronological and reading ages of the three reader groups are presented in Table 4.1.

Raw data for reader group performance on the lexical decision and reading tasks is presented in Appendix 4.3. Mean Reader group performance on the lexical decision and reading tasks are set out in Table 4.2.

A correlation analysis was performed with lexical decision performance (LD3) and reading performance (PRON3) as the dependent variable and the fifteen word measures as independent variables. These are the accuracy scores for lexical decision and reading aloud. This was done in three separate analyses for the three reader groups - a summary table is presented in Table 4.3. The independent variables which correlated most highly ($p < .001$) with the dependent variables were Imagery, First Order Approximation to English, Second Order Approximation to English and Word Length.

The American Heritage Frequency Count proved to be more highly correlated with performance than did Kučera and Francis for the three reader groups, suggesting that for this data set the American Heritage count is the better frequency measure to use in future analysis.

There were only a few high correlations within the independant variables (see Appendix 4.4(b)) with the exception of First Order Approximation to English and second Order Approximation to English ($r=0.90, p<.001$), First order Approximation to English and Word Length ($r = 0.97$, $p<.001$), Second Order Approximation to English and Word Length ($r = 0.89$, $p < .001$) and the two frequency

TABLE 4.1: CHRONOLOGICAL AND READING AGES OF SUBJECTS

EXPERIMENT 1

READER GROUP	N	MEAN CHRONOLOGICAL AGE	MEAN READING AGE
GOOD	10	8:8	9.54
AVERAGE	19	9:2	8.52
POOR	9	9:1	7.62

TABLE 4.2: MEAN PROPORTION CORRECT SCORES FOR THREE READER GROUPS ON LEXICAL DECISION & READING ALOUD TASKS

READER GROUP	LEXICAL DECISION		READING ALOUD	
	$\bar{x}$	S.D	$\bar{x}$	S.D
GOOD (n=10)	0.868	2.14	0.868	1.83
AVERAGE (n=19)	0.792	3.83	0.710	4.63
POOR (n=9)	0.741	1.95	0.504	2.73

$\bar{x}$ refers to the mean proportion of children who read or made correct lexical decisions about the words.

TABLE 4.3

SUMMARY OF PEARSON CORRELATION COEFFICIENTS FROM THE MATRIXES OF THE THREE READER GROUPS

		GOOD (n=10)		AVERAGE (n=19)		POOR (n=9)	
		LD3	PRON3	LD3	PRON3	LD3	PRON3
LEXICAL DECISION	CLASS 1 LD1	.990**	.700**	.923**	.744**	.835**	.592**
	CLASS 2 LD2	.754**	.361**	.960**	.767**	.850**	.568**
	CLASS 1&2 LD3	1.000	.675**	1.000	.802**	1.000	.688**
READING	CLASS 1 PRON1	.694**	.994**	.748**	.959**	.636**	.948**
	CLASS 2 PRON2	.183	.532**	.732**	.881**	.662**	.937**
	CLASS 1&2 PRON3	.675**	1.000	.802**	1.000	.688**	1.000
	KF	.184	.155	.255	.254	.247	.286
	AH	.238	.240	.310*	.343*	.335*	.364*
	I	.377**	.468**	.371**	.466**	.360**	.435**
	FOA	.343**	.448**	.517**	.417**	.443**	.340*
	SOA	.354**	.422**	.530**	.373**	.459**	.320*
	ONR	.240	.135	.113	.202	.067	.229
	ILF	.042	.027	.123	.034	.070	.011
	ILV	.033	-.024	-.016	-.025	-.102	-.062
	TLF	-.105	-.094	-.108	-.083	-.099	-.036
	TLV	-.231	-.197	-.196	-.164	-.190	-.085
	IBF	-.070	-.092	-.061	-.244	-.143	-.170
	IBV	.009	-.155	-.090	-.253	-.160	-.138
	TBF	-.104	-.131	-.135	-.029	-.127	-.094
	TBV	-.396**	-.339*	-.243	-.232	-.203	-.262*
	WL	-.392**	-.463**	-.542**	-.448**	-.499**	-.348**

* SIGNIFICANT AT $p < .01$ ** SIGNIFICANT AT $p < .001$

measures - Kučera and Francis and American Heritage Frequency ($r = 0.85$, $p < .001$) indicating that there will be no problems of multicollinearity.

ITEM ANALYSIS

There is sufficient experimental evidence to argue that some kind of orthographic knowledge helps children to recognise letter strings. The item analysis was aimed at finding which one of (or which combination of) the orthographic measures best reflects this, so as to obtain substantial evidence for constructing a model of the young reader's knowledge of orthographic structure.

Firstly a difficulty value for each word was calculated by counting the number of subjects who read it correctly and a series of logistic models were fitted to this measure. This was done for each of the three reader groups. The GLIM computer program (Baker and Nelder, 1978) was used to perform the analyses. The procedures are equivalent to regression analyses but a binomial error term is used as it is appropriate for proportional data, and the data was transformed using the formula ($\log \left| \frac{p}{1-p} \right|$). This procedure was used for all the item analyses performed in the experiments that follow. Various models can be fitted to the data and the goodness of fit can be compared for each one. It should be noted that GLIM obtains estimates for each predictor adjusted for other predictors in the model and so the estimates are not affected by order of entry into the regression analysis.

MODEL 1: PARTIAL CORRELATIONS FOR INDEPENDENT VARIABLES (CORRECTED FOR WORD LENGTH AND WORD FREQUENCY)

In this model each of the measures of orthographic regularity was used as a predictor to investigate which one of the measures best reflects the kind of orthographic knowledge used by the children in this sample, if it is used at all. Log Word Frequency - (American Heritage) and Word Length were included in order to control for their effects which have been observed in previous research and therefore of less interest in this study. The estimates, t values, significance levels and partial r for the three reader groups are presented in Table 4.4,Table 4.5 and Table 5.6.

There were several variables which remained significant after taking frequency and word length into account. (Since frequency and word length effects are strong and reliable effects they may swamp orthographic regularity effects which may be present). The partial correlation for Imagery was high in most reader groups and both types of task.

Of the several probabilistic measures of orthographic regularity - the Orthographic Neighbour Ratio was significant in the lexical decision task for good and average reader groups, while First Order Approximation to English was the significant variable for poor readers. In the reading task initial bigram frequency and initial bigram versatility were significant variables for average readers;

TABLE 4.4

MODEL 1: PARTIAL CORRELATIONS FOR INDEPENDENT VARIABLES (CORRECTED FOR WORD LENGTH AND WORD FREQUENCY)

GOOD READERS

LEXICAL DECISION

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.392	4.60	<.001	0.272
FOA	-0.161	2.66	< .01	0.161
SOA	-0.009	0.27	NS	0.017
ONR	-1.612	2.80	< .01	0.170
ILF	-0.005	0.70	NS	0.042
ILV	0.001	0.96	NS	0.058
TLF	-0.002	0.91	NS	0.054
TLF	-0.00005	0.89	NS	0.052
IBF	-0.015	2.26	< .05	0.133
IBV	-0.004	0.49	NS	0.029
TBF	0.006	0.83	NS	0.050
TBV	-0.019	1.61	NS	0.096

READING

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.411	5.34	<.001	0.360
FOA	-0.002	0.04	NS	0.000
SOA	-0.0007	0.03	NS	0.000
ONR	-0.751	1.40	NS	0.096
ILF	-0.010	1.49	NS	0.102
ILV	-0.0005	0.52	NS	0.035
TLF	-0.0006	0.26	NS	0.018
TLV	-0.00001	0.23	NS	0.015
IBF	-0.008	1.15	NS	0.077
IBV	-0.013	1.94	NS	0.128
TBF	-0.001	0.21	NS	0.014
TBV	-0.018	1.59	NS	0.107

TABLE 4.5

MODEL 1: PARTIAL CORRELATIONS FOR INDEPENDENT VARIABLES CORRECTED FOR WORD LENGTH AND WORK FREQUENCY

AVERAGE READERS

LEXICAL DECISION

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.222	4.74	<.001	0.244
FOA	-0.022	0.66	NS	0.034
SOA	0.026	1.36	NS	0.071
ONR	-0.948	2.94	< .01	0.155
ILF	-0.002	0.51	NS	0.027
ILV	-0.0003	0.46	NS	0.023
TLF	-0.0006	0.43	NS	0.022
TLV	-0.000003	0.07	NS	0.000
IBF	-0.003	0.60	NS	0.031
IBV	-0.004	0.80	NS	0.041
TBF	-0.0003	0.08	NS	0.00
TBV	-0.00005	0.0007	NS	0.00

READING

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.348	8.16	<.001	0.374
FOA	-0.041	1.36	NS	0.062
SOA	-0.029	1.66	NS	0.077
ONR	-0.231	0.84	NS	0.038
IL	-0.007	1.84	NS	0.084
ILV	-0.0002	0.36	NS	0.017
TLF	-0.0007	0.64	NS	0.029
TLV	-0.000003	0.08	NS	0.000
IBF	-0.017	4.68	<.001	0.215
IBV	-0.018	4.34	<.001	0.198
TBF	0.007	1.75	NS	0.081
TBV	-0.004	0.65	NS	0.030

TABLE 4.6

MODEL 1: PARTIAL CORRELATIONS FOR INDEPENDENT VARIABLES (CORRECTED OR WORD LENGTH AND WORD FREQUENCY

POOR READERS

LEXICAL DECISION

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.202	3.29	<.01	0.232
FOA	-0.103	2.16	<.01	0.156
SOA	0.001	0.06	NS	0.000
ONR	-0.386	0.93	NS	0.067
ILF	-0.006	1.04	NS	0.074
ILV	-0.001	1.49	NS	0.106
TLF	-0.0003	0.18	NS	0.012
TLV	-0.00001	0.26	NS	0.019
IBF	-0.006	1.12	NS	0.080
IBV	-0.007	1.2	NS	0.085
TBF	-0.002	0.3	NS	0.021
TBV	-0.003	0.31	NS	0.023

READING

PREDICTOR	ESTIMATE	t(75)	p	partial r
I	0.405	6.59	<.001	0.390
FOA	0.00009	0.002	NS	0.000
SOA	0.003	0.14	NS	0.008
ONR	-0.669	1.84	NS	0.110
ILF	-0.007	1.37	NS	0.078
ILV	-0.0004	0.58	NS	0.033
TLF	-0.00008	0.06	NS	0.000
TLV	0.00003	0.66	NS	0.038
IBF	-0.014	2.66	<.01	0.153
IBV	-0.009	1.67	NS	0.096
TBF	-0.0002	0.04	NS	0.000
TBV	-0.013	1.3	NS	0.075

initial bigram frequency was a significant variable for poor readers. No measures of orthographic regularity reached significance for good readers in the reading task.

These variables were then used to build up a model which would best describe the reading and lexical decision performance of the children.

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

The independent variables that were fitted as predictors of difficulty value in Model 2 were all the variables that were still significant after adjustment for word length and word frequency in the previous model; and also Log Word Frequency and Word Length (in order to control for their effects). i.e. the full model for good readers in the lexical decision task was I+FOA+ONR+IBF+WL+WF. The results of this model fitting for the three reader groups are shown in Appendix 4.5 to Appendix 4.10.

In the analysis for good readers, Imagery ($r = 0.187$, $t(73) = 3.16$, $p < .01$), Orthographic Neighbour Ratio ($r = 0.182$, $t(73) = 2.97$, $p < .01$), First Order of Approximation ($r = 0.144$, $t(73) = 2.37$, $p < .05$) and Word Frequency ($r = 0.401$, $t(73) = 6.06$, $p < .001$) were the significant predictors in the lexical decision task. In the reading task, the significant predictors were Imagery ($r = 0.360$, $t(76) = 5.34$, $p < .001$), Word Frequency ($r = 0.288$, $t(76) = 4.13$, $p <$

.001) and Word Length (r = 0.233, $t(76) = 3.38$, $p < .01$) - Imagery accounting for the greatest proportion of the variance.

The item analysis or average readers showed that Imagery (r = 0.216, $t(75) = 4.18$, $p < .001$), Orthographic Neighbour Ratio (r = 0.105, $t(75) = 2.01$, $p < .05$), Word Length (r = 0.172, $t(75) = 3.28$, $p < .01$) and Word Frequency (r = 0.274, $t(76) = 5.12$, $p < .001$) were the significant predictors in the lexical decision task; Word Frequency accounting for the greatest proportion of the variance. In the reading task the significant predictors were Imagery (r = 0.350, $t(74) = 7.67$, $p < .001$), Initial Bigram Versatility (r = 0.107, $t(74) = 2.33$, $p < .05$, Word Length (r = 0.103, $t,(74) = 2.25$, $p < .05$) and Word Frequency (r = 0.383, $t(74) = 8.08$, $p < .001$), Word Frequency again accounting for the greatest proportion of the variance.

In the analysis for the lexical decision performance of poor readers Imagery was a significant predictor (r = 0.207, $t(75) = 2.92$, $p < .01$); as was Word Length (r = 0.19111, $t(75) = 2.65$, $p < .02$),

and Word Frequency (r = 0.206, $t(75) = 2.86$, $p < .01$). The significant predictors in the reading task were Imagery (r = 0.362, $t(75) = 6.13$, $p < .001$) and Word Frequency (r = 0.448, $t(75) = 7.30$, $p < .001$).

SUMMARY OF RESULTS OF FITTING MODELS 1 AND 2

From the analyses of results so far, it appears that, apart from Word Frequency, Imagery accounts for the largest proportion of the variance in both lexical decision and reading tasks for the three reader groups - yet when variables were first examined individually in Model 1, which took into account word length and freqency, it was shown that many more orthographic variables were significant. It could be that in Model 2, which looked at the independent variables together, the very strong Imagery effect was partialling out the effect of other orthographic variables. Other orthographic variables may be significant in explaining a fair proportion of the variance, but when Imagery is fitted into the model the effects of these variables are partialled out, becoming non-significant.

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

In the final analysis, Imagery was removed from the model fitting so that other variables whose significant effect may have been masked by Imagery effects become more evident. Estimates, t values, significance levels and partial r for the 3 reader groups are presented in Appendix A4.11 to Appendix A4.15.

In the analysis for good readers Word Frequency ($r = 0.486$, $t(74) = 7.08$, $p < .001$), First Order of Approximation ($r = 0.178$, $t(74) = 2.94$, $p < .01$), Orthographic Neighbour

Ratio ($r = 0.688$, $t(74) = 4.0$, $p < .001$), Word Length ($r = 0.140$, $t(74) = 2.32$, $p < .05$) and Word Frequency ($r = 0.486$, $t(74) = 7.08$, $p < .001$) were significant in the lexical decision task. This analysis was inappropriate for the reading task, because no other orthographic variables were significant from the previous analysis.

Average readers also appeared to use orthographic information as measured by the Orthographic Neighbour Ratio when coming to a lexical decision ($r = 0.155$, $t(76) = 2.94$, $p < .01$), as well as Word Length ($r = 0.179$, $t(76) = 3.4$, $p < .01$) and Word Frequency ($r = 0.324$, $t(76) = 6.0$, $p < .001$).

In the reading task, Initial Bigram Frequency ($r = 0.104$, $t(75) = 2.29$, $p < .05$), Word Length ($r = 0.179$, $t(76) = 3.90$, $p < .01$) and Word Frequency ($r = 0.386$, $t(75) = 8.20$, $p < .001$) were the salient variables in predicting average reader's performance.

For poor readers, model fitting which did not include Imagery showed that the only other significant variables in the lexical decision task were Word Frequency ($r = 0.207$, $t(76) = 2.88$, $p < .01$), First Order of Approximation ($r = 0.156$, $t(76) = 2.17$, $p < .05$) and Word Length ($r = 0.246$, $t(76 = 3.40$, $p < .01$) - which accounted for the greatest proportion of the variance. However the model fitting for the reading task showed a highly significant effect of Word Frequency ($r = 0.419$, $t(76) = 7.00$, $p < .001$), and of Initial Bigram Frequency ($r = 0.153$, $t(76) = 2.66$, $p < .01$).

Comparing the results of Model 2 and the final model fitting - Model 3, it can be seen that the results showed that the very strong Imagery effect was indeed masking effects of orthographic regularity for both average and poor reader groups. The Orthographic Neighbour Ratio is significant variable in the lexical decision task, but not in the reading performance of good and average readers. It is not a significant predictor in the reading and lexical decision performance of poor readers, or so it appeared from the results of Model 2.

However comparing the results of Model 2 and the final model fitting, after Imagery was removed the Orthographic Neighbour Ratio accounted for a greater proportion of the variance than it did when Imagery was present in the model fitting, in the lexical decision task for good and average reader groups.

4.3(c) DISCUSSION

In the rationale for the three groups, different predictions were made as to the sort of orthographic information that would be used by children, according to reading ability. The hypothesis was that there would be a positive correlation between a word's initial and terminal letter frequency and versatility and reading accuracy for poor readers. With an increase in reading skill it was hypothesised that larger orthographic units, such as bigram frequencies and versatilities, and more global orthographic information, such as first and second order approximations to English would be used by average readers. Good readers, in Marsh et al's (1981) Stage 4 would use more complicated measures of orthographic structure, such as the Orthographic Neighbour Ratio, in the reading process.

It had been predicted that there would be a high correlation between measures of orthography and children's performance in lexical decision and reading tasks. The correlation analyses for the three reader groups did show that particular measures of orthography - First and Second Order Approximations to English and Terminal Bigram Versatility were significantly correlated with lexical decision performance and reading aloud.

An item analysis which took into account the effect of Frequency and Word Length showed that Imagery explained a greater proportion of variance for all reader groups in both types of task. The best predictors of performance from the large range of variables which measured orthographic structure were now seen, as predicted, to be the Orthographic Neighbour Ratio, First Order of approximation to English, Initial Bigram Versatility, and Initial Bigram Frequency. In these analyses ONR was seen to be significant in the lexical decision task, but not in the reading task confirming the experimental hypothesis that orthographic information would be more salient in the lexical decision task.

In the analysis which looked at significant independent variables after adjustment for word length and word frequency (Model 2), Imagery was shown to be a highly significant predictor variable in the reading and lexical decision performance for poor readers - it was the only variable of significance, with the exception of Word Frequency in the reading task. Of the number of variables which measured orthographic regularity, First Order

Approximation and ONR were important predictors of lexical decision performance; no orthographic variables reached significance in the reading performance of good readers.

For average readers, the Orthographic Neighbour Ratio was significant in lexical decision; Initial Bigram Vesatility was a significant variable in reading.

The final model fitting which excluded Imagery (Model 3), showed that there was an Orthographic Neighbour effect in the lexical decision of good and average readers. However it is not used by poorer readers, who used First Order Approximation to English in the lexical decision task. The effects of Imagery were of most importance in explaining the reading and lexical decision performance of poor reders.

Thus predictions made as to the orthographic information that would be used by children of differing ability was supported by the data collected. Probabilistic measures such as initial bigram frequency and versatility and First Order of approximation to English were used by the three reader groups. However, the good and average reader group showed an ability to use the Orthographic Neighbour ratio. It was considered that this complicated measure of orthographic structure would only be used by good readers; although the results show that it was used by average readers in the lexical decision task.

It seems important to investigate why it is that probabilistic measures of orthography are not more closely correlated with accuracy in reading aloud and lexical decision. There may be two reasons for this. Firstly summed or averaged frequencies can produce intuitively misleading estimates of orthography, especially where very high frequency clusters are involved.

For example, GLURCH and THXZQP have about the same average bigram frequency (according to the Underwood-Schulz (1960) combined count) because of the high frequency bigram TH in the letter string. Secondly raw frequencies of occurrence, as described by Massaro, Venezky & Taylor (1979) may not be as relevant psychologically as certain conditional probabilities or relative frequencies e.g., positional frequencies of Mayzner & Tresselt (1965). It may not matter how often a reader has seen a particular cluster if other, similar clusters are equally frequent. Partial visual information could trigger perception or report of any of the similar clusters with roughly equal likelihood; therefore despite their high frequencies, members of the set might not appear to show perceptual advantages. For example, the trigram THI is more frequent than the trigram QUE. However, QUE is the most frequent trigram beginning with QU, while THE is almost ten times more frequent than THI. Thus QUE might be reported with high accuracy, while THI, presented under identical conditions, might show many errors when it is reported as THE.

There is sufficient experimental evidence to argue that some kind of orthographic knowledge helps children to recognise letter strings. The results of Experiment 1 indicate that orthographic structure affects lexical decision accuracy and that this type of information is used by average and good eight year old readers. The clear evidence for the use of some kind of orthgraphic knowledge in lexical decision but not in reading aloud may be due to the difference in task demands. Barron (1981) suggested that orthographic information (specifically information about neighbourhood size) might be more likely to occur in lexical decision task, since it does not require that an item be given a specific overt response.

It may be that the reading task encourages good and average readers to use a phonological strategy (i.e. nonlexical process of assembled phonology) in preference to using orthographic information. In order to provide clear evidence for the use of orthographic regularity in children's reading, a younger age group could be used, some of whom were being taught using the phonics approach; some by 'look and say' - to see if there is a difference in their usage of orthographic information in a reading task.

Thus it appears that the orthographic neighbour ratio a previously unexplored variable, is important in predicting lexical decision accuracy. It takes into account the number of orthographic neighbours that a particular word has and their frequencies. It is felt that this orthographic variable may be very useful in the elaboration of Glushko's activation-synthesis approach, and other analogy theories of reading. This concept will be explored in a later experiment.

The following chapter explored further the role of Imagery in children's reading, as it was seen from this experiment to be an important predictor variable in reading, especilly for poor readers.

C H A P T E R   F I V E

THE ROLE OF WORD IMAGERY AND AGE-OF-ACQUISITION IN READING

5.1 INTRODUCTION

A well established fact in memory research is that concrete, high imagery words are easier to remember than abstract low imagery words. This phenomenon was demonstrated in a paper by Stoke (1929), who studied free recall in children aged 10 and 12 as well as in college students. Later, more systematic studies which controlled word frequency also found that concrete words were easier to remember in recognition memory tasks (Gorman, 1961), paired assocates learning (Yarmey and Paivio, 1965). Paivio, Yuille and Rogers (1969) demonstrated that word imagery has powerful effects on free recall and serial learning when word meaningfulness and word frequency are controlled. Thus word imagery is a strong determinant of memory performance in a range of tasks.

Another psychologically interesting property of words - the chronological age at which words are acquired - has been shown to affect retrieval from semantic memory. There are many sources of support for the importance of word age of acquisition (AOA) in the experimental literature showing strong effects of AOA on retrieval from the internal lexicon. Carroll and White (1973), in one of the first studies in this area, found that adults were faster at naming pictures when the picture names were learned early in life, and that the rated AOA measures were more important than objective frequency measures provided by the

Thorndike-Lorge and the Kučera and Francis word frequency counts. Subsequent studies have confirmed this finding (Lachman 1973; Lachman, Schaffer, and Hennrikus, 1974).

Carroll and White (1973) went on to collect AOA norms for 220 picturable nouns. Subjects had to estimate when in their life they probably first learnt a word and its meaning either in spoken or written form. Their instructions can be criticised because the order of acquisition of spoken words does not always relate directly to the order of acquisition of printed forms, since reading is usually begun when a substantial spoken vocabulary already exists. (However, Coltheart and Winograd (1986) noted that the validity of this procedure is supported by two validity studies carried out by Gilhooly and Gilhooly (1980). They found that adults' AOA ratings correlated .94 with the age at which children could offer a minimally satisfactory definition. In a second study, Gilhooly and Gilhooly (1980) found a correlation of .93 between adults' ratings of word AOA with the rank ordering, according to age norms, of words on a standard vocabulary test). Carroll and White's multiple regression analyses using the new AOA nouns confirmed their previous finding that AOA was a variable of greater relevance than word frequency in predicting speed of picture naming by adults. Their claims about the roles of AOA and word frequency rest in part on rather small differences between the respective correlations, but they have shown these differences to be replicable. It is possible that AOA ratings are to some extent affected by judgements of frequency, familiarity, or other word attributes.

Gilhooly and Gilhooly (1979) observed that AOA affected word retrieval when cued by initial bigrams i.e. early acquired target words were more likely to be produced than later acquired words, even when frequency and other word measures were taken into account. Their results added to the growing evidence that AOA was an important variable in lexical retrieval performance by adults. However, not all lexical retrieval tasks are affected by AOA, since Gilhooly and Logie (1982) found no effects of AOA on lexical decision times, although word frequency, length and familiarity emerged as significant factors.

Gilhooly and Watson (1981) reviewed research concerned with AOA effects in adult word naming and lexical decision, citing a number of studies which showed its effects on word naming latencies (Butler & Hains, 1979; Gilhooly and Logie, 1981b; and Rubin, 1980) but no effects on lexical decision latencies (Gilhooly and Logie, 1982; Whaley, 1978 and Rubin, 1980).

Coltheart and Winograd (1986) presented evidence from two experiments and from a reanalysis of data published by Christian, Bickley, Tarka and Clayton (1978) that word AOA did not affect the free recall or recognition of young adults, while Imagery did. Their experimental results were inconsistent with Morris's (1981) report, that late acquired words are better recalled than early acquired words. However a closer inspection of Morris's list showed that his results may be attributable to a difference in the emotionality value of his lists.

Brown and Watson (in press) performed two multiple regression studies; the first of which showed that word Age of Acquisition was the major predictor of rated familiarity; and the second of which showed that word AOA was a better predictor of word naming latency than spoken word frequency, written word frequency and rated familiarity.

It has been claimed that unambiguous AOA effects can only be found in tasks involving overt word naming (Gilhooly & Watson, 1981; and Watson, 1985) - on the basis of their reviews of the literature.

Brown and Watson suggested that the phonological output lexicon could be a possible source of AOA effects. This lexicon is phonologically organised and when access to it fails, word substitution errors occur which are phonologically related to the target (Fay and Cutler, 1977). Ease of mapping onto the output lexicon cannot be affected by target-word AOA, for if the early AOA words in the phonological output lexicon were more available, then speech errors of this type would be words that were learned earlier than the target words. Fay and Cutler (1977) showed that this was not the case, so the suggestion that early-learned words have more <u>accessible</u> representations in the phonological store, does not have experimental support. Brown and Watson considered the possibility that the <u>quality</u> of phonological information in the store might differ for early acquired words in their "completeness hypothesis". Some evidence for their hypothesis was provided by Aitchison and Straf (1982) in their study of children's malapropisms. They found that children, in their storage of word phonologies, accorded relatively less weight to the beginnings of words, which provide much information, suggesting that their relatively undeveloped phonological storage systems are less economical, and make less use of redundancy in word endings. If this assumption is correct, then a more complete phonological representation might be permanently available for early acquired words. When later acquired words must be pronounced, then, more time might have to be spent in generating phonological information not directly represented in the phonological lexicon. Thus the completeness hypothesis assumes that a common word ending will be represented in the phonological output lexicon by a single, more abstract symbol, which can be expanded via look-up procedures into a full phonological representation prior to further processing. Phonetic accommodation, for example, will take place at later stages. Words are stored in a relatively complete form by young children, and these representations remain complete throughout adult life.

Brown and Watson suggest that the effects of rated AOA in tasks that require overt word naming reflect the fact that the phonological output representations are stored in a relatively complete form during the early stages of vocabulary acquisition. As vocabulary size increases, storage limitations require more efficient strategies and only minimal information is stored explicitly. This is reflected in word naming latency, as the phonological information that is not directly represented has to be generated whenever required, which takes time.

As for Brown and Watson's finding that word AOA is the major predictor of familiarity ratings, it is possible that subjects might use the quality of phonological lexical representation available to them as an index of how frequently they have encountered that word. Assuming that subjects do not have direct access to frequency information, they may perform the familiarity ratings in a number of different ways. Some variable such as ease-of-predication might perhaps be the relevant variable, for Jones (1985) has shown that it is highly correlated with word AOA. His research is discussed in more detail later on in this chapter.

The findings outlined in this brief review of the literature raises the possibility that the previously reported imagery effects on various memory tasks may actually be the consequence of AOA and that these two variables may have been confounded in earlier research.

The effect of the AOA variable on children's reading is, as yet, unexplored. It is not known if there is an AOA effect in young children's reading, and, if there is, the strength of it. Secondly it is not known if children of differing reading ability will be differentially affected by AOA. Thirdly it is not known if the AOA effect is still present in skilled adult reading, although Gilhooly and Watson (1981) consider that it is.

The influence of word imagery on reading has only recently been studied. It appears to affect reading accuracy in certain types of acquired dyslexics. Marshall and Newcombe (1973) and others e.g. Coltheart, Patterson and

Marshall (1980) reported that deep dyslexics, noted for making semantic errors when reading aloud words presented in isolation, also find high imagery words easier to read than low imagery words.

Jorm (1977) studied the effects of imagery on children's reading accuracy. In three experiments on the effects of word imagery, length and frequency on reading difficulty, he found that high frequency words were easier to read for both good and poor readers. High imagery words were easier to read for poor readers only. Word length had little effect on reading difficulty for either good or poor readers. The differential effect of word imagery on reading difficulty for good and poor readers was interpreted in terms of the types of reading strategy used - phonics for good readers and whole word reading for poor readers. When children were forced to learn to read words by a whole word method, word imagery predicted ease of learning for both good and poor readers.

Jorm (1979a) has drawn attention to similarities between developmental dyslexia and acquired deep dyslexia. Baddeley, Ellis, Miles and Lewis (1982) compared the two syndromes by studying the reading performance of fifteen developmental dyslexics. One of their experiments showed that both developmental dyslexics and normal nine year old readers (the reading age matched controls) were less able to read low imagery words, a pattern of reading performance similar to that of deep dyslexic patients. Thus, the one feature of the performance of Baddeley et al's developmental dyslexics which appear to resemble deep dyslexics turned out to be a much more general phenomenon. As such, it does not argue for a common basis for deep and developmental dyslexia - the performance of the developmental dyslexics appeared qualitatively to resemble that of younger children.

Jorm (1979b) tried to interpret the imageability effect by suggesting that "word imagery affects the ease with which a word can be read via the direct visual route". A problem for Jorm's view is raised by the question of just why imageability should influence reading by the direct route rather than by the phonological route. Baddeley et al (1982) suggested that there may be a tendency for imageable words to be acquired earlier than low imageable abstract words. Frith (1985) also considers that the imageability effect may be attributable, to a large extent, to word age of acquisition.

Klose, Schwartz and Brown (1983) attempted to control age of acquisition and vary imagery in their investigation on reading by teenage boys. Their hypothesis was that the imageability effect was a function of the differing ages at which imageable and abstract words are acquired. Specifically, it was hypothesised that if the imageability effect was actually mediated by age of acquisition, then equating high - and low - imagery words for age of acquisition should eliminate the effect. They failed to find an imagery effect in two tasks. However, the first task required only visual matching and not lexical access. In the second task the subjects read aloud the previously presented words and their naming latencies were recorded.The problem with this task is that the earlier presentation of the items could have produced repetition priming effects which could have masked any imagery effects.

It is also difficult to evaluate Klose et al's study because no details of either chronological or reading ages of their subjects were given. Similarly, little detail was presented about the word lists used, apart from the fact that the words differed significantly on imagery but not on other characteristics.

Kroll and Merves (1986) looked at differences in lexical representation and processing for concrete and abstract words in the lexical decision task with normal, skilled adult readers. Previous research on memory and verbal learning (Paivio, 1971, 1978) and clinical neuro-psychology (eg. Coltheart, 1980; Patterson, 1981) suggested that the process of understanding and remembering concrete words (eg. words that refer to picturable objects and actions) differed from the process of understanding and remembering abstract words. They suggested a "dual-code" theory, according to which concrete words, like pictures of objects, could be coded in memory both verbally and imaginally, whereas abstract words could only be coded in a single verbal code.

Kroll and Merves's three experiments compared the speed and accuracy of lexical decisions for concrete and abstract nouns. In their first experiment separate groups of subjects judged each word type. In the second experiment all subjects judged mixed blocks of both word types. In both of these experiments there was a small speed advantage for concrete nouns in lexical decision. In their third experiment all subjects judged blocked presentations of each word type - the purpose of this experimental design was to observe transfer effects from one word type to another. In this condition there were large effects of word type on the speed of lexical decisions. When blocks of abstract words followed blocks of concrete words, judgements for the abstract words were significantly longer than those for concrete words. When concrete blocks followed abstract blocks, however, there was no difference in response time for the two word types. Kroll and Merves concluded that the effect of concreteness in lexical decision appears to be critically sensitive to order of presentation - and since they consider that concrete words can be coded in memory both visually and imaginally, this study appears to be of great relevance to any investigation of imagery effects.

The literature review has shown that the ease with which a word gives rise to an image is a powerful determinant of reading performance, both in normal reading and deep dyslexia. However, despite the fact that imagery is a potent variable, it has virtually no explanatory power. Jones (1985) tried to account for the apparent role of imagery in reading in terms of a variable termed ease of predication i.e. how easy or difficult it was to put a word into simple factual statements. His basic assumption was that there was some underlying semantic variable capable of being used to explain systematic differences both in the ease with which a word could be read in particular circumstances. Modern psychological theories of meaning have generally assumed that the elements representing a word in semantic memory were associated with a number of features or predicates (eg. Anderson, 1976; Anderson and Bower, 1973; Collins and Loftus, 1975; Jones, 1983b; Smith, Shoben and Rips, 1974). Jones considered that the explanation which could in principle account for both phenomena being related was the variability in ease of imagery and in ease of reading both reflected variability in the associated distributions of predicates for individual words. He predicted that there should be a close relation between ease of predication and imageability. He obtained ease-of-predication scores for a corpus of nouns, by asking subjects to rate on a seven point scale, how easy or difficult it was to put a word into simple factual statements eg. the word DOG would be very easy to make simple factual statements about - A DOG IS A TYPE OF ANIMAL, A DOG HAS FOUR LEGS, A DOG BARKS etc. As a contrasting example, the word IDEA would probably be judged as difficult to make simple factual statements about.

Jones asked thirty students to rate a set of 125 nouns, selected from the Paivio, Yuille and Madigan (1968) study, for ease of predication. He found that ease of predication scores were closely related to imageability scores. The ease of predication measure provided evidence in favour of the frequently voiced hypothesis (Anderson & Bower, 1973; Kieras, 1978; Shallice & Warrington, 1980) that apparent effects of imagery may be mediated via a previously unspecified variable with which it was very closely correlated. Jones then went on to show the usefulness of the ease of predication concept in accounting for the reading performance of deep dyslexic patients. However his paper is extremely useful because it provides a theoretical framework for the imagery effect in children's reading that was found Jorm (1977).Jones argued that deep dyslexic and children's reading reflected the operation of a semantic reading route and that access to a word's pronunciation was mediated by activation of semantic features (or predicates) in the semantic lexicon. Words which had few predicates were less likely to mediate any pronunciations. Thus, high imagery words were more likely to elicit a correct pronunciation.

Jones therefore attributes the imagery effects in children's reading to the use of direct visual access via a predicational (semantic) route to pronunciation, while Baddeley et al (1982) and Frith (1985) suggest that the imagery effect is more likely to be mediated by differences in word age of acquisition.

It appears that age of acquisition may be a powerful variable in the reading performance of young children, but that the role of imagery and age of acquisition has not conclusively been determined. Our experiments aimed to investigate the effects of these word properties more directly. In Experiment 2 children aged nine-to-ten years were presented with two reading tasks: One varied age of acquisition, the other varied imagery. The words presented were those used by Coltheart and Winograd (1986) with a few additional words. In Experiment 3 twenty skilled adult readers were presented with the two reading tasks, and naming latencies i.e. reaction time data were collected.

HYPOTHESES FOR EXPERIMENT 2

The literature review suggests that age of acquisition may well be an important variable in children's reading. Therefore the first hypothesis was that children's reading of early acquired words would be better than of late acquired words.

In view of the results of Experiment 1, where Imagery was a significant predictor variable in the reading performance of good, average and poor reader groups, it was hypothesised that there would be an Imagery effect shown by all reader groups, but more so for the poorer readers.

The third hypotheses considers whether children of differing reading ability will be differentially affected by these two variables. It is hypothesised that age of acquisition effects are very strong and will be shown by all reader groups. If imagery is to some extent mediated by age of acquisition, its effects may only be present in poor readers - as observed by Jorm (1977).

5.2(a) EXPERIMENT 2

METHOD

SUBJECTS

Sixty two children from two inner city schools acted as subjects. A class of eleven year olds (n = 15) were used in a pilot study and two parallel classes of nine-to-ten year olds (n = 47) were used in the main study.

STIMULUS MATERIALS

Task 1. Early and late Acquired Words

Forty words with age of acquisition ratings $\leq$ 3.50 and forty with age of acquisition ratings $\geq$ 4.50 were chosen from the Gilhooly and Logie (1980) nouns. These words were matched on word length, word frequency (Kucera and Francis, 1967) and imagery using the MRC Psycholingiustic Database (M.Coltheart). The mean values of the various measures are presented in Appendix A5.1 and Appendix A5.2.

Task 2. High and Low Imagery Words

Forty words with high imagery ratings $\geq$ 5.70 and forty words with low imagery ratings $\leq$ 3.50 were selected. As in the previous task these words were matched on other characteristics: Word length, word frequency and age of acquisition. The mean values are presented in Appendix A5.3 and Appendix A5.4

It was not possible to combine these lists in a single experiment because it would have been very difficult to control for frequency; there being a high inverse correlation between AOA and both Imagery and Frequency, as pointed out by Gilhooly and Watson (1981). Attempts to manipulate these variables would inevitably lead to small word lists.

APPARATUS AND PROCEDURE

The apparatus and procedure is similar to that described in Experiment 1. Words were presented one at a time in lower case print on a video display unit connected to a BBC-B micro computer which randomised the order of presentation separately for each subject. The subjects had to read each word aloud and the rate of stimulus presentation was subject paced.

PILOT STUDY

A pilot study was carried out, investigating the performance of 15 subjects on the Task 1 list. Their mean chronological age was 11:2 and their mean reading age was 10:8. This reading age is lower than expected because the children were performing at ceiling level for the SPAR reading test and 10:8 was the maximum reading age score, although the test is said to be applicable to children up to fifteen years of age, but only if they are less able readers.

RESULTS

Early acquired words were read more easily than late acquired words.

MEAN NUMBER CORRECT RESPONSES

EARLY AOA WORDS	LATE AOA WORDS
n = 40	n = 40
39.3	32.3

A Wilcoxon test showed that this age of acquisition effect was significant ($T = 0$, $P < .01$, one-tailed test). However it is clear from the mean number of correct responses that there is a ceiling effect for early acquired words. Thus, the main experiment used a younger age band instead, with two parallel classes of nine-to-ten year olds.

SUBJECTS

Two parallel classes of nine-to-ten year olds (n = 47) acted as subjects. Reading ability was assessed by means of the SPAR Reading Test (Young,1970), which was used to assign the children to three groups differing in reading ability. Good readers included children whose standardised score was 100, the average reader group consisted of those scoring between 91-100 and the poor reader group included children whose scores were between 71-90. This was done because a standardised score of 100 indicates that the chronological age and reading age are the same i.e. the child is an 'average' or better reader. In Experiment 1, class teacher ratings were used to assign childlren to reader groups. When comparing the reading scores obtained using the SPAR reading test with teacher ratings of ability there was a great similarity. Thus it was decided that reading test scores would be used in future experiments to assign children to reader groups. Chronological ages and reading ages for the three reader groups can be found in Table 5.1. The children were assigned to the three reader groups so that any developmental trend in the effects of Imagery and AOA would be clearly shown, and because the experimental hypotheses predict differential effects for the three reader groups.

TABLE 5.1

CHRONOLOGICAL AND READING AGES OF SUBJECTS

READER GROUP	N	MEAN CHRONOLOGICAL AGE	MEAN READING AGE
GOOD	13	9:9	10:5-10:6
AVERAGE	15	9:10	8:10
POOR	19	9:11	7:11

APPARATUS AND PROCEDURE

The apparatus and procedure was the same as that used in the pilot study. Half the subjects read the Task 1 words first and then the Task 2 words. The other half read Task 2 words followed by Task 1 words. This was done so as to prevent any order effects. Subjects' responses were tape-recorded. Verbatum instructions to children are presented in Appendix A5.5

5.2(b) **RESULTS**

EFFECTS OF AGE OF ACQUISITION

Means for correctly read words differing in age of acquisition are presented in Table 5.2

TABLE 5.2

MEAN CORRECTLY READ WORDS DIFFERING IN AGE OF ACQUISITION

	AGE OF ACQUISITION	
READER GROUP	EARLY	LATE
GOOD (N=13)	39.08	31.08
AVERAGE (N=15)	36.67	20.53
POOR (N=19)	24.37	11.16
MEAN	32.36	19.66

MAXIMUM SCORE = 40

An analysis of variance in which Reader Group was a between subjects factor and Age of Acquisition a within subjects factor was performed. This indicated a significant main effect of Reader Group (F (2,44) = 22.1, $P < .001$). Early acquired words were easier to read than late acquired words (F (1,44) = 182.56) and the interaction between Reader Group and Age of Acquisition was also significant (F (2,44) = 6.02, $P < .001$). The full anova table can be found in Appendix A5.6.

Planned comparisons indicated that good readers performed significantly better than average readers (t (44) = 2,29, $p < .05$) and that average readers performed significantly better than poor readers (t (44) = 4.21, $p < .001$). The interaction between Reader Group and Age of Acquisition arose because the Age of Acquisition effect was significantly smaller for good readers than for average readers (t (44) = 3.44, $p < .001$). However, although not as much affected by Age of Acquisition, good readers nonetheless read significantly fewer late acquired words than early acquired words (t (44) = 4.62, $p < .01$), and the interaction may simply have resulted from the ceiling effect on early acquired words for good readers.

EFFECTS OF IMAGERY

Means for correctly read words differing in Imagery may be found in Table 5.3.

TABLE 5.3: MEAN CORRECTLY READ WORDS DIFFERING IN IMAGERY

	IMAGERY	
READER GROUP	HIGH	LOW
GOOD (N=13)	35.31	35.15
AVERAGE (N=15)	29.73	28.13
POOR (N=19)	20.11	15.58
MEAN	27.38	25.0

MAXIMUM SCORE IS 40

An analysis of variance with Reader Group as between subjects factor and Imagery as a within subjects factor, indicated a significant effect of Reader Group (F (2,44) = 22.69, $p < .001$). Higher imagery words were easier to read than low imagery words (F (1,44) = 14.99, $p < .001$) and the interaction between Reader Group and Imagery was also significant (F (2,44) = 6.01, $p < .005$). The full anova table is presented in Appendix A5.7.

Planned comparisons showed that good readers read more words than average readers (t (44) = 2.24, $p < .05$) and that average readers' performance was superior to that of poor readers (t (44) = 4.33, $p < .001$). The interaction between Reader Group and Imagery occurred because the Imagery effect was significant only for the poor readers, t (44) = 5.39, $p < .001$.

DISCUSSION

The first hypothesis - that children's reading of early age of acquisition words would be better than late age of acquisition words was supported by the results. There was a strong age of acquisition effect in all reading ability levels.

The results also showed an imagery effect, which was not eliminated when high - and low - imagery words were controlled for age of acquisition, for the poor reader group. However good and average readers did not show an imagery effect when age of acquisition was controlled for. Thus children of differing reading ability were differentially affected by these two variables.

The question remains whether these effects are still present in skilled readers. The following experiment studied the naming latency in skilled readers in an attempt to answer this question. It was assumed that their performance would be virtually error free so that it would be possible to study the effects of AOA and Imagery on naming latencies.

5.3(a) EXPERIMENT 3 - THE EFFECTS OF WORD IMAGERY AND AGE OF ACQUISITION ON ADULT'S READING

SUBJECTS

Twenty students from the City of London Polytechnic volunteered to act as subjects. They ranged from eighteen year old first year students to postgraduate students.

STIMULUS MATERIALS

Task 1. Early and Late Acquired words.

Task 2. High and Low Imagery words.

These word lists were the same as those used for Experiment 1.

APPARATUS AND PROCEDURE

The apparatus and procedure were similar to that described in Experiment 2. Words were presented one at a time in lower case print on a Hantarex video display unit connected to a BBC-B micro computer which randomised the order of presentation separately for each subject. The subjects had to read each word aloud and the rate of stimulus presentation was subject paced. Half the subjects read the Task 1 one words first and then the Task 2 words. The other half read Task 2 followed by Task 1 words. Subject's responses were tape recorded, errors noted and naming latencies collected.

TREATMENT OF REACTION TIME DATA

Only reaction times for correct responses were used in the analysis. The error rate was very small - only 1% for each task. Any reaction times which were two standard deviations above or below the mean were removed. This was done to exclude any unusually high or low reaction times - sometimes caused when the voice key had been triggered too quickly eg. by a cough, or when it had not been triggered by a vocal response eg. when the subject's reply had not been loud enough to trigger the voice key. This happened very rarely - usually only one or two items were affected. Approximately 2% of the responses collected in Task 1 were discarded because of this; the same was true of Task 2.

5.3(b) <u>RESULTS</u>

EFFECTS OF AGE OF ACQUISITION

There was an age of acquisition effect shown in the adult reaction time data, (t (19) = 3.05, $P < .006$). Adults were able to read early acquired words more quickly than late acquired words - a group mean reaction time of 504.2 ms as compared to 519.4 ms for late age of acquisition words. Individual mean reaction times are presented in Appendix A5.8.

EFFECTS OF IMAGERY

There was no significant difference between the mean reaction times for high imagery and low imagery words (t (19) = 0.34 NS) - a group mean reaction time of 529.4 ms as compared to 527.2 ms for low imagery words. Individual mean reaction times are presented in Appendix A5.9.

ITEM ANALYSIS

The mean reading latency for each word was calculated and logistic modelling techniques were again applied for the Task 1 list. The predictors used were Imagery, Age of Acquisition, Kucera and Francis Word Frequency and Word Length (in order to control for its effect). The estimates, their t values, significance levels and partial r values are presented in Table 5.4. It can be seen that age of acquisition ($t(75)=1.72$, NS) and Imagery ($t(75)=0.10$, NS) are not significant predictors of reading performance. Word Frequency ($t(75)=3.07$, $p<.05$) is a significant predictor as is Word Length ($t(75)=2.70$, $p<.05$) - the partial r values showing that they make the most substantial contribution to reading difficulty, as measured by naming latency.

TABLE 5.4

LINEAR LOGISTIC MODELLING ANALYSIS OF READING DIFFICULTY OF WORDS (AS MEASURED BY NAMING LATENCY) VARYING IN IMAGERY

PREDICTOR	ESTIMATE	t(75)	p	partial r
AGE OF ACQUISITION	4.85	1.72*	NS	0.169
IMAGERY	0.87	0.13	NS	0.012
WORD FREQUENCY	-15.14	3.07	<.05	0.302
WORD LENGTH	6.61	2.70	<.05	0.265
PROPORTION OF TOTAL VARIANCE EXPLAINED i.e. R^2				0.193

* t needs to be 1.98 (df 120,)p<.05.

5.4 DISCUSSION

The results of Experiment 2 support the suggestion made by Baddeley et al (1982) and Frith (1985) that age of acquisition of words would affect children's reading performance. It appeared to have a very strong effect on the reading accuracy of nine-to-ten year olds at all the reading ability levels tested. The interaction between reader ability and age of acquisition was due to the near ceiling level performance on early acquired words by the best readers. However they still showed a substantial effect of age of acquisition, there being a mean difference of eight items between early and late acquired words.

It is surprising that such large effects on children's reading are produced by what Gilhooly and Gilhooly (1980) have themselves referred to as a rather implausible measure. They obtained age of acquisition values for words by asking adults to rate on a seven-point scale the age they thought they were when they acquired the word in their vocabulary. Despite the subjectivity and lack of precision of this procedure, the measures derived appear to predict both children's accuracy in reading as well as their knowledge of vocabulary as reported by Gilhooly and Gilhooly (1980). Brown and Watson (in press) recently suggested that early acquired words have a more complete representation in the phonological lexicon, and so are quicker to read. This is an interesting suggestion, though difficult to prove experimentally.

The suggestion that Imagery effects on reading accuracy are caused by differences in Age of Acquisition was examined in the second task. In this task it appeared that variations in Imagery produced a much smaller effect on reading accuracy (an overall mean difference of two words) than did variations in Age of Acquisition in Task 1. The Imagery effects appeared to be significant only for the poorer readers. These results confirm the earlier findings of Jorm (1977) who also obtained imagery effects with poor readers only.

The question remains as to why word imagery affects reading accuracy in poor young readers. It is possible that poorer readers who are likely to have lower levels of verbal ability generally, may have fewer low imagery/abstract words represented in semantic memory. Thus, low imagery words may lack entries in the phonological lexicon or they may lack entries in the cognitive system which represents word meanings. Alternatively, a word may be represented in both systems but may be inadequately or incompletely represented in the cognitive system. Evidence for the fact that low imagery words may lack entries in the phonological lexicon and/or cognitive system is provided by Vellutino and Scanlon (1985). They found inferior memory for low imagery words by poor readers in tasks using auditory presentation. From this and other evidence (Vellutino & Scanlon, in press) they concluded that poor readers are impaired in verbal encoding. In older poor readers (aged about eleven) they found impaired semantic encoding, younger poorer readers (aged seven) displayed impaired phonological encoding.

If the acquisition of word recognition units is dependent on the prior existence of a reasonably adequate entry in the cognitive system, then poor readers would be less likely to acquire word recognition units for low imagery words. Competent readers can often read words which they do not comprehend by successful application of grapheme - phoneme correspondence rules. Such a non-lexical procedure is far less successfully achieved by poor readers (Perfetti & Hogaboam, 1975).

These suggestions imply that the acquisition of word recognition units for beginning readers is determined by the pre-existence of entries in the cognitive system i.e. the assumption is that the direct visual access procedure necessarily involves the cognitive system. Evidence to the contrary would be provided by the observation that children can read words irregular in spelling-to-sound correspondence that they do not understand i.e. there is evidence of direct visual access with no involvement of the cognitive system. Such evidence has not yet been reported.

In Experiment 3 there was still a significant age of acquisition effect found in adults - a 15.2 ms difference. This is totally in accordance with the results of Experiment 2 where there was a strong age of acquisition effect in both poor and average readers performance on early and late acquired words, but the effect was significantly smaller for good readers, partly because of ceiling effects for early acquired words.

However, in the more rigorous item analysis there was no age of acquisition or imagery effect found in the adult naming latency data. This is also consistent with the results of Experiment 2 with children, where the imagery effect was only significant for the poor reader group.

Thus our findings suggest support for the supposition of Baddeley et al (1982) and Frith (1985) that the imagery effect is mediated by differences in word age of acquisition. The results of Task 2 show that the imagery effect is only significant for poor readers, as noted by Jorm (1977). The concept of "ease-of-predication", detailed by Jones

(1985) does provide a suitable framework for the imagery effect in poor readers. A high imagery word may well be represented by a complex cognitive network which links up the predicates for that word, while a low imagery word will have very few predicates. This explanation could in principle account for the role of imagery in the reading performance of poor readers - the variability in the ease of imagery and in the ease of reading both reflecting variability in the associated distribution of predicates for individual words.

In the light of our experiment on word age of acquisition and looking back to the experiment by Kroll and Merves (1986) in which they concluded that imagery effects were critically sensitive to order of presentation, it is clear that their highly concrete nouns are words acquired early in life i.e. there is a confounding with word age of acquisition, so unfortunately their results are no longer as unequivocal as they would first seem.

The puzzle that remains is why age of acquisition ratings are so reliable. It seems obvious that the age ratings by adults are by no means direct measures of a word's age of acquisition; it is unlikely that adults can remember exactly when each of the words to be rated were learned. Certainly our adult subjects found the task a difficult one. It seems more likely that the subjects make plausible estimates which are based on their general knowledge and intuitions about the child's use and knowledge of language. Thus, names of common objects, animal names etc are likely to be given low ratings because the raters assume that knowledge of the objects (and their labels) will have been available to children at an early age. However, ratings of less obvious types of words also appear to be reliable. This suggests that many people have a good, possibly tacit, knowledge of the pool of words understood by children at different ages. However, the fact remains that the age of acquisition variable does account for a large porportion of the variance of children's reading scores. What exactly underlies the measure remains for further studies to demonstrate.

C H A P T E R SIX

AN EVALUATION OF THE CONCEPT OF ORTHOGRAPHIC NEIGHBOURHOOD

6.1 INTRODUCTION

The concept of an orthographic neighbourhood is fundamental to analogy theory. Barron (1981) suggested the size of a word's orthographic neighbourhood affects how quickly a word would be responded to in lexical decision and reading tasks. Glushko (1979) explained the regularity effect in terms of exception words having inconsistent neighbours i.e. visually similar neighbours which were sounded differently. Henderson's (1982) Lexical pooling model included classes of words (i.e. hermits and heretics) that might be expected to have special properties due to the nature of their orthographic neighbourhood.

The results of Experiment 1 showed that orthographic structure (as measured by the Orthographic Neighbour Ratio) is an important variable in the lexical decision process of average and good eight year old readers, but not in the reading process. Experiment 4 explores further the concept that orthographic neighbours can affect children's reading and lexical decision performance. A different measure of orthography was used - the "N" of a letter string, which was defined as the number of different English words that can be produced by changing just one of the letters in the string to another letter, preserving letter positions. This is a more general measure of orthography as it only takes into account the number of neighbours of a letter string, and not their frequency. N values can also be calculated for nonwords so that the effects of neighbours can also be studied in nonword reading.

The measure was devised by Coltheart, Davelaar, Jonasson and Besner (1977). In their experiment eighteen skilled readers were asked to perform a lexical decision task on a list of randomly mixed high-N and low-N words and nonwords. The words and nonwords in the list were matched pairwise for Kucera and Francis frequency and number of letters, i.e. for any high-N word there was a corresponding low-N word with the same frequency and number of letters. Coltheart et al (1977) found high-N nonwords produced slower reaction times than low-N nonwords, but that there was no N effect with words. They suggested a dual route model with a flexible deadline to explain these results. According to this model each word in the internal lexicon has its own logogen, an evidence collecting device that is excited by a letter string to varying degrees) but with a judgement deadline value which can be adjusted up or down during processing i.e. at stimulus onset, information begins to flow into the logogen system. Excitation rises at various rates in various logogens. If the overall amount of excitation is rising rather rapidly, the value of the deadline is increased considerably, since in these circumstances, the letter string is probably a word, so NO is probably the wrong response; if the logogen system is fairly quiet, the deadline value is not increased, since NO is likely to be the correct response.

If such adjustments are constantly being made as information from the letter string is flowing into the logogen system, illegal nonwords, which are very unlike words and so do not excite logogens much, will have a short deadline and so a fast NO compared to legal nonwords, which are more wordlike and so excite logogens to a greater degree. Furthermore, among legal nonwords those which are like many words will excite many logogens, while those which are like few words will excite few. Therefore, the deadline time, and hence the NO latency, will be longer for high-N nonwords than for low-N nonwords.

However Coltheart et al (1977) made differing predictions about the effect of N on YES latencies. The YES response is triggered when some logogen has actually reached threshold. A principal determinant of the time which elapses between the stimulus onset and the reaching of a threshold is word frequency, since the amount of excitation required for the threshold to be attained is inversely related to word frequency. This time will not be influenced by whether many other logogens are partially excited (as is the case of a high-N word). Therefore Coltheart et al hypothesised that the time required to respond YES to a word will be strongly affected by the word's frequency, but not affected at all by whether it is a high-N or a low-N word. Their model correctly predicted the pattern of results in a lexical decision experiment reported in their 1977 paper.

Thus, the effect of the number of orthographic neighbours on (word) lexical decision is questioned by dual route theorists. No explicit hypothesis about the N effect on the reading process was advanced. This is not the case for researchers who are adherents of analogy theory such as Glushko (1979) and Barron (1981). Barron suggested that a word with a large neighbourhood (e.g. FILL) would be responded to more rapidly than one with a small neighbourhood (e.g. DEBT). This possibility involves the assumption that the speed with which a word can be activated is based on some combination of the activation values of the item and its corresponding neighbours. This hypothesised neighbourhood-size effect would occur even if visual information were the only information activated. Furthermore, the effect might be more likely to occur in a lexical decision task, since that task does not require that an item be given a specific, overt response. Barron calculated the number of neighbours and summed log frequency of the neighbours for each item in a word list used in a lexical decision task. He found that the averaged values of the summed log frequencies was significantly higher for regular than exception words. However some of the neighbours of the regular items were very high frequency, so Barron's results remain inconclusive.

Glushko (1979) explained the tendency for exception words to take longer to name as due to their tendency to have visually similar neighbours that are sounded differently. For example, in analogy theory the difficulty in the exception word BREAST is that it has a large number of inconsistent neighbours, such as BEAST, FEAST, LEAST, and the pronunciations of these neighbouring words influences the decision about BREAST. An important point is that regularity is not the true causal agent of the latency effects according to analogy theory - it is the correlation of regularity with consistency of pronunciation among a group of visually similar words that is responsible for the effect, Glushko (1979) suggested that, in the reading process, there is lexical pooling - information from the various sources is entered into a common decision process and synthesised. He assumed that decisions take longer when they depend upon the integration of information about portions of the word and also when the various sources provide inconsistent information. The most interesting prediction that can be derived from these assumptions is that exception words will be slower to read than regular words because of conflicting sources of information. The lexically stored pronunciation conflicts with the conventional product of orthographic translation. Of greater interest is Glushko's extension of the prediction to regular and exception pseudowords. The fact that exception pseudowords such as HEAF were more subject to exceptional pronunciation than comparable regular pseudowords offered strong support to the idea that lexical analogy has reality as a source of information on pronunciation. The analogy hypothesis itself only predicts the nature of pronunciations. The existence of a latency effect favouring regular pseudowords requires that the analogy theory be elaborated. This elaboration was provided by Glushko's lexical pooling concept, together with the assumption that inconsistent neighbours slow the decision process.

Glushko (1979) found that although most exception words, when mispronounced, were regularised (eg. TOMB as in BOMB), some were mis- pronounced in accord with an even more irregular analogy (eg. TOMB as in COMB). Generalising from this example, Glushko suggested that a visual pattern, when presented for naming, activates an orthographic neighbourhood of closely resembling words. It is probably reasonable to assume that the strength of activation is proportional to the similarity of the presented pattern. When the stimulus is a familiar word, the lexical entries activated will include that for the word itself, as well as various analogies which may be consistent or inconsistent. When the stimulus is a pseudoword only real word analogies will be activated.

Henderson (1982) extended Glushko's ideas with a lexical pooling model which included classes of words that might be expected to have special properties in lexical pooling theory due to the nature of their orthographic neighbourhood. According to this model, when a letter string is presented it activates an orthographic neighbourhood of words which differ from the target by one letter in the same position. He distinguished between two classes - hermits and heretics. Hermits eg. LYNX and NERP have no close orthographic neighbours. He suggests that they are disadvantaged firstly at the activation stage, where it takes longer for activation to spread to remote analogies (MANX, LYMPH etc for the word LYNX); and secondly at the assembly stage where small segments - such as used in the GPC procedure would be used.

Henderson (1982) reported finding large effects for lexical hermits in unpublished studies but the lists were confounded by other variables, as hermits often tended to have low positional and sequential frequencies. He did not specify whether these effects were for reading or lexical decision, Henderson (1982) p.159.

Lexical heretics are described by Henderson as words like HAVE which are very inconsistent. They are at odds with almost all their orthographic neighbours. HAVE possesses many close neighbours which are united "in a hostile orthodoxy" eg. CAVE, SAVE, RAVE etc. His definition of neighbours is not just confined to words sharing the terminal segment, as is Glushko's, but includes words sharing beginnings as well as endings.

In summary, it can be seen that the theoretical approaches described make different predictions for the effect of orthographic neighbourhood either for lexical decision or reading. (See TABLE 6.1) It is inappropriate to have a straightforward pitting of one theory against another because they are often similar in their predictions, and also because their definitions of what constitutes a neighbour differ. Some theories make predictions about lexical decision latency, others make predictions about reading latency. The summary table provides a general indication of which theories consider that orthographic neighbourhood is an important variable in the reading process and in the lexical decision task.

If knowledge of letter to sound correspondences build up as a result of the child's experiences (Patterson & Morton, 1985; Francis, 1984) we would predict that words involving more frequent correspondences would be read more easily than others. These correspondences could be defined

TABLE 6.1

PREDICTIONS MADE BY VARIOUS MODELS

COLTHEART et al's (1977) DUAL ROUTE MODEL WITH FLEXIBLE DEADLINE predictions for lexical decision latency only	No effect of N on words because frequency is more important. High frequency words have low threshold, low frequency words have high thresholds exciting logogens. Effect of N on nonwords - high-N nonwords have a slower RT because similar to many lexical entries, low-N nonwords have a faster RT because similar to few lexical entries, and have a shorter deadline
GLUSHKO (1979) predictions for reading latency only	N limited to words sharing terminal segment. Number of neighbours is important in terms of consistency only. Words & NW have fast RT if they have many consistent neighbours, slow if they have many inconsistent neighbours
HENDERSON (1982) predictions not specific as to task	Hermits with no close orthographic neighbours will be disadvantaged - true for both words and nonwords. (i.e. LOW-N). N.B. His idea of neighbourliness is not limited to words sharing the terminal segment

according to GPC rules or larger segments such as bodies. However this assumes that children segment in a particular way e.g. segmenting HEAD into H + EAD.

N is of considerable value as an experimental variable because it provides a measure of frequent correspondences, which makes no assumptions about the nature of segmentation and size of segment and it provides a very rough measure of the word-likeness of a word - so it should be a strong predictor of reading skill whether one considers reading experience rather than learning abstract GPC rules as important for reading, or whether one considers N from an analogy theory approach.

There has been much less evidence for the use of orthographic structure in spelling. Campbell (1983) found that only very good young spellers could use grapheme-phoneme correspondence rules to spell nonwords. Poorer spellers used sub-word segments.

Frith (1985) in her elaboration of Marsh, Friedman, Welch and Desberg (1981) predicted different strategies that would be used in the spelling of an unfamiliar word. If a child were still using the alphabetic strategy, then the word could not be written at all. Use of the alphabetic strategy in basic form would mean that individual phonemes are converted into individual graphemes and strung together e.g. "gust" for JUST. Use of the orthographic strategy would mean that instantly segmented units were matched with internally represented units, and so unfamiliar words would be spelled by using components of familiar words e.g. "inteligents" for intelligence. As a result a good speller can spell nonsense words apparently "in analogy" with existing words, regardless of whether a regular or irregular spelling is present.

Frith predicted that different types of otherwise phonetically acceptable spelling errors would occur at the different phases of acquisition. For example, the unfamiliar word TRACTION might be spelled "trackshen" at Phase 2 - the orthographic stage, however, instant segmentation would take place, so that the first segment of the word could be spelled with regard to an exiting unit TRACK, and the ending could be spelled as a known morpheme such as "SION". Thus the effect of orthographic neighbourhood size on spelling was also investigated and children were presented the lists described earlier in a spelling task.

The effect of orthographic neighbourhood size on reading, lexical decision and spelling were investigated. These tasks were used because although neighbourhood size has been shown to have an effect on nonwords but not on words in a lexical decision task, analogy theorists have hypothesised that it will have an effect in the reading task. There is less evidence for the use of orthographic neighbourhood size in spelling as compared to reading. It is hypothesised that there will be an N effect on the three experimental tasks, being more marked for lexical decision. Children aged eight to ten years of age were used because good readers of eight and nine in Experiment 1 had been shown to use orthographic knowledge.

6.2(a) EXPERIMENT 4 - METHOD

SUBJECTS

Pupils in the second (n = 23) and third year (n = 26) classes of an inner city school acted as subjects. The mean chronological ages were 8:10 and 9:1 for the two classes, ranging between 8:5 and 10:5. Reading and spelling ability was assessed by means of the SPAR standardised group tests of reading and spelling (Young 1976). There was very little difference in the mean reading ages of the two classes (2nd year = 9:1, 3rd year = 8:10) so they were treated as one group and subdivided into reader groups on the basis of the raw Spar reading scores. Raw Spar scores were used rather than standardised scores, because standardised scores are more appropriate for children of similar age groups, and in this experiment there is a wide age range (8:5-10:5). The range of raw scores was subdivided into three. This yielded rather uneven sized groups, but other methods of subdividing would hve been totally arbitrary - dividing the subjects into three equal sized groups would have meant that poorer readers would have a wide range of ability i.e. not be a homogeneous group. Thus good readers included children whose raw score was greater than 36, average readers consisted of those scoring between 27-36 and poor readers included the remaining children whose scores were between 14-26. Mean reading, spelling and chronological ages for the three reader groups are presented in Table 6.2. Subjects were subdivided into three reader groups to see if there was a detectable trend, as outlined by Marsh, Friedman, Welch and Desberg (1981) which predicted that young unskilled readers are unaffected by orthographic knowledge, but that as children become increasingly skilled they become more reliant on orthographic knowledge.

TABLE 6.2

SAMPLE CHARACTERISTICS : MEANS FOR CHRONOLOGICAL AGE AND PERFORMANCE ON STANDARDISED TESTS OF READING AND SPELLING ACCORDING TO READER GROUP

	READER GROUP		
	GOOD	AVERAGE	POOR
CHRONOLOGICAL AGE	9:7	9:5	9:6
RAW READING SCORE	40.19	31.45	28.88
READING AGE	10:1	8:6	7:6
NUMBER OF CHILDREN	21	20	8
RAW SPELLING SCORE	25.1	21	15.67
SPELLING AGE	10:8-10:9	9:5-9:6	8:4
NUMBER OF CHILDREN	19	18	8

STIMULUS MATERIALS

A list of 39 high-N and 39 low-N words were used (N being the number of orthographic neighbours). These were matched pairwise for Kucera and Francis frequency and number of letters, i.e. for any high-N word there was a corresponding low-N word with the same frequency and same number of letters. There were also 40 high-N nonwords and 40 low-N nonwords, with pairwise matching for number of letters (See Appendix A6.1). All nonwords were pronounceable. These stimulus items were those used by Coltheart, Davelaar, Jonasson and Besner (1977).

APPARATUS AND PROCEDURE

The children were tested individually in a quiet room. An initial practice session preceded the main experiments to ensure that the child understood the tasks.

EXPERIMENT 4a LEXICAL DECISION

The 78 words and 80 nonwords were randomised and presented one at a time in lower case print on a video display unit connected to a BBC-B microcomputer. The subjects had to make a lexical decision about what had just been presented i.e. they were instructed to make a NO response if the item was not a word and a YES response if it was a word. A voice key triggered by the onset of vocalisation caused the timer to stop and reaction times and the correctness of the response were recorded. Presentation was self-paced i.e. the child pressed the space bar to make a new item appear.

EXPERIMENT 4b READING

The children were tested individually on a later occasion. The procedure was the same as for the lexical decision task - the micro- computer randomising order of presentation separately for each subject. In this instance the children had to read each item aloud. They were told that half the items were nonsense words, and encouraged to read them as best they could. Verbatum instructions can be found in Appendix A6.2. Responses were tape recorded and naming latencies were recorded by the computer.

EXPERIMENT 4c SPELLING

The 158 items were given as a group spelling test a month later, words being presented before nonwords. If a word was ambiguous or homophonic it was put into a sentence so as to make the meaning clear. The possibility that blocked presentation might have affected results was considered and so these items were administered to a second sample of twenty four children in another inner city school. Two different mixed random orders were constructed for words and nonwords. Error scores for each word were calculated for both samples and the resulting correlation was .922, showing that the blocking of similar items for the main sample did not appear to affect the spelling performance in any way.

Thus for low and high N words and nonwords data was collected for lexical decision performance (n=49), reading performance (n=46) and spelling (n=45) out of a total of 49 subjects - due to children being absent on the day of test.

6.2(b) RESULTS

SUBJECT ANALYSIS

Mean accuracy scores on the three experimental tasks are given in subsequent tables task by task. Analyses of variance were carried out in which Reader Group was a between subjects factor and N values a within subjects factor for each of the three experimental tasks.

The scoring of nonwords for both the naming and spelling of these items proved a very difficult task. There is no wrong way of saying or spelling a nonword, apart from a straightforward error e.g. when the child lexicalises a nonword. If a subject uses the irregular form of pronunciation, how does one decide which subject is correct? For the reading task responses were scored correct only when the pronunciation conformed to Wijk's (1966) classification of the regular form. For the spelling task various alternative forms were acceptable for nonwords.

e.g. KLUN "cloon", "klune", "kloon".

LEXICAL DECISION TASK

The mean correct lexical decisions for words and nonwords in total are set out in Table 6.3.

TABLE 6.3

MEAN TOTAL CORRECT LEXICAL DECISIONS FOR WORDS AND NONWORDS DIFFERING IN N

READER GROUP	HIGH N	LOW N
GOOD (N = 21)	63.2	65.8
AVERAGE (N = 20)	59.9	58.9
POOR (N = 8)	48.8	46.4
N = 49		
OVERALL MEAN	59.5	59.8

Maximum score = 79

This analysis examined the number of correct responses to the 78 stimuli i.e. Y(W) + N(NW). There was a significant main effect of Reader Group ($F(2,46) = 17.58$, $p < .001$), but no N effect or interaction was found. Planned comparisons showed that good readers performed significantly better than average readers ($t(46) = 2.38$, $p < .021$) but that there was no difference in the performance of the poor and average reader groups, so when word and nonword performance was analysed together, there was no N effect. (A full anova table is presented in Table 6.3a)

TABLE 6.3a

ANALYSIS OF VARIANCE TABLE FOR LEXICAL DECISION FOR WORDS AND NONWORDS Y(W) + N(NW)

	SS	df	MSe	f	signif
READER	3328.69	2	1664.34	17.58	< .001
n	1.82	1	1.82	.16	NS
READER x n	98.74	2	49.37	4.28	< .05
MSE_b	531.03	46	11.54		

Contrasts

Good	vs Average	t = 2.38	df = 46	p < .021
Average	vs Poor	NS		

This analysis was done so as to check if there was any response bias shown by subjects. There was no response bias found i.e. readers did not have a preference for saying YES all the time.

Table 6.4 shows the means for lexical decision responses to words and nonwords separately. For words there was a significant Reader Group effect ($F(2,46) = 15.5$, $p < .001$), there was a highly significant N effect ($F(1,46) = 86.6$, $p < .001$) indicating that all reader groups found high-N words easier to accept as words than low-N words. The interaction of N with reader group was also significant ($F(2,46) = 6.6$, $p < .003$). This was due to the fact that good readers were significantly less affected by N than average readers ($t(46) = 2.11$, $p < .04$).

For nonwords a significant Reader Group effect was found ($F(2,46) = 7.8$, $p < .001$). There was a significant N effect ($F(1,46) = 30.3$, $p < .001$) indicating that low-N

TABLE 6.4

MEAN CORRECT RESPONSES TO WORDS AND NONWORDS DIFFERING IN N LEXICAL DECISION TASK

READER GROUP	WORDS		NONWORDS	
	HIGH N	LOW N	HIGH N	LOW N
GOOD (N = 21)	34.7	32.7	28.6	33.1
AVERAGE (N = 20)	31.5	27.7	28.5	31.4
POOR (N = 8)	29.5	23.6	19.3	22.8
N = 49				
OVERALL MEAN	32.5	29.2	27.0	30.7
	Maximum score = 39		Maximum score = 40	

TABLE 6.4a

ANALYSIS OF VARIANCE TABLES LEXICAL DECISION:- WORDS AND NONWORDS ANALYSED SEPARATELY

LEXICAL DECISION FOR WORDS (NUMBER CORRECT)

	SS	df	MSe	f	signif
READER	692.34	2	346.17	15.47	$< .001$
n	303.52	1	303.52	86.55	$< .002$
READER x n	46.24	3	23.12	6.59	$< .003$
MSE_b	161.31	46	3.52		

Contrasts

Good vs Average $t = 21$ $df = 46$ $p < .04$

Average vs Poor NS

LEXICAL DECISION FOR NONWORDS (NUMBER CORRECT)

	SS	df	MSe	f	signif
READER	1197.39	2	598.70	7.77	$< .001$
n	270.47	1	270.47	30.34	$< .001$
READER x n	12.91	2	6.45	0.74	NS
MSE_b	410.10	46	8.92		

nonwords were significantly easier to reject than high-N nonwords. There was no significant interaction between N and reader group. Full anova tables can be found in Table 6.4a and a plot of the means are shown in Appendix A6.3.

Response latencies were analysed for words, but not for nonwords in view of the variability of the nonword data. Mean correct lexical decision latencies are set out in Table 6.5.

TABLE 6.5

MEAN CORRECT LEXICAL DECISION LATENCIES OF WORDS DIFFERING IN N

	WORDS	
READER GROUP	LOW N	HIGH N
GOOD (N = 21)	1106.5	1093.0
AVERAGE (N = 20)	1361.1	1276.6
POOR (N = 8)	1390.4	1195.9
N = 49	1256.8	1184.7

An analysis of variance showed a nonsignificant effect of reader group, but there was still a significant N effect - lexical decision latencies were longer for low-N words than for high-N words ($F(2,46) = 11.45$, $p < .001$). A full anova table is presented in Table 6.5a.

TABLE 6.5a

ANALYSIS OF VARIANCE TABLE FOR LEXICAL DECISION LATENCIES FOR WORDS

	SS	df	MS	f	signif
READER	1083806.0	2	541903.0	2.48	NS
N	192353.1	1	192353.1	11.45	< .001
READER x N	97470.9	2	48735.5	2.9	NS
MSE	772835.0	46	16800.8		

READING TASK

Means for correctly read words and nonwords differing in N are set out in Table 6.6.

TABLE 6.6.

MEAN CORRECT READING OF WORDS AND NONWORDS DIFFERING IN N

READER GROUP		WORDS		NON WORDS	
		HIGH N	LOW N	HIGH N	LOW N
GOOD	(N = 19)	37.3	35.6	29.6	25.2
AVERAGE	(N = 18)	35.1	30.7	25.2	20.2
POOR	(N = 9)	20.1	15.0	9.9	7.6
	N = 46				
OVERALL MEAN		33.5	30.1	24.4	20.2
		Maximum score = 39		Maximum score = 40	

TABLE 6.6a

READER GROUP MAIN EFFECTS AND CONTRASTS FOR THE READING TASK

READER GROUP EFFECT	WORD	NONWORD
F	38.18	19.95
MS_e	52.48	99.33
df	2,43	2,43
p	$<.001$	$<.001$
Contrasts		
1 Good - Average		
t	2.15	2.04
df	43	43
p	$<.05$	$<.05$
2 Average - Poor		
t	6.88	4.64
df	43	43
p	$<.001$	$<.001$

An analysis of variance was performed using Reader Group as a between subjects factor, and N value (high v low) as a within subjects factor. In the interests of brevity F values, MS_E and contrasts for overall differences between the reader groups are set out in Table 6.6a.

It can be seen that the three reader groups differed significantly from each other on the reading task.

Low N words were significantly harder to read than high N words ($F(1,43) = 76.44$, $MSe = 3.98$, $p < .001$). The interaction of N with reader group ($F(2,43) = 7.51$, $p < .005$) was because good readers were significantly less affected by N than average readers ($t(43) = 2.93$, $p < .005$), and average and poor readers did not differ.

The results of planned comparisons indicated that there was an N effect for all three reader groups; for good readers ($t(43) = 2.62$, $p < .02$), average readers ($t(43) = 6.60$, $p < .001$) and for poor readers ($t(43) = 5.00$, $p < .001$).

Low N nonwords were significantly harder to pronounce than high N nonwords ($F(2,43) = 40.03$, $p < .001$). The interaction of N with reader group was not significant.

SPELLING TASK

Means for correctly spelt words and nonwords differing in N are presented in Table 6.7.

TABLE 6.7

MEAN CORRECT SPELLING OF WORDS AND NONWORDS DIFFERING IN N

READER GROUP		HIGH N	WORDS LOW N	diff
GOOD	(N = 20)	35.1	27.3	8.8
AVERAGE	(N = 17)	31.1	20.7	10.4
POOR	(N = 8)	15.0	5.5	9.5
	n = 45			
OVERALL MEAN		30.2	21.1	
		Maximum score = 39		

READER GROUP		HIGH N	NONWORDS LOW N	diff
GOOD	(N = 20)	26.8	22.0	4.8
AVERAGE	(N = 17)	22.1	14.2	7.9
POOR	(N = 8)	10.8	6.6	4.2
	N = 45			
OVERALL MEAN		22.0	16.5	
		Maximum score = 40		

The F values and contrasts for overall differences between the reader groups on the spelling task is set out in Table 6.7a.

TABLE 6.7a

READER GROUP MAIN EFFECTS AND CONTRASTS FOR THE SPELLING TASK

	READER GROUP EFFECT	WORD	NONWORD
	F	37.03	9.06
	MS_E	68.97	155.97
	df	2,43	2,43
	p	$< .001$	$< .001$
	Contrasts		
1	Good-Average		
	t	2.78	2.09
	df	43	43
	p	$< .01$	$< .04$
2	Average-Poor		
	t	6.22	2.49
	df	43	43
	p	$< .001$	$< .02$

Low-N words were significantly harder to spell than high-N words ($F(1,43) = 179.06$, $MS_E = 9.27$, $p < .001$). The interaction of N and reader group was not significant. Low-N nonwords were significantly harder to spell than high-N nonwords. ($F(1,43) = 51.32$, $MS_E = 11.05$, $p < .001$). The interaction between N and reader group was significant ($F(2,43) = 3.34$, $p < .05$). The results of planned comparisons indicated that low N nonwords were significantly harder than high N nonwords for all groups - for good readers ($t(43)$ =

4.49, $p < .001$), for average readers ($t(43) = 6.37$, $p < .001$) and also for poor readers ($t(43) = 2.48$, $p < .02$). The interaction between N and Reader Group was significant because good readers were significantly less affected by N than average readers ($t(43) = 2.41$, $p < .02$), whereas average and poor readers did not differ significantly.

The results of Experiment 4 show that there is an effect of the size of orthographic neighbourhood for children in a variety of tasks - words with large orthographic neighbourhoods (i.e. high-N words) are easier to read, spell and accept as words than are low-N words, which have very small orthographic neighbourhoods. For nonwords, the pattern of results was slightly different. Although a large orthographic neighbourhood of words makes a high-N nonword easier to read and spell, it makes the lexical decision task - where a child has to decide if the nonword stimulus is a word or not, much more difficult. It was easier for children to reject low-N nonwords.

There also appeared to be a trend in terms of reading competence. Poor and average reader groups were using this orthographic information, but the N effect was less pronounced in the lexical decision and reading performance of the good reader group.

DISCUSSION

The results of the subject analysis yielded highly significant effects of N for words and nonwords indicating that orthographic structure is an important variable in

TABLE 6.8

MEAN RATINGS OF AGE OF ACQUISITION FOR ADDITIONAL EARLY AOA AND LATE AOA WORDS

			EXPERIMENTAL RATINGS		GILHOOLY & LOGIE (1980) RATINGS
	ball		1.7		1.50
	fork		1.85		2.25
	doll		1.675		1.61
EARLY	kick		2.15		2.28
	sand		2.125		2.17
AOA	angel		2.475		2.42
	band		2.975		2.36
WORDS	fairy		2.175		2.42
	uncle		1.95		1.92
	water		1.525		1.53
	paper		2.15		2.29
		$\bar{x}$	2.07	$\bar{x}$	2.07
		S.D	0.41	S.D	0.36
	ague		6.325		6.61
	cowl		6.0		5.92
LATE	jade		4.95		5.72
	kale		6.375		5.75
AOA	seer		6.05		6.11
	chili		5.425		6.06
WORDS	dowry		5.7		5.56
	ether		5.575		6.25
	flora		5.725		5.69
	irony		6.15		6.06
	thong		5.05		5.53
		$\bar{x}$	5.76	$\bar{x}$	5.95
		S.D	0.48	S.D	0.33

$r = 0.984$

referred to in Experiment 6

these children's reading. It was shown that items with few neighbours were significantly more difficult than those with many neighbours. The results are consistent with Glushko's (1979) prediction that words with few neighbours are more difficult to read, although he emphasised that the consistency of neighbours was an important factor. They are also consistent with Henderson's (1982) lexical pooling model, which predicted that hermits with no close orthographic neighbours would be disadvantaged; true for both words and nonwords. It may be that there is some developmental trend whereby this sort of knowledge of orthographic structure is more important at a particular stage. There was a consistent pattern showing that good readers were significantly less affected by N. Highly skilled adult readers may only ever require this orthographic information when dealing with unfamiliar reading material i.e. nonwords; this suggestion could be followed up in future experiments.

The results also showed an increasing N effect with increased spelling skills providing some support for Frith's as yet untested model, in which orthographic information is used by increasing skilled spellers. A further area of research of research would be an error analysis of the spelling data, in which spelling errors would be classified according to the extent to which they preserve morphemes or small orthographic units (ght, ou, etc). Longitudinal studies of individual children would certainly provide tests of the usefulness of Frith's present model.

The following experiment investigates whether there is still an N effect shown in the reading performance of skilled adult readers, as measured by naming latency in a reading task. This was done because the results of Experiment 4 showed that good readers were significantly less affected by N. Coltheart et al (1977) found an N effect for skilled readers, but only for nonwords in the lexical decision task. It seemed important to establish whether or not size of orthographic neighbourhood can affect word naming in adult skilled readers.

6.3(a) EXPERIMENT 5

METHOD

SUBJECTS

The subjects were 30 students at the City of London Polytechnic. They ranged from 18 year old first year students to postgraduate students.

STIMULUS MATERIALS

The words were the thirty nine high-N words and thirty nine low-N words used in Experiment 4.

APPARATUS AND PROCEDURE

Each subject was tested individually in a quiet room. An initial practice session preceded the main experiment to ensure that the task requirements were understood.

The 78 words were randomised and presented one at a time in lower case print on a video display unit connected to a BBC-B microcomputer. Subjects had to read each word aloud. Responses were taped and naming latencies were recorded by the computer.

6.3(b) RESULTS

Naming latency data was analysed. Only reaction times for correct responses were used in the analysis. The error rate was very small - only 1%. Any reaction times which were two standard deviations above or below the mean were removed. This was done to exclude any unusually high or low reaction times - caused by equipment malfunction. Approximately 1.6% of the data was excluded on this criterion.

There was no significant N effect shown in the adult reaction time data. ($t(30) = 0.69$, NS), although the trend was in the same direction as that shown by children - high-N words were read faster (mean RT = 543.7 ms) than low-N words (mean RT = 547.7 ms). Mean reading latencies for each subject for words differing in N value are presented in Appendix A6.4.

DISCUSSION

Thus the results of Experiments 4 and 5 indicate a developmental trend in which knowledge about orthographic neighbourhood, or more loosely, of more frequent correspondences, is used by poor and average readers, less important in good readers, and not used by skilled adult readers.

The results showed that there was a reliable N effect for children in the lexical decision task, the reading task and the spelling task, although the N effect was less marked in the lexical decision and reading performance of the good reader group, and was no longer present in the reading task for skilled adult readers. However there is a possibility that high-N words, which children found easier to read, spell and judge as words, are acquired earlier in life than low-N words and that the N effect is in fact due to the age of acquisition effect. The following experiment investigates this possibility.

6.4(a) EXPERIMENT 6

METHOD

SUBJECTS

The subjects were 40 students at the City of London Polytechnic. All spoke English as their first language.

STIMULUS MATERIALS

The words were the thirty nine high-N words and thirty nine low-N words used in the previous experiment. An additional eleven early age of acquisition words and eleven late age of acquisition words were added to the list, so that there were fifty four-letter words and fifty five-letter words. These words were useful, as there were already AOA norms available, and so a brief check could be made as to the reliability of the ratings obtained in this experiment (See Appendix A6.5).

PROCEDURE

The words were printed, in random order, and alongside each word was a 7 point rating scale. The pages were shuffled and assembled into four booklets, so that each booklet contained the pages in a different random order.

Instructions were printed clearly on the front page of the booklet. These are shown in Appendix A6.6. Subjects were asked to estimate when they first learned the words that were about to be presented. The instructions went on to explain the code by which the words were to be rated, and to emphasise that all the words were to be interpreted as nouns. The instructions used were the same as those used by Gilhooly and Logie (1980) to obtain their age of acquisition norms.

6.4(b) RESULTS

Examining the ratings of the additional eleven early AOA words and the eleven late AOA words, a correlation analysis showed that the ratings obtained using City of London Polytechnic students were very highly correlated with the measures that Gilhooly and Logie (1980) obtained - $r = 0.984$. These mean ratings are presented in Table 6.8 on page 220.

The mean AOA rating for the high N words was 2.97. The mean AOA rating for the low N words was 3.36. The AOA ratings for the high N words were not significantly different from the AOA ratings for low N words (t = 1.91, df 39, p = 0.064). Mean ratings of age of acquisition measures for high and low-N words are presented in Appendix A6.5.

CONCLUSION

The results showed that high-N words, which children found easier to read, spell and make a lexical decision about did not differ from low-N words in their age of acquisition rating i.e. the N effect is not attributable to the age of acquisition effect.

C H A P T E R   S E V E N

REGULARITY VERSUS CONSISTENCY : THE EFFECT OF BODIES ON CHILDREN'S READING

7.1 INTRODUCTION

There has been extensive research on the effects of word regularity and consistency, mainly with skilled readers. This introduction to Experiment 7 and 8 will briefly summarise recent research with adults and describe some relevant developmental studies carried out in the light of results from studies of skilled reading.

According to the Dual Route Theory, as described by Coltheart (1978), both direct visual access and the grapheme-phoneme correspondence systems are available. Problems arise if an output from the GPC route precedes processing by means of direct visual access. In the GPC route the first step in recognising a word is to translate it into phonemes by means of spelling-sound correspondence rules. It thereore follows that irregular words, like PINT and CASTE, which are exceptions to the rule, will cause difficulties; either they will be recoded wrongly, or some additional mechanism must be added to "correct" their mispronunciation. If the GPC route is used, then the recognition of irregular words will be slower than the recognition of regular words.

Several studies have tested this prediction, with contradictory results. Evidently irregularity of spelling to sound correspondence appears to have the predicted result on naming latency. Baron and Strawson (1976), Gough and Cosky (1977), Stanovich and Bauer (1978), Coltheart et al (1979), Parkin (1983) all reported longer naming latencies for exception words than for controls, though Seidenberg, Waters, Barnes and Tanenhaus, (1985) reported that the effect holds only for low-frequency words. With the lexical decision task however, no such consistent finding is obtained. Stanovich and Bauer (1978) and Barron (1979) found the predicted effect, but neither Coltheart et al nor Seidenberg et al (1985) obtained this result.

These inconsistencies are in need of resolution. Some of the differences are almost certainly due to differences in the criteria employed by different investigators to select exception words. It seems unlikely that the resolution will reveal the simple effect of irregularity predicted when the grapheme to phoneme correspondence route is used. Since the regularity effect may not be consistently obtained in the lexical decision task, the two experiments reported in this chapter both involve reading aloud. Accuracy measures were also used, because latency data even from skilled young readers tend to be very variable and differences in accuracy reduce the number of responses on which mean latency can be based.

The orthography of English is alphabetic; the spellings of English words correspond to the phonemes of their spoken forms. Few of these correspondences are one to one: many phonemes can be spelled in more than one way, and many spelling patterns correspond to more than one phoneme. For example, EA corresponds to one phoneme in LEAK, another in HEAD and still another in STEAK. The spelling-sound correspondence rules will, therefore, include (at least) three different rules which must be applied to words containing EA. So it is intuitively plausible that the recoding of such phonologically ambiguous strings may be complicated relative to that of unambiguous strings. Thus, when the GPC route is used, phonologically ambiguous words will be more difficult to read than unambiguous ones (although Coltheart suggests that phonemic values are stored in a frequency based tabulation so that the most common value will be assigned regardless of the number of possible pronunciations). This prediction has not been supported by research. Seidenberg, Waters, Barnes and Tanenhaus (1982) found no difference in naming latency in adults, between words with ambiguous spelling patterns e.g. DEAD, GOWN and unambiguous controls e.g. NOTE, BOIL, which were matched for frequency.

The most recent developmental work on regularity effects in children was carried out by Waters, Bruck and Seidenberg (1985). They evaluated the use of spelling-sound information in both reading and spelling by having eight year old children read and spell nonwords and five types of words that differed in terms of their regularity for reading and spelling:-

TYPE 1 Regular words (e.g. DISH) highly regular both in reading and spelling

TYPE 2 Regular words (e.g. BEEF) regular in terms of reading but less so in spelling because they have more than one legal spelling (e.g. -IEF, EAF etc.).

TYPE 3 Ambiguous words contain spelling patterns with two or more associated pronunciations, both of which occur in many words e.g. FEAR, BEAR. They are also ambiguous in terms of spelling e.g. DEER.

TYPE 4 Exception words (e.g. HAVE) have common spelling patterns that are pronounced irregularly. There are no -ave words that rhyme with HAVE.

TYPE 5 Strange words (e.g. ACHE) have irregular pronunciations, like exception words. However they contain spelling patterns that occur in very few other English words.

The response measures used were mean number of errors and pronunciation latencies for the reading task and mean number of errors for the spelling task.

Results indicated that all children attempted to use spelling-sound correspondences in both reading and spelling. Children were divided into three groups - good readers-good spellers, good readers-poor spellers i.e. mixed, and poor readers-poor spellers. In the spelling task, they found that all subjects made the most errors on strange words and that all subjects made more errors on exception words compared to regular words. However, only the good subjects made fewer errors on regular words as compared to ambiguous and Type 2 regular words. Poor and mixed subjects made the same number of errors on ambiguous and regular words, and on Type 2 regular and regular words.

In the reading task, good subjects made more errors on exception words than on regular, Type 2 regular and strange words, and more errors were made on ambiguous words than on regular words. The patterns of performance for the mixed and the poor subjects were very similar with both groups making more errors on strange and exception words compared to both regular and Type 2 regular words, and ambiguous words. The poor subjects also made more errors on exception words then on ambiguous words. Thus, all three groups made more errors on exception words compared to regular words, and failed to show any differences between regular and Type 2 regular words. In addition, the good subjects made more errors on ambiguous words than on regular words, while the mixed and the poor subjects made more errors on strange words than on regular words. Waters, Bruck and Seidenberg suggested that all three groups attempted to use spelling-sound correspondences in both reading and spelling but the poorer readers and spellers had weak knowledge of these correspondences and were less systematic in their use of them. The Waters et al (1985) paper can be criticised because some of the words in TYPE 5 - the orthographically strange category - were sometimes regular, sometimes irregular. They also did not take into account the number of neighbours a target item had, nor its body type - as described by Patterson and Morton (1985), so their results must be treated with caution.

The regularity effect, as explained by dual route theorists may not be quite as simple as previously described. Glushko (1979) showed that consistency, not regularity affected the ease with which words and nonwords could be read. Kay and Marcel's research (1981) showed that inconsistent pseudowords could be significantly shifted towards irregularity if there was prior presentation of an irregular word neighbour. This therefore suggests that the advantage of regular words over irregular words reflects the use of analogy strategies rather than the application of GPC rules.

However, Parkin's (1983) work on regularity and consistency effects in word and nonword reading, described in Chapter Three, suggested that the regularity effect was robust and replicable, but that Glushko's consistency effect was not a reliable phonemenon; only affecting nonword reading.

In summary, research to date which has examined the effects of regularity and consistency has produced conflicting results. Such results indicate that both regularity and consistency may be a matter of degree rather than a straightforward dichotomy. It also seems that word frequency and inter-list priming can influence the extent to which assembled phonology may interfere with performance.

Patterson and Morton (1984) tried to explain the data by their modified standard model which consisted of three routes to the pronunciation of a word - two lexical routes, a semantic route and a visual route, and a non-lexical OPC (Orthography-to-Phonology Correspondence) route which uses GPCs and bodies. This model was described at length in Chapter 3 'Theories of word recognition'. Patterson and Morton suggested five body types which required differentiation within the OPC system.

1. CONSISTENT BODIES e.g. -aze where the body sub-system offers a single mapping.
2. CONSENSUS/HERETIC BODIES e.g. -int where the overwhelming consensus is /Int/ with the single heretic pint.

3. GANG BODIES - the opposite of consistent bodies; the gang pronunciation is irregular rather than regular with respect to GPC rules. e.g. -old.
4. GANG/HERO bodies e.g. -OOK with the single regular hero spook.
5. AMGIBUOUS bodies e.g. -OVE, -EAF with both regular and irregular pronunciations.

They hypothesised that representations in the body sub-system are established on the basis of experience with words, so that only the majority mapping is represented. The body routine therefore provides regular pronunciations for both consistent and consensus/heretic bodies; irregular pronunciations for gang and gang/heroes; and a regular/irregular pronunciation for ambiguous bodies. However Patterson and Morton note that there is little data on reading times and accuracy for body-level solutions although there is some data collected by Kay and Lesser (1986). Experiment 7 and 8 attempted to discover whether children acquire knowledge about body sublexical units i.e. to what extent the different body types affect children's recogition of single words.

The feature which sets the modified standard model apart from other two process models is that it includes bodies as well as GPCs and it makes specific predictions about different types of bodies. It claims that ambiguous bodies like -EAD are the only type for which more than one phonological correspondence is represented. On the other hand, consensus/heretic bodies such as -INT will, like consistent bodies such as -AZE, only map on to the dominant regular pronunciation. Alternative 'heretical' correspondences such as for PINT will not be represented. Similarly, for gang bodies (with or without heroes), only the major irregular correspondence will be stored.

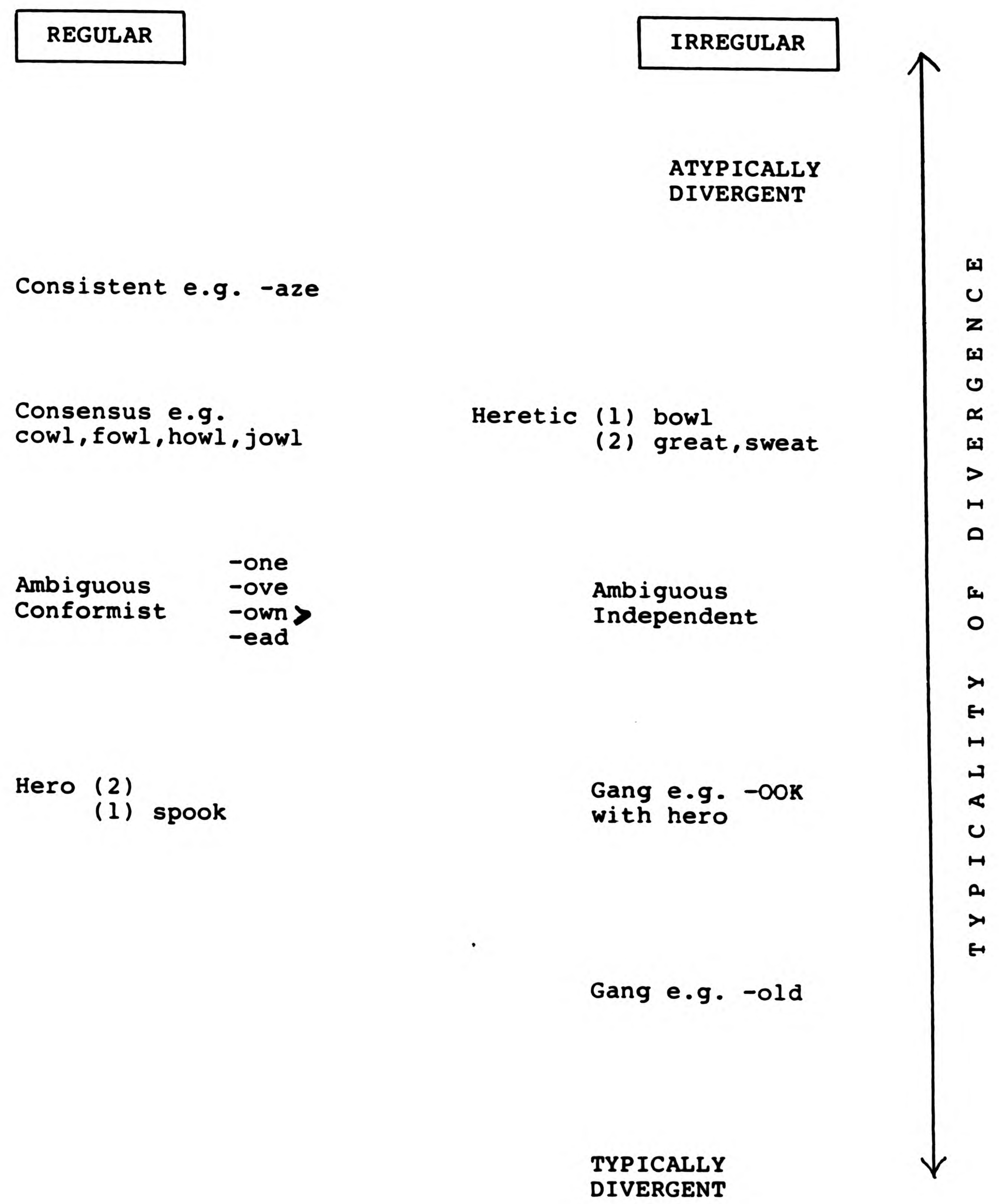
Patterson and Morton assume that the separate products of the GPC and body subsystems are not combined in any way to arrive at a final OPC pronunciation: rather, either the GPC or the body pronunciation is assigned. They make the further assumption that the GPC pronunciation is selected on 70% of occasions and that the body pronunciation is selected on the remaining 30% of occasions; this ratio derives from a "first approximation fit" of the nonword pronunciation data supplied by Kay (1982). She found that there were more regular than exception responses made to consensus heretic, ambiguous and gang with hero bodies. This is described in more detail later on in the chapter.

Patterson and Morton's model accounts less easily for Kay's (1985) findings of regularisations of gang without hero nonwords. Even though lexical analogy models would predict a considerably higher percentage of irregular pronunciations than was actually observed for such nonwords, the modified standard model predicts a much lower percentage (30% irregular pronunciation supplied entirely by the body subsystem), but an even lower percentage than that was actually observed - 23.5%.

Kay (1985) suggested that if the modified standard model is to account fully for the data on nonword pronunciation it has to broaden its view on orthographic neighbours to include initial-consonant vowel analogies or maybe even smaller segment analogies. It is considered that her suggestion emphasises the importance of orthographic variables such as N, as well as small sized orthographic units. For nonword pronunciation there may be considerable variability even within a single body type in the number of assigned irregular responses; and this may be due to the influence either of the initial consonant, or of the initial consonant and vowel.

Kay and Lesser (1985) provided a continuum of regularity within which Patterson and Morton's body types could operate. This continuum was based on Shallice, Warrington and McCarthy's (1983) typicality of divergence - see Figure 7.1. Shallice et al (1983) proposed the concept of typicality of divergence because regularity of spelling-to-sound correspondence appears to be (or to demonstrate its inluence as) a continuum rather than a dichotomy, and they considered that effects averaged over irregular words spanning the range of this continuum might be misleading. They classified irregular words into typically divergent and atypically divergent words. It can be seen from Figure 7.1 that typically divergent words were those where the most frequent correspondence was irregular e.g. -OLD (regularity being defined by gpc rules according to Venezky); the term atypically divergent was applied to words such as BOWL which contains less frequent and divergent correspondences. They predicted that typically divergent words would be easier to read than atypically divergent words, and found evidence of this in the reading performance of their surface dyslexic HTR. Kay and Lesser's patient PT did show an effect of

FIGURE 7.1: KAY & LESSER (1985) CONTINUUM OF REGULARITY



'degree' of regularity in terms of accuracy of reading response, but did not show an effect of typicality of divergence. i.e. typically divergent words e.g. BOLD, were not significantly easier to read than atypyically divergent words e.g. BOWL.

The published data on bodies is sketchy and is limited to nonword reading in adults because this was considered pertinent to the further development of the regularity/irregularity continuum.

Glushko (1979) reported that exception nonwords were pronounced in an irregular fashion in 8.7% of trials. In an extension of this finding, Kay (1985) demonstrated that the percentage of assigned irregular pronunciations was in fact variable, and that the variability was governed to a significant extent by the number of regular and exception words sharing terminal vowel consonant orthography i.e. what proportion of the orthographic neighbours were regular or irregular.

The words from which inconsistent nonwords were derived were assigned into three groups by Kay, according to the ratios of their regular and exception word neighbours. The three groups were the consensus/heretic group, the ambiguous group and gangs without heroes - terms taken from Henderson (1982) and Morton and Patterson (1984). Kay found that inconsistent nonwords of the consensus heretic group e.g. GOUCH, were given exception pronunciations (as in TOUCH) on less than 1% of the occasions by subjects in a reading task. Inconsistent nonwords of the ambiguous group eg. GLERE were given exceptional pronunciations (as in THERE) on 12.8% of trials. Inconsistent nonwords of the gang hero group eg. POOK were given exception pronunciations (as in BOOK) on 23.5% of trials.

In half of the word families in each of the three groups, the combined frequency of exception examples was greater than that of the regular examples; in the remaining half, Kay reversed this manipulation. Token frequency, i.e. the frequency of occurrences of each type of pattern - the frequency of regular versus irregular exemplars, was not observed to influence the number of assigned irregular pronunciations. Even though the irregular word TOUCH, is more common than either COUCH or SLOUGH etc., the nonword GOUCH was no more likely to be pronounced irregularly than a nonword such as NINT which shares the same orthographic endings with the fairly uncommon irregular word PINT.

Her results appeared to indicate that type frequency (the number of words with each type of pronunciation) i.e. the number of regular versus irregular exemplars) had an important part to play in nonword pronunciation. The results are also in accord with the prediction made by Marcel (1980) that, in the case of a nonword that will not find a complete lexical orthographic and phonological match, a decision between competing lexical phonological alternatives is based on the pronunciation found in the largest number of lexical cases.

A striking feature of Kay's results was that far more regular than exception responses to nonwords were made in all conditions. This is perhaps particularly surprising in the case of the 'gang with a hero' group nonwords, where the number of exception words actually exceeds the number of regular words e.e. BOOK, COOK, LOOK. According to an analogy view (Kay & Marcel 1981) the number of regular pronunciations should be considerably less than exception pronunciations of gang with a hero nonwords, because there are so many irregular neighbours.

Kay explained the preponderence of regular pronunciations firstly by a lexical analogy approach, and secondly by a modified two-process account, recently put forward by Patterson and Morton (1985).

Glushko (1979) made the explicit assumption that the terminal vowel - consonant segment was the most influential sub-lexical unit in determining pronunciation at a phonological stage. However, its importance was justified on the somewhat ad hoc grounds (Glushko, 1980) of: "the salience of rhyme for adults, and the primacy with which this develops in children". Thus his experiments and the nonword pronunciation experiments described here have manipulated only the terminal vowel-consonant segment, and have not taken into account the influences of phonological correspondences of alternative orthographic segments (such as the initial consonant-vowel). In outlining his model, Glushko (1979) did note that "a more general activation and synthesis model, with a broader experimental base, would allow for the contribution of neighbours in all positions" (p.684). So it was possible that the high numbers of regular responses were due to the predominance of regular pronunciation in initial consonant-vowel orthography. It is also possible that the initial segment was able to dominate phonological alternatives specified by terminal vowel-consonant segments.

regular/hero pronunciation

e.g. /pu:k/ e.g. spook

nonword

pook

exception/gang pronunciation

/puk/ e.g. book, cook, rook

Lexical analogies sharing the final segment would clearly have specified the exception pronunciation. However the initial CV segment is usually pronounced /pu:/,and it is perhaps this influence that contributes to the large number of observed regular pronunciations in this example.

Thus there has been some work on skilled nonword reading suggesting that a dichotomous view of regularity is no longer tenable and that an elaboration of the regularity/ irregularity continuum may be more appropriate.

However there is no published data on children's reading of words using Patterson and Morton's (1984) body classification. In order to examine whether children's reading performance is influenced by subword segments of this type a word list was constructed within the regular/irregular continuum. These words were classified according to Patterson and Morton's (1984) body categories and the following measures of psychological interest were calculated:

a) the hostility of each word's neighbours as measured by the number of hostile neighbours that a word has

b) the frequency of hostile neighbours as measured by the summed American Heritage frequency of hostile neighbours

Both these measures are of interest because they reflect whether a word is typically divergent or atypically divergent - as described in Shallice, Warrington and McCarthy's (1983) multiple levels model. An atypically divergent word has many hostile neighbours while a typically divergent word has very few hostile neighbours, and would be easier to read.

The second measure of frequency of hostile neighbours enables us to explore whether a few high frequency hostile neighbours are more disruptive than many lower frequency hostile neighbours.

In looking at degree of regularity Shallice et al predict that there is a typicality of divergence rather than the straightforward regularity effect predicted by the standard GPC model. However there has been no experimental work on children's reading to show that this is the case.

The aim of Experiment 7 was to see whether a typicality of divergence affects children's reading performance. It was an exploratory study because trying to obtain large equal word sets which were controlled for frequency, word length and body type was very difficult; for example it was difficult to find low frequency gang examples where the hero was high frequency. Thus some categories had only two members.

It is hypothesised that bodies which have a straightforward orthographic mapping i.e. consistent bodies and gang bodies will be much easier than consensus bodies and gang-hero bodies which have one or two inconsistent neighbours.

If the Dual Route model has the best predictive power, then performance will be based on grapheme phoneme correspondences for vowel digraphs and there will be no difference seen among the body types, other than that which could be accounted for by a regular/irregular distinction. However, for bodies which reflect a degree of consistency, predictions are more complex. If bodies, as described in the modified standard model, are acquired and used by readers then word reading should be affected as follows:-

a) High Frequency items are not affected by regularity/consistency.
Prediction: There will be no difference in reading performance for high frequency consistent, consensus and ambiguous body types.
N.B. Dual route and Modified Standard Models make the same prediction.

b) Low Frequency items <u>are</u> affected by regularity/consistency. The Dual Route Model and the Modified Standard Model make different predictions here.

i) GPC LF

There will be no difference in reading performance for consistent, consensus and Ambiguous body types which are regular, but these will be easier to read than irregular gang/hero and gang body types (which will not differ from each other).

ii) OPC LF (only the body subsystem is used). There will be no difference in reading performance for consistent and Gang bodies. These will be easier than consensus and Gang/hero bodies, which will not differ, and these will be easier than Ambiguous bodies.

iii) OPC LF (GPC + OPC subsystem is used, with GPC providing 70% regular pronunciation).
Consisent bodies shouldn't differ significantly from Consensus bodies. These will be easier than Gang/hero and gang bodies, which should yield the same performance in terms of accuracy. Ambiguous bodies will be the most difficult, since they are the only bodies to have more than one mapping on offer.

If in addition these are built up on the basis of experience, then perhaps frequency and number of members of the body need to be taken into account. Considering Marsh, Friedman, Welch and Desberg's (1981) developmental model it was predicted that performance on bodies would differ according to level of reading ability. At Stage 1 - the Linguistic Guessing Stage and Stage 2 - the Discrimination Net Guessing Stage only frequency would be a significant variable, with high frequency words being easier to read than low frequency words. At Stage 3 - the Sequential Decoding Stage, low frequency gangs and gang/heroes would be more difficult than their consistent matches. At Stage 4 - the Hierarchical Decoding Stage the predictions were the same as the Dual Route model; that there would be no differences seen among the body types, other than that which could be accounted for by regular/irregular distinction. However, if children did read analogously, then body type differences as specified by the OPC system would be apparent. If reading were truly analogous - one high frequency neighbour should serve to bias pronunciation according to whether it is regular or not.

Therefore our word list has taken into account both the friendliness of neighbours and their frequency, because analogy theory would predict that only one high frequency neighbour is required to affect children's reading performance - a high frequency, hostile neighbour will make the word more difficult to read, a high frequency friendly neighbour will make the word easier to read.

7.2(a) **EXPERIMENT 7 - EXPLORATORY STUDY**

METHOD

Twenty six subjects from the first year junior of an inner city school acted as subjects. Their mean chronological age was 8:3, range 7:10-8:7. Their mean reading age was 8:2, range 6:7-10:0.

STIMULUS MATERIALS

A word list of 62 items was constructed within the regularity/irregulrity continuum. These words were classified according to Patterson and Morton's (1984) body categories; the body types included consistent, gang, hero, heretic, consensus and gang/hero body types, i.e. gang with hero. Measures of psychological interest were calculated:-

a) whether the item was high or low frequency, according to Hofland and Johansson's (1982) frequency count.

b) the hostility of each word's neighbour - either no hostile neighbours, or mostly hostile neighbours, or mostly friendly neighbours.

c) the frequency of hostile neighbours - either no hostile neighbours, or hostile neighbour is high frequency, or hostile neighbour is low frequency.

The word list is set out in Appendix A7.1. The mean frequencies of words in each body type used in the statistical analyses are presented in Table 7.1.

APPARATUS AND PROCEDURE

Words were presented one at a time in lower case print on a video display unit connected to a BBC-B microcomputer which randomised order of presentation separately for each subject. The subjects had to read each word aloud and the rate of stimulus presentation was subject-paced. The task was carried out in a quiet room and each subject's responses were tape recorded.

TABLE 7.1: MEAN FREQUENCIES OF WORD LIST USED IN EXPERIMENT 7

ANALYSIS I NO HOSTILE NEIGHBOURS

	REGULAR	IRREGULAR
	Consistent	Gang
HF	162.8	177.2
LF	1.4	2.6

ANALYSIS II MOST NEIGHBOURS HOSTILE AND (ONE AT LEAST) HIGH FREQUENCY

	Hero	Heretic
LF	1.3	3.7

ANALYSIS III MOST NEIGHBOURS FRIENDLY WITH LOW FREQUENCY HOSTILE NEIGHBOURS

	Consensus	Gang/hero
HF	228.7	321.0
LF	2.0	2.8

MOST NEIGHBOURS FRIENDLY WITH HIGH FREQUENCY HOSTILE NEIGHBOURS

	Consensus	Gang/hero
HF	145.0	314.9
LF	3.4	2.2

ANALYSIS IV HIGH FREQUENCY ITEMS ONLY

	Consis	Gang	Consensus		Gang/hero	
	No hostile neighbour		Most neighbours friendly			
			Hostile neighbour		Hostile neighbour	
			LF	HF	LF	HF
HF	162.8	177.2	228.7	145.0	321.0	314.9

ANALYSIS V LOW FREQUENCY ITEMS ONLY

	Consis	Gang	Hero	Heretic	Consensus		Gang/hero	
	No hostile neighbour		Most neighbours hostile - 1 at least HF		Most neighbours friendly			
					Hostile neighbour			
					LF	HF	LF	HF
LF	1.4	2.6	1.3	3.7	2.0	3.4	2.8	2.2

7.2(b) RESULTS

Results of an item analysis are reported. A subject analysis was not appropriate because of small cell numbers e.g. in the low frequency gang with high frequency hero category: a limitation of the word set. Thus developmental aspects were not examined in this exploratory study.

ITEM ANALYSIS

An item (or stimulus) analysis was performed in which a difficulty value for each word was calculated by counting the number of subjects who read it correctly. The mean number of children who read words correctly, expressed as a proportion is presented in Table 7.2. A series of logistic models were fitted to this measure using as predictors the word characteristics Word Frequency (Hofland & Johansson Frequency, 1982) hostility of neighbours, frequency of hostile neighbours and regularity/ irregularity. The GLIM computer program (Baker and Nelder, 1978) was used to perform the analyses.

Several different analyses were performed:-

1) ANALYSIS OF WORDS WITH NO HOSTILE NEIGHBOURS

i.e. numbers 1-21 in the word list (see Appendix 7.1). The difficulty values or low and high frequency (regular in GPC terms) consistent bodies and irregular gang bodies were compared. The Dual Route model predicts that performance will

TABLE 7.2: MEAN NUMBER OF CHILDREN WHO READ WORDS CORRECTLY, EXPRESSED AS A PROPORTION OF TOTAL NUMBER OF CHILDREN IN EXP.7

ANALYSIS I NO HOSTILE NEIGHBOURS

	REGULAR	IRREGULAR
	Consistent	Gang
HF	0.646	0.639
LF	0.423	0.292

ANALYSIS II MOST NEIGHBOURS HOSTILE AND (ONE AT LEAST) HIGH FREQUENCY

	Hero	Heretic
LF	0.315	0.26

ANALYSIS III MOST NEIGHBOURS FRIENDLY WITH LOW FREQUENCY HOSTILE NEIGHBOURS

	Consensus	Gang/hero
HF	0.577	0.635
LF	0.362	0.356

MOST NEIGHBOURS FRIENDLY WITH HIGH FREQUENCY HOSTILE NEIGHBOURS

	Consensus	Gang/hero
HF	0.731	0.750
LF	0.500	0.400

ANALYSIS IV HIGH FREQUENCY ITEMS ONLY

	Consis	Gang	Consensus		Gang/hero	
	No hostile neighbour		Most neighbours friendly			
			Hostile neighbour		Hostile neighbour	
			LF	HF	LF	HF
HF	0.646	0.639	0.577	0.731	0.635	0.750

ANALYSIS V LOW FREQUENCY ITEMS ONLY

	Consis	Gang	Hero	Heretic	Consensus		Gang/hero	
	No hostile neighbour		Most neighbours hostile - 1 at least HF		Most neighbours friendly			
					Hostile neighbour			
					LF	HF	LF	HF
LF	0.423	0.292	0.315	0.260	0.362	0.500	0.356	0.400

be based on grapheme-phoneme correspondences, so that any differences will be accounted for by a regular/irregular distinction i.e. regular words will be easier to read than irregular words.

It was hypothesised that high frequency items would not be affected by regularity/consistency. Any consistency effects observed would be more likely be observed for the lower frequency items, as found by Seidenberg et al (1984). If the body subsystem alone was used, then there would be no difference in the reading performance of the two body types as they have no hostile neighbours. If the body subsystem was used in conjunction with the GPC system, then the prediction would be the same as for the Dual Route model.

The mean difficulty value for words with no hostile neighbours are presented in Table A7.3 i.e. the mean number of subjects who read the words correctly. The maximum score is 26. The means are also expressed in terms of proportion p.

TABLE A7.3

MEAN NUMBER OF CHILDREN WHO READ WORDS WITH NO HOSTILE NEIGHBOURS

	REGULAR CONSISTENT		IRREGULAR GANG	
	$\bar{x}$	p	$\bar{x}$	p
HIGH FREQUENCY	16.8	0.65	16.6	0.64
LOW FREQUENCY	11	0.42	7.6	0.29

MAXIMUM SCORE 26

Fitting a simple model of the two predictor variables Categorical Frequency (each word was categorised as High Frequency or Low Frequency) and Regularity explained 45% of the total variance (R^2 = 0.447). R^2 values are reported for each variable, as is the significance of the difference in explained variance using χ^2. This was done for all the analyses which follow. The Regularity effect was not significant (R^2 = 0.027, χ^2 = 2.72, NS). The frequency effect, as measured by Categorical Frequency was highly significant (R^2 = 0.423, χ^2 = 42.74, $p < .001$).

Interactions were not analysed because of small cell numbers e.g. There were 5 high frequency consistent words and 5 low frequency gang words. Interactions from small cell numbers such as these would be of doubtful validity.

2) ANALYSIS OF LOW FREQUENCY WORDS WITH MOSTLY HOSTILE AND HIGH FREQUENCY NEIGHBOURS

i.e. numbers 22-31 on the word list (see Appendix A7.1). In this analysis the reading performance of children on low frequency heroes and low frequency heretics with hostile neighbours (at least one of which was high frequency) was compared. The Dual Route model would predict a regularity effect, with heroes being easier to read than heretics, as would Patterson and Morton (1984). If only the body subsystem were being used then there would be little difference in children's reading performance of words in these two body types, because they were all low frequency and had similar hostile neighbourhoods.

Children found regular heroes ($\bar{x}$ 8.2) not significantly different from irregular heretics ($\bar{x}$ 6.8) in the reading tasks. It appears that for low frequency words with a hostile environment, the regular/irregular continuum is not significant - Regularity explained only 3% of the total variance (R^2 = 0.034, χ^2 = 0.92, NS).

3) ANALYSIS OF WORDS WITH MOSTLY FRIENDLY NEIGHBOURS

This analysis investigated the reading of high and low frequency regular consensus/heretic bodies and irregular gang/hero bodies which either had a high frequency hostile neighbour or a low frequency hostile neighbour. If the body subsystem is used then consensus/heretic and gang/hero bodies which have a high frequency hostile neighbour will be more difficult than those with a low frequency hostile neighbour. Means and proportions are presented in Table 7.4.

The simple additive model of Categorical Frequency Regularity and Frequency of hostile neighbours explained 38% of the Total Variance (R = 0.383). R^2 values are reported for each variable, as is the significance of the difference in explained variance using χ^2. Regularity was not significant (R^2 = 0, χ^2 = 0, NS). However the Frequency of hostile neighbours was a significant variable (R^2 = 0.074, χ^2 = 13.1, $p < .001$). Again, it was deemed inappropriate to fit any interactions because of small cell numbers.

4) ANALYSIS OF HIGH FREQUENCY WORDS

i.e. numbers 1-5, 12-16, 37-45, 67-71, 94-95, 54-58, 88-91 of the word list in Appendix A7.1. The variables of interest in this analysis were the number of and hostility of

TABLE 7.4: MEAN NUMBER OF CHILDREN WHO READ WORDS WITH MOSTLY FRIENDLY NEIGHBOURS CORRECTLY

	REGULAR CONSENSUS/HERETIC				IRREGULAR GANG/HERO			
	HF HOSTILE NEIGHBOUR		LF HOSTILE NEIGHBOUR		HF HOSTILE NEIGHBOUR		LF HOSTILE NEIGHBOUR	
	$\bar{x}$	p	$\bar{x}$	p	$\bar{x}$	p	$\bar{x}$	p
HIGH FREQUENCY	19	0.73	15	0.58	19.5	0.75	16.5	0.64
LOW FREQUENCY	13	0.50	9.4	0.36	10.5	0.40	9.25	0.36

MAXIMUM SCORE = 26

neighbours. It was considered that words with no hostile neighbours might be easier to read than words with a low frequency hostile neighbour, and both of these categories would be easier than words with a high frequency hostile neighbour. The Dual Route model and the Modified Standard Model of Patterson would predict that there would be no body effects for high frequency words.

The body subsystem would also predict that no differences in body type would be found in high frequency words.

The means used in this analysis are presented in Table 7.5. A model which included the predictor variables Regularity, Hostility of Neighbours and Frequency of Neighbours only explained 6% of the Total Variance in the data (R^2 = 0.071). There were no significant effects of Regularity (R^2 = .0032, χ^2 = 0.4, NS) or Hostility of Neighbours (R^2 = 0, χ^2 = 0, NS) Frequency of Neighbours was a significant variable (R^2 = 0.069, χ^2 = 8.6, $p < .005$) but in the opposite direction to that predicted, i.e. words with high frequency, hostile neighbours were easier to read. These results appear to support Patterson and Morton's suggestion that no differences in body type would be found in high frequency words.

5) ANALYSIS OF ALL THE LOW FREQUENCY WORDS

i.e. numbers 6-10, 17-21, 22-26, 27-31, 59-63, 92-93, 49-53, 84-87 of the word list in Appendix 7.1.

TABLE 7.5: MEAN NUMBER OF CHILDREN WHO READ HIGH FREQUENCY WORDS CORRECTLY

		NO HOSTILE NEIGHBOURS	MOST NEIGHBOURS FRIENDLY	
		NO HOSTILE NEIGHBOUR	HOSTILE NEIGHBOUR HIGH FREQUENCY	HOSTILE NEIGHBOUR LOW FREQUENCY
REGULAR	$\bar{x}$	16.8 (CONSISTENT)	19 (CONSENSUS/ HERETIC)	15 (CONSENSUS/ HERETIC)
	p	0.65	0.73	0.58
IRREGULAR	$\bar{x}$	16.6 (GANG)	19.5 (GANG/HERO)	16.5 (GANG/HERO)
	p	0.64	0.75	0.63

MAXIMUM SCORE = 26

The modified standard model predicts that differences due to body type will be more marked in low frequency words. If just the body subsystem is used, then consistent and gang bodies would be easier than consensus/heretic and gang/hero type bodies which have mainly friendly neighbours and these would be easier than heroes and heretics which have a mainly hostile neighbourhood. The frequency of the hostile neighbour was also considered to have an effect on reading performance i.e. consensus body with a high frequency heretic would be more difficult to read than consensus body with a low frequency heretic.

The Dual Route model predicts that there will be an effect of regularity, but not of body type, hostility of neighbourhood or frequency of hostile neighbours. The Modified Standard Model would predict that performance is based on degree of conflict between GPC and body subsystem. The means used in this analysis are set out in Table 7.6.

A model which included the predictor variables Regularity, Hostility of Neighbours and Frequency of Neighbours explained 22% of the variance in the data (R^2 = 0.221). Hostility of neighbours appeared to be significant (R^2 = 0.138, χ^2 = 12.9, $p < .001$) - i.e. words with no hostile neighbours or with mostly friendly neighbours were easier to read than words with mostly hostile neighbours. Words with high frequency hostile neighbours were easier than words with a low frequency hostile neighbour (R^2 = 0.046, χ^2 = 4.34, $p < .05$). This result is surprising. One would have expected that low frequency words which had high

TABLE 7.6: MEAN NUMBER OF CHILDREN WHO READ LOW FREQUENCY WORDS CORRECTLY

		NO HOSTILE NEIGHBOURS	MOSTLY HOSTILE NEIGHBOURS	MOSTLY FRIENDLY NEIGHBOURS	
		NO HOSTILE NEIGHBOURS	HOSTILE NEIGHBOUR	HOSTILE NEIGHBOUR HF	HOSTILE NEIGHBOUR LF
REGULAR	$\bar{x}$	11 (CONSISTENT)	8.2 (HERO)	13 (CONSENSUS/ HERETIC)	9.4 (CONSENSUS/ HERETIC)
	p	0.42	0.32	0.50	0.36
IRREGULAR	$\bar{x}$	7.6 (GANG)	6.8 (HERETIC)	10.5 (GANG/HERO)	7.4 (GANG/HERO)
	p	0.29	0.26	0.4	0.29

MAXIMUM SCORE = 26

frequency hostile neighbours would be more difficult than those with low frequency hostile neighbours. The Regularity effect was significant for low frequency items (R^2 = 0.053, χ^2 = 4.99, $p < .05$). Consistent, consensus and hero body types were easier than gang and heretic body types.

DISCUSSION

Thus the exploratory study Experiment 7 showed that regularity, as defined in more detail in Morton and Patterson's (1984) OPC system is dependent upon the frequency of the target word and the hostility and frequency of the word neighbours. The first item analysis showed that there was no regularity effect for words with no hostile neighbours. This result is consistent with a pattern of results, if just the body subsystem were being used. However since both high and low frequency items were used, then no very firm conclusions can be drawn. The second item analysis, performed on low frequency words with mostly hostile and high frequency neighbours showed that there was no difference in performance for body types where the neighbours were hostile and high frequency. Again these results support predictions made if just a body subsystem were used. The third item analysis in which target words had mostly friendly neighbours showed that the frequency of hostile neighbours was a significant variable. Surprisingly, high frequency hostile neighbours increased reading accuracy. This result does not appear to support any of the models.

The regularity effect is also dependent upon the frequency of the target word, as shown by the fourth analysis, in which only high frequency words were analysed. There were no significant effects of Regularity, as defined by body type on high frequency words, a result that supported the prediction of Morton and Patterson (1984). However there were significant effects of frequency of hostile neighbours. Words with high frequency hostile neighbours were easier to read, although it was predicted that they would be more difficult to read.

The analysis of the low frequency words in the fifth analysis showed that there was a significant regularity effect - consistent, consensus and hero body types were easier to read than gang and heretic body types. This result does not support predictions made if only the body subsystem were used, that consistent and gang bodies would be similar in reading difficulty. Hostility of neighbours appeared to be significant - words with no hostile neighbours were easier to read than words with hostile neighbours. As in the analysis for high frequency words, it was found, counterintuitively, that words with high frequency hostile neighbours were easier than words with low frequency hostile neighbours. The results of analysis 4 and 5 indicate that regularity/body effects, when they do occur, appear to be confined to low frequency items - indicating support for the notion that high frequency items may be processed by a separate lexical route. The results also suggest that the ease with which a word is read may depend upon the hostility and frequency of its neighbours.

SUMMARY OF RESULTS - EXPERIMENT 7

To summarise the results of this preliminary experiment in terms of predictions set out in the introduction:-

a) Both the Dual Route and Modified Standard Model predicted that high frequency items would not be affected by regularity/consistency. The results of Analysis 4 supported this prediction. There were no effects of body type for high frequency words.

b) Low frequency items are affected by regularity/ consistency. Three different predictions were presented

i) GPC LF

There will be no difference in reading performance for the body types over and above that which can be explained by a regularity/irregularity distinction

ii) OPC LF (only the body subsystem is used)

There will be no difference in reading performance for Consistent and Gang bodies. These will be easier than consensus and Gang/hero bodies, which will not differ significantly from each other.

iii) OPC LF (GPC & OPC subsystem is used, with GPC providing 70% regular pronunciation)

Consistent and consensus bodies should not differ significantly. They will be easier than gang/hero and gang bodies, which will not differ significantly from each other. However heroes and heretics will differ significantly. Regular heroes will be easier than irregular heretics. Ambiguous bodies will be most difficult.

The results of an analysis of low frequency words with hostile and high frequency neighbours (Analysis 2) appeared to support a pattern of performance consistent with just body subsystem being used. Heroes did not differ significantly from heretic. These two body types were similar in that they were low frequency words in a hostile neighbourhood.

However, the results of an analysis examining a wider range of low frequency body types - consistent, consensus/heretic, gang, gang/hero, hero and heretic body types (Analysis 5) provided support for the predictions of the Modified Standard Model. Regular consistent, consensus and hero body types were easier than irregular gang and heretic body types.

An analysis of body types with friendly neighbours i.e. consensus/heretic and gang/hero bodies (Analysis 3), showed that variables other than that of regularity were important. In this analysis the frequency of hostile neighbours was a significant variable. It appeared that words with high frequency hostile neighbours were easier to read than words with low frequency hostile neighbours.

In an analysis of all the high frequency words (Analysis 5), results were as predicted for all the models - there were no regularity/body effects for high frequency words. Again, the surprising finding observed was that the frequency of neighbours was a significant variable; words with high frequency hostile neighbours being easier to read.

This facilitatory effect of a frequent hostile neighbour was not predicted by the dual route model, the Modified Standard model or analogy models. It may be due to other distinguishing features in the high frequency hostile neighbour word sets. It is not the case that these high frequency words with hostile high frequency neighbour were more frequent than the comparable word set with hostile low freqency neighbours. In fact they were slightly <u>less</u> frequent (Consensus/heretic body type with HF hostile neighbour xf = 145.0, Consensus/heretic body type with LF hostile neighbour xf = 228.7; Gang/hero body type with HF hostile neighbour xf = 314.9, Gang/hero body type with LF hostile neighbour xf = 321.0). It is possible that meaningfulness or Imagery differences may account for the facilitory effect.

Experiment 8 examines <u>both</u> the effects of body type and orthographic neighbourhood size on the reading process. In Chapter Six orthographic neighbourhood size, as defined by N, was shown to have an effect on the reading, lexical decision and spelling task. Experiment 7 showed that body type manipulations also affect reading performance. In Experiment 8, therefore, an attempt was made to manipulate orthographic neighbourhood size and body type. This was a developmental study which compared three age groups, in order to observe any developmental trend in children's use of body segments in reading.

7.3(a) EXPERIMENT 8

Experiment 8 examined the concept of regularity as defined by body types. However there were larger word sets matched for word frequency and N - the number of word neighbours. Frequency of word neighbours was not manipulated because it is not relevant to all the word sets used. Thus Experiment 8 manipulates both a probabilistic measure of orthographic regularity - N and regularity according to grapheme-phoneme rules i.e. bodies.

METHOD

Sixty children from the first, second and third year in three inner city schools acted as subjects. Mean chronological ages and reading ages are presented in Table 7.7.

TABLE 7.7: CHRONOLOGICAL AND READING AGES OF SUBJECTS

YEAR GROUP	N	MEAN CHRONOLOGICAL AGE	MEAN READING AGE
1	20	7:8	7.5
2	20	9.2	8.4
3	20	10:4	9.1

Reading ability was assessed by means of the SPAR Reading Test (Young, 1970). The children were assigned to groups according to which year they were in i.e. first year children were assigned to Group 1. Children were assigned to year groups rather than reading ability groups because it was considered that if children were allocated to reading groups this would not necessarily provide homogenous groups, as three different schools were used in this study.

STIMULUS MATERIALS

The stimulus items consisted of 128 words, grouped according to body type, N, and word frequency. *1* Thus there were low frequency, low N ambiguous, consensus/heretic, gang/hero and gang body types; low frequency high N ambiguous, consensus/heretic body types and high frequency, low N ambiguous and consensus/heretic body types. For each body type there was a consistent body control set which was matched for N value and frequency. The word list used in Experiment 8 is set out in Appendix 7.2. Item characteristics for the word list are presented in Table 7.8.

A brief rationale of body type categories is provided, so as to indicate that Patterson and Morton's definitions of various bodies was adhered to as closely as possible.

RATIONALE FOR WORD CATEGORIES

1. **Consistent**

   All endings are regularly pronounced.

2. **Consensus/heretic**

   i Large number of regular items with one or two exceptions; these may be 3, 4 or 5 letters long

1 Devised by V. Laxon in 1985

TABLE 7.8: ITEM CHARACTERISTICS FOR WORD LIST USED IN EXPERIMENT 8

LOW FREQUENCY, LOW N BODIES

BODY TYPE		x	S.D
AMBIGUOUS	N	9.00	2.83
	f	19.38	23.49
CONSISTENT 1	N	8.50	2.50
	f	19.88	16.92
CONSENSUS/HERETIC	N	8.88	2.32
	f	22.38	22.34
CONSISTENT 2	N	8.63	1.93
	f	22.50	22.51
GANG/HERO	N	8.50	2.18
	f	21.75	23.13
CONSISTENT 3	N	8.50	2.69
	f	19.25	14.45
GANG	N	7.63	3.71
	f	24.38	26.12
CONSISTENT 4	N	7.38	2.74
	f	23.38	19.22

LOW FREQUENCY, HIGH N BODIES

BODY TYPE		x	S.D
AMBIGUOUS	N	14.63	2.74
	f	19.50	22.50
CONSISTENT 5	N	15.38	3.23
	f	18.75	21.36
CONSENSUS/HERETIC	N	16.25	3.27
	f	19.00	18.96
CONSISTENT 6	N	15.88	2.09
	f	18.00	19.11

HIGH FREQUENCY, LOW N BODIES

BODY TYPE		x	S.D
AMBIGUOUS	N	9.5	2.0
	f	243.38	123.60
CONSISTENT 7	N	9.25	1.99
	f	235.13	88.05
CONSENSUS/HERETIC	N	9.63	1.50
	f	214.13	88.99
CONSISTENT 8	N	9.5	2.87
	f	236.75	86.82

ii Smaller sets of words with only one exception.

3. <u>**Ambiguous**</u>

   Large number of items with deviant pronounciations and/or several <u>very</u> high frequency deviants and/or several kinds of deviant pronounciations and/or homographs (read, lead)

4. <u>**Gang/hero**</u>

   All irregular (covered by a special environment rule) except one, or, where there are two heroes, one or both is <u>very</u> LF e.g. asp.

5. <u>**Gang**</u>

   All irregular - covered by a special environment rule.

<u>**The rationale for items included in the N count**</u>

The rationale or items included in the N count, derivation of frequencies or neighbours, selection of items, definition of heretics and heroes, and selection of consistent matched items are presented in Appendix A7.3.

It was possible to find high frequency sets and low frequency sets of words varying in body type and N-orthographic neighbourhood size, allowing for the following comparisons:-

1. <u>**Analysis of the four low frequency, low N consistent body types**</u>. These words were chosen to be regular controls or the irregular body types. The four low frequency low N consistent bodies were analysed to see

if there were any significant differences between them in reading performance, so that the mean difficulty value for low frequency, low N consistent bodies could be used in later analyses.

2. **The N effect in low frequency words.** This analysis was possible for the ambiguous, Consistent and consensus/heretic body types. (It was not possible to do this for high frequency words because of difficulty in obtaining a set of high frequency, high N words for other body types e.g. gang words). It was hypothesised that low frequency words with many neighbours would be easier to read than those with few neighbours. It is implicit in Patterson and Morton's (1984) body subsystem that consistent bodies would be easier than consensus/heretic bodies and that ambiguous bodies would be the most difficult to read.

3. **Analysis of the frequency effect on low N words.** It was hypothesised that high frequency words with few neighbours would be easier to read than low frequency words with few neighbours. Any effects of body type might be more marked for low frequency words.

4. **Analysis of low frequency, low N words.** This was performed in order to determine whether there was a body effect for low frequency words if neighbourhood size was constant.

It should be noted that the N count involves a variety of sub-types of words, so that for items with the same N count these would vary in number. Taking as examples the target words LONE and BEAD. The N count for these words would involve

a) Words neighbours sharing <u>body</u> endings which also share the grapheme-phoneme correspondence rule

e.g. LONE, BONE, TONE

MEAD, READ, LEAD (verb, present tense)

b) Words neighbours sharing <u>body</u> endings which do not share the grapheme-phoneme correspondence rule

e.g. NONE, DONE, GONE

HEAD, READ (verb, past tense) LEAD (noun, present tense)

c) Word neighbours which share other segments of the word and also share the grapheme-phoneme correspondence rule

e.g. LOPE, LODE

BEAK, BEAT

d) Word neighbours which share other segments of the word, but which do not share the grapheme-phoneme correspondence rule

e.g. LOSE, LOVE

BEAR

e) Distant neighbours

e.g. LONG, LANE, LINE

BEND

All of these word neighbours vary markedly in frequency and they might be expected to affect reading difficulty. It was considered that there may be a Frequency x sub unit interaction, or a Frequency x proportion of helpful/unhelpful neighbours interaction.

APPARATUS AND PROCEDURE

Words were presented one at a time in lower case print on a video display unit connected to a BBC-B microcomputer which randomised order of presentation separately for each subject. The subjects had to read each word aloud and the rate of the stimulus presentation was subject paced. Subject's responses were tape-recorded.

7.3(b) **RESULTS**

SUBJECT ANALYSIS

ANALYSIS 1: ANALYSIS OF LOW FREQUENCY CONSISTENT BODIES ie. CONSISTENT GROUPS 1-4

Analysis 1 was performed to see if there were any significant differences among the four low frequency, low N consistent body types which were the regular controls for the irregular body types. Theoretically one would not expect a difference as they are all of the same body type and carefully matched for frequency and N. Means for correctly read consistent words are presented in Table 7.9.

An analysis of variance in which Age Group was a between subjects factor and consistency a within subjects factor was performed. This indicated a significant main effect of Age Group ($\underline{F}(2,57) = 21.59$, $p < .001$). Subjects did not perform significantly differently on the four consistent body types ($\underline{F}(2,57) = 0.81$, NS) and there was no significant

TABLE 7.9: MEANS CORRECTLY READ LOW FREQUENCY, LOW N CONSISTENT BODIES BY AGE GROUP

AGE GROUP	CONSISTENT 1	CONSISTENT 2	CONSISTENT 3	CONSISTENT 4
1	3.2	2.7	3.75	3.1
2	5.7	5.45	5.15	5.0
3	6.4	7.05	4.4	4.8
	5.1	5.07	4.83	4.97

MAXIMUM SCORE = 8

interaction. The full anova table can be found in Appendix A7.4. Thus the mean difficulty value for low frequency, low N consistent words were used in later analyses.

The second analysis used low frequency words only. This was appropriate because there were certain body types (e.g. high frequency gang words) which would be too few in number to make up a set.

ANALYSIS 2: The N effect in low frequency words

An analysis of variance in which Age Group (1st year, 2nd year & 3rd year) was a between subjects factor and N (High N, Low N) and Body Type (Ambiguous, Consistent & Consensus) as within subjects factors was performed. The means are set out in Table 7.10. There was a significant main effect of Age Group ($\underline{F}(2,57) = 31.99$, $p < .001$). High N words were easier to read than low N words ($\underline{F}(1,57) = 5.30$, $p < .024$). There was a significant main effect of Body Type

($\underline{F}(2,114) = 18.37$, $p < .001$) and the interaction between N and Body Type was also significant ($\underline{F}(2,114) = 17.56$, $p < .001$). (The full anova table is presented in Appendix A7.5).

Planned comparisons indicated that Age Group 3 performed significantly better than Age Group 2 ($\underline{t}(57) = 2.32$, $p < .05$) and that Age Group 2 performed significantly better than Age Group 1 ($\underline{t}(57) = 5.47$, $p < .001$).

Planned comparisons also indicated that the ambiguous body type words were significantly more difficult to read than both the consensus heretic body type ($\underline{t}(114) = 2.88$, $p < .01$), and the consistent body type ($\underline{t}(114) = 6.06$, $p < .001$) which were the easiest to read. There was no significant difference between consensus and consistent body types. The interaction between N and Body Type arose because the N effect was significant only for consistent body types ($\underline{t}(114) = 2.07$, $p < .05$).

The body x N interaction indicated that the body effect appeared less marked for low n than for high N words - the high N consistent group was easier to read than both the high N ambiguous group ($\underline{t}(228) = 6.37$, $p < .001$) and the high N consensus group ($\underline{t}(228) = 6.37$, $p < .001$). There was no significant difference between the high N ambiguous and consensus groups. For low N means, the low N ambiguous group was more difficult to read than the low N consistent group ($\underline{t}(228) = 2.48$, $p < .02$) and the consensus group ($\underline{t}(228) = 4.21$, $p < .001$). There was no significant difference between the low N consistent and consensus groups.

TABLE 7.10: MEAN CORRECTLY READ WORDS DIFFERING IN SIZE OF NEIGHBOURHOOD AND BODY TYPE BY AGE GROUP

AGE GROUP	LOW n			HIGH n			$\bar{x}$
	Ambig	Consis*	Consens	Ambig	Consis*	Consens	
1	2.65	2.95	2.95	2.90	3.65	3.0	3.01
2	5.20	5.58	5.90	5.25	6.63	5.35	5.65
3	6.25	6.73	7.20	6.60	7.45	6.4	6.77
	4.70	5.08	5.35	4.92	5.90	4.92	

MAXIMUM SCORE = 8

* The average mean or the two low N (CONSISTENT 1 and CONSISTENT 2) and for the two high N (CONSISTENT 5 and CONSISTENT 6) consistent matches for ambiguous and consensus bodies were used. T tests showed that this grouping was permissible because there was no significant difference for CONSISTENT 1 and CONSISTENT 2 body types ($\underline{t}$ = 0.21, $p < 0.84$) and none for CONSISTENT 5 and CONSISTENT 6 ($\underline{t}$ = 0.12, $p <$ 0.91).

Overall high N words were easier, and while ambiguous bodies were the most difficult, there was not, as predicted, a difference in the reading of consistent and consensus bodies.

ANALYSIS 3: The effects of frequency on low N words

A similar analysis was performed using Frequency and Body Type as within subjects factors. The means used in Analysis 3 are presented in Table 7.11. This indicated a significant main effect of Age Group ($F(2,57) = 39.77$, $p < .001$). High frequency words were easier to read than low frequency words ($F(1,57) = 150.05$, $p < .001$). There was a significant main effect of Body Type ($F(2,114) = 4.72$, $p < .01$) and the interactions between Age Group and Frequency ($F(2,57) = 4.75$, $p < .01$), Age Group and Body ($F(4,114) = 3.87$, $p < .006$), and Frequency and Body ($F(2,114) = 21.93$, $p < .001$) were also significant. The full anova table can be found in Appendix A7.6.

Planned comparisons indicated that Age Group 3 performed significantly better than Age Group 2 ($t(57) = 2.04$, $p < .05$) and that Age Group 2 performed significantly better than Age Group 1 ($t(57) = 6.50$, $p < .001$). Looking at the planned comparisons for the body types, the ambiguous body type was significantly more difficult than the consistent bodies ($t(114) = 2.73$, $p < .01$); however children did not differ significantly on ambiguous and consensus heretic bodies ($t(114) = 0.15$, NS).

TABLE 7.11: MEAN CORRECTLY READ WORDS DIFFERING IN FREQUENCY AND BODY TYPE BY AGE GROUP

AGE GROUP	LOW F			HIGH F			$\bar{x}$
	Ambig	Consis*	Consens	Ambig	Consis*	Consens	
1	2.65	2.95	2.95	5.25	5.4	3.95	3.858
2	5.20	5.575	5.90	7.45	7.725	6.80	6.442
3	6.25	6.725	7.20	7.65	7.95	7.75	7.254
	4.7	5.08	5.35	6.78	7.025	6.17	

MAXIMUM SCORE = 8

* The average mean for the two low F (CONSISTENT 1 and CONSISTENT 2) and for the two high F (CONSISTENT 7 and CONSISTENT 8) consistent matches for ambiguous and consensus bodies were used. T tests showed that this grouping was permissible because there was no significant difference between CONSISTENT 1 and CONSISTENT 2 body types ($\underline{t}$ = 0.21, $p <$ 0.84) and none for CONSISTENT 7 and CONSISTENT 8 body types ($\underline{t}$ = 1.42, $p < 0.16$).

The interaction between Age group and Frequency arose because the frequency effect was most marked for the younger readers (t(57) = 8.84, $p < .001$), although it was still significant for Age Group 2 (t(57) = 7.74, $p < .001$) and Age Group 3 (t(57) = 4.64, $p < .001$).

Planned comparisons also indicated that the Group x Body interaction was due to the body effect being most marked for Age Group 1. Ambiguous words were more difficult than consensus heretic words (t(114) = 2.52, $p < .02$), which in turn were more difficult than the consistent words (t(114) = 3.6,) $p < .001$). For Age Group 2 none of the body comparisons were significant while for age Group 3 ambiguous words were more difficult than consensus heretic words (t(114) = 2.65, $p < 02$)

There was also an interaction between Frequency and Body Type. This was because body effects were more marked for high frequency than for low frequency words. High frequency ambiguous words were significantly easier than high frequency Consensus/ heretic words (t(228) = 3.97, $p < .001$), and consensus/heretic words were significantly more difficult than consistent words (t(228) = 5.53, $p < .001$). The most significant frequency effects were found for the consensus body type (t(114) = 5.38, $p < .001$).

For the low frequency words, low frequency ambiguous words were significantly more difficult than consistent words (t(228) = 2.47, $p. < 02$) and consensus/heretic words (t(228 = 4.19, $p < .001$). The largest frequency effects were found for the consensus body type (t(114) = 5.38, $p < .001$), although they were also present for the ambiguous body type ($t(114) = 3.17$, $p < .001$) and the consistent body type ($t(114) = 2.21$, $p < .05$).

High frequency words were easier, as predicted. However body effects were more marked for high frequency items, which is surprising, and not in accordance with the model.

ANALYSIS 4: Analysis of low frequency, low N bodies

Means for correctly read words differing in body type are presented in Table 7.12.

TABLE 7.12: MEAN CORRECTLY READ LOW FREQUENCY, LOW N WORDS DIFFERING IN BODY TYPE, BY AGE GROUP

AGE GROUP	AMBIG	CONSIS*	CON/HER	GANG/H	GANG	$\bar{x}$
1	2.65	2.9375	2.95	2.25	2.4	2.6375
2	5.20	5.325	5.90	3.75	4.6	4.955
3	6.25	6.7125	7.20	5.5	6.4	6.4125
	4.7	4.992	5.35	3.83	4.47	4.47

MAXIMUM SCORE = 8

* The average mean of the four matched consistent groups (CONSISTENT 1-4) for low Frequency, low N Ambiguous, Consensus/Heretic, Gang/hero and gang bodies was used in this analysis because there were no significant differences found in Analysis 1.

An analysis of variance in which Age Group was a between subjects factor and Body Type a within subjects factor was performed. This indicated a significant main effect of Age Group ($F(2,57) = 31.03$, $p < .001$). There was a significant effect of Body type ($F(4,228) = 20.05$, $p < .001$), but no Group and Body Type interaction. A full anova table can be found in Appendix A7.7.

Planned comparisons indicated that gang/heroes were the most difficult body type; significantly more difficult than gang bodies ($t(228) = 3.51$, $p < .01$). There was no significant difference between gang bodies and ambiguous bodies, and the trend was not in the predicted direction i.e. the gang body words were harder. Ambiguous words were found to be no more difficult than the consistent body type; the consistent body type was more significantly difficult than the consensus/heretic body type ($t(228) = 1.98$, $p < .05$).

Thus there was a body effect when N was constant, but not the body effect anticipated.

ITEM ANALYSIS FOR LOW FREQUENCY, LOW N WORDS (comparable to subject analysis 4)

This was carried out for the low frequency low N words alone because regularity effects have only been found for low frequency words in the item analysis Experiment 7, as was also found by Seidenberg et al (1985); and to determine whether the effects noted in the fourth subject analysis are robust when words are treated as a random variable.

Difficulty values for each word were calculated as described in previous experiments and logistic modelling techniques were again applied. Table 7.13 shows the raw means and mean proportion correct for body types - the most difficult body type has the lowest p value. It also shows the raw means for the four matching consistent body types,

TABLE 7.13: RAW MEANS AND MEAN PROPORTION CORRECT FOR LOW FREQUENCY LOW n BODY TYPES

BODY TYPE		MEAN NUMBER OF CHILDREN WHO READ WORDS CORRECTLY				NO. CORRECT EXPRESSED AS PROPORTION p
		GP1 (n=20)	GP2 (n=20)	GP3 (n=20)	OVERALL	
AMBIGUOUS	$\bar{x}$	6.75	12.75	15.75	11.75	.59
	p	0.34	0.64	0.79	0.59	
CONSISTENT	$\bar{x}$	7.281	13.313	16.75	12.448	.62
	p	0.36	0.67	0.84	0.62	
CONSENSUS/HERETIC	$\bar{x}$	7.375	14.75	17.875	13.33	.67
	p	0.37	0.74	0.89	0.67	
GANG/HERO	$\bar{x}$	5.625	0.375	13.0	0.33	.47
	p	0.28	0.47	0.65	0.47	
GANG	$\bar{x}$	5.875	11.5	16.0	11.125	.56
	p	0.29	0.58	0.80	0.56	

RAW MEANS FOR FOUR MATCHING CONSISTENT GROUPS

CONSISTENT (1)	$\bar{x}$	7.75	14.25	16	12.67	.63
	p	0.39	0.71	0.80	0.63	
CONSISTENT (2)	$\bar{x}$	6.875	13.625	17.5	12.67	.63
	p	0.34	0.68	0.88	0.63	
CONSISTENT (3)	$\bar{x}$	6.875	13.0	16.5	12.125	.61
	p	0.34	0.65	0.83	0.61	
CONSISTENT (4)	$\bar{x}$	7.625	12.375	17.0	12.33	.62
	p	0.38	0.62	0.85	0.62	
GRAND MEAN		7.281	13.313	16.75		

indicating that, in terms of difficulty, they are a highly homogenous consistent set. Therefore the means of the consistent bodies were averaged for the analysis.

Estimates, adjusted for all predictors in the model, were obtained by the logistic modelling technique available in GLIM (Baker and Nelder, 1978).

The variables used in the regression equation were N - the number of neighbours a target word had, Word Frequency, Body Size and Body Type. They accounted for 25% of the total proportion of the variance ($R^2 = 0.250$). R^2 and χ^2 values for significant difference in explained variance are reported. The variables that were significant in this model were Frequency ($R^2 = 0.177$, $\chi^2 = 65.6$, $p < 001$) and Body Type ($R^2 = 0.107$, $\chi^2 = 39.5$, $p < .001$) - the fact that N was not significant is not surprising since the N values were all low.

Looking more closely at the Body Types it appears that the gang/hero body type was the one which differed most from the other body types. It was significantly more difficult than the gang ($t(59) = 2.77$, $p < .02$), the ambiguous body type ($t(59) = 3.74$, $p < .001$), the consensus/heretic ($t(59) = 6.21$, $p < .001$) and consistent bodies ($t(59) = 6.16$, $p < .001$). Significant differences for other body types are set out in Table 7.14.

The pattern of body types effects observed in the fourth subjects analysis remained robust in the item analysis.

TABLE 7.14: COMPARISONS BETWEEN LOW FREQUENCY, LOW N BODY TYPES

	GANG/HERO	CONSISTENT	CONSENSUS/ HERETIC	AMBIGUOUS	GANG
GANG/HERO	NS	t 6.16 df 59 p .001	t 6.21 df 59 p .001	t 3.74 df 59 p .001	t 2.77 df 59 p .01
CONSISTENT		NS	t 1.80 df 59 NS	t 1.40 df 59 NS	t 2.65 df 59 p .02
CONSENSUS/ HERETIC			NS	t 2.53 df 59 p .02	t 3.50 df 59 p .01
AMBIGUOUS				NS	t 0.99 df 59 NS
GANG					NS

An analysis of the error data showed many examples of regularisation and irregularistion of the body types, for the three age groups. These are presesnted in Table A7.15. It can be seen from the table that the instances of regularistion and irregularisation errors increased with age. The mean proportion of irregularisations was 0.15 for the first years, 0.27 for the second years and 0.45 for the third years. The proportion of regularisations was 0.14 for the first years, 0.36 for the second years and 0.53 for the third years.

DISCUSSION

The first analysis investigated if there was a difference in the reading performance for the four, low frequency, low N consistent body types, which were used as controls matched for frequency and N with the irregular body types. this was done so as to be able to use the mean difficulty value for low frequency, low N consistent bodies in later analyses.

The second analysis investigated the N effect for low frequency words differing in body type. It was hypothesised that there would be a body effect with consistent words being easier to read than consensus/heretic bodies, if only the body subsystem were used. Both the Dual Route and Modified Standard model would not predict this significant different.

Also if the body subsystem were used ambiguous bodies would be the most difficult of all. Ambiguous words were the most difficult to read, but there was no significant difference between consistent and consensus body types. Size

TABLE A7.15: NUMBER OF REGULARISATIONS AND IRREGULARISATIONS FOR LOW FREQUENCY, LOW N WORDS DIFFERING IN BODY TYPE FOR THE THREE AGE GROUPS (ALSO EXPRESSED AS A PROPORTION)

	LOW FREQUENCY, LOW N BODY TYPE			
	AMBIGUOUS		CONSENSUS/ HERETIC	
YEAR 1		p		p
IRREGULARISATIONS	15	0.14	16	0.16
OTHER	92	0.86	85	0.84
TOTAL	107		101	
YEAR 2				
IRREGULARISATIONS	18	0.33	8	0.20
OTHER	36	0.67	33	0.80
TOTAL	54		41	
YEAR 3				
IRREGULARISATIONS	19	0.56	5	0.33
OTHER	15	0.44	10	0.67
TOTAL	34		15	

	GANG/HERO		GANG
YEAR 1			
REGULARISATIONS	16	0.14	NO
OTHER	99	0.86	REGULARISATION
TOTAL	115		ERRORS
YEAR 2			POSSIBLE
REGULARISATIONS	30	0.36	BECAUSE
OTHER	54	0.64	ALL EXAMPLES
TOTAL	84		ARE PRONOUNCED
YEAR 3			IRREGULARLY
REGULARISATIONS	27	0.53	
OTHER	24	0.47	
TOTAL	51		

of neighbourhood was a significant variable in the reading task (i.e. high N words were easier to read), but this effect was confined to consistent bodies. This seems intuitively correct; the more examples of a consistent set there are the more "consistent" it is. However one would have expected body size to affect consensus/heretic bodies, especially if the heretic were high frequency.

The third analysis investigated the effects of frequency on low N words of differing body type. It was predicted that any effects of body type would be more marked for low frequency words; this was not shown in the data. High frequency ambiguous words were significantly easier than high frequency consensus/heretic words, with consistent words being the easiest. It is noted that this is not the order of difficulty implied by Patterson and Morton. They would have expected ambiguous words to be the most difficult as they are the most inconsistent. In this analysis there was a significant Age Group x Body interaction. However, body effects were most marked for Age Group 1 - the youngest readers, and not in the predicted order of difficulty. It had been predicted that younger readers would only be affected by frequency. The frequency effect was most marked for this age group. What was surprising was that the body effect was also most marked for this age group. Body effects were not significant for Age Group 2, and less marked for Age Group 3. This appears to detract from Marsh Frith & Seymour's developmental model and implies that logographic and orthographic strategies differ quantitively rather than qualitatively.

Analysis 4 investigated the body effect for low frequency, low N words. An investigation of a wider range of bodies was possible in this analysis - body types included were ambiguous, consistent, consensus/heretic, gang/hero and gang bodies. There was a significant body type effect, although the order of difficulty was not as clear cut as that predicted by the modified standard model. Gang/heroes were significantly more difficult than gangs, which were more difficult (but not significantly so) than ambiguous words. Ambiguous words were found to be no more difficult than the consistent body type, with the consensus heretic body type being the easiest. However the results were considered to follow closely enough the order of difficulty predicted by Patterson and Morton to give credence to their theory. It may well be that the consensus heretic word set in this particular word list was easier, but for reasons other than body type: differences in Imagery or familiarity may be. This brings up the question of why a non-lexical routine would be affected by imageability or familiarity. Although it was assumed that children used non-lexical routine because the stimulus items were low frequency words, there is no guarantee that this was the case. A nonword stimulus set would be required to ensure the use of a non-lexical route, but this would bring attendant problems of relevance to the reading process and of difficulty scoring right or wrong responses.

It is possible that lexical and non-lexical routes are not totally isolated from each other, and that an interaction of the two might explain the suggested importance of imgeability in the reading task.

The item analysis, which confirmed the robustness of effects in the fourth subjects analysis just discussed, also provided evidence of a body effect, with the gang/hero body type differing most from other body types. This is consistent with Dual Route Model predictions, but undermines the Modified Standard Model prediction that the ambiguous body type would be the most difficult. They were not different from gang bodies and consistent bodies.

From the analyses reported a clear effect of body type on reading performance was observed, although it was not always in the predicted direction and was unexpectedly present in high frequency items.

The analysis of the N effect on low frequency words also showed an effect of neighbourhood size on reading which did not disappear when body type was manipulated. Thus probabilistic descriptions of regularity - N, and rule-based measures - body type were supported by the data.

Evidence that performance on bodies would differ according to level of reading as described by Marsh et al (1981), with more skilled readers using body type information, was not obtained. What was observed was that even young readers of seven to eight years showed differences in reading performance due to the body effect.

The pattern of performance on bodies was the same for the three age groups. Gang/heroes were the most difficult body type; significantly more difficult than gang bodies. Gang and ambiguous bodies did not differ significantly - gang body words being marginally harder. ambiguous words were no more difficult than the consistent body type; the consensus/heretic body type was the easiest to read. Tentative conclusions can therefore be drawn as to which of the body representations appear first in the OPC system, and which appear last. It is suggested from the results that regular body types - consistent and consensus/heretics, are among the first body type representations to appear in the OPC system, followed by ambiguous, gang and gang/hero body type representations.

The analysis of the error data, however, did show that regularisation and irregularisation errors increased with age, for the young readers 85% of their errors were in the "other" category which included visual errors, some semantic errors and refusals. This does indicate that although the body effect was less marked in the older readers, there is an awareness of consistency, as shown by the increase in regularisation and irregularisation errors, whereas young readers are attending to visual, phonemic and body type information in the reading task.

The results of Experiments 7 and 8 showed that the regularity effect is too complex to be described in terms of a regular-irregular dichotomy. More detailed descriptions of regularity such as typicality of divergence and the body sub-system are useful in trying to explain experimental results. Even so, the continuum may be even more complex than the Modified Standard Model implies.

In summary, the main points found in both experiments showed:-

a) the far from straightforward body effect and the fact that it is particularly pronounced for high frequency items. This undermines the notion of separable routines.

b) The facilitatory effect of high frequency hostile neighbours.

c) Both probabilistic and rule-based measures of regularity were supported by experimental results.

d) The developmental pattern observed - even young readers showed differences in reading performance due to the body effect.

CHAPTER EIGHT

SUMMARY AND CONCLUSION

8.1 INTRODUCTION

The purpose of the research reported in this thesis was to investigate the changes in knowledge of word characteristics as reading ability improves. The word characteritics investigated were:-

a) Imagery and Age of Acquisition

b) Orthographic regularity - probabilistic measures such as ONR and N

c) Orthography-to-phonology regularity - rule-based measures such as bodies

The following principle findings were observed:-

Firstly Imagery was found to be a highly significant word characteristic for young readers. In Experiment 1 it explained a large proportion of the variance for all reader groups in the lexical decision and reading task, and was most significant for the less skilled readers. In further investigations, Experiments 2 and 3 which directly observed the effects of word imagery and Age-of-Acquisition in reading, it was found that there was an Age-of-Acquisition effect which became less important for skilled readers and that there was an Imagery effect which was only significant for poor readers.

Skilled adult readers showed a small Age of Acquisition effect, but the Imagery effect had completely disappeared. These results support the suggestion made by Baddeley et al (1982) and Frith (1985) that age of acquisition of words would affect children's reading performance. It is rather surprising that the large effects on children's reading are produced by this conceptually vague measure. However it does appear to predict children's accuracy in reading. The results also showed that the Imagery effect was not mediated by word age of acquisition.

The question remains as to why word imagery affects raeding accuracy in young readers. It may be that the acquisition of word recognition units or beginning readers is determined by the pre-existence of entries in the cognitive system. If words are low imagery, they would be less likely to have a reasonably adequate entry in the cognitive system, and therefore it would be less likely that young readers have acquired word recognition units for them. Jones (1985) provided a suitable framework for the imagery effect in poor readers when he described the concept of ease-of-predication, in which a high imagery word may be represented by a complex cognitive network linking up predicates for that word. Although Jones noted the act that Imagery is potent variable, he also noted that it has virtually no explanatory power. Again, the question remains as to whether "ease-of-predication" has any greater explanatory power, and how this network of predicates can be experimentally investigated.

Secondly a large range of variables which measured orthographic structure were compared, (probabilistic measures as defined by Venezky, Massaro & Taylor; 1979) so as to see which of them was the best predictor of reading and lexical decision performance. The Orthographic Neighbour Ratio, first Order of Approximation to English, Initial Bigram Frequency and Initial Bigram Versatility appeared to be the orthographic measures which best predicted reading performance. The hypotheses made as to the orthographic information that would be used by children of differing reading ability was supported by the data collected. Probabilistic measures such as Initial Bigram Frequency and Versatility and First Order of Approximation to English were used by all reading ability levels. However as reading ability improved responsiveness to a word's Orthographic Neighbour Ratio increased, in the lexical decision task. Thus it appears that the Orthographic Neighbour Ratio, a previously unexplored variable, is important in predicting lexical decision accuracy. It takes into account the number of orthographic neighbours that a particular word has and their frequencies.

It is essential to stress the difference between orthographic regularity defined by probabilistic measures and regularity as defined by the dual route model. Orthographic regularity has been defined as the allowable patterns of letters within single words. It can be described using a probabilistic approach e.g. bigram frequencies, First Order Approximation to English; or a rule-governed approach, which attempts to define orthographic structure in terms of the more general linguistic patterns of English spelling e.g. Wijk (1966), and Venezky (1970). Regularity, as defined by the dual route model, can only be described by the spelling-to-sound correspondence rules elaborated by Venezky (1970).

It was considered important to investigate more closely the effects of orthographic regularity, as defined by probabilistic measures, because previous research by

Coltheart, Laxon, Keating and Pool (1986) and Seidenberg et al (1984) had shown that regularity, as defined by grapheme-phoneme correspondence rules was a less reliable effect than had been previously thought, limited only to low frequency words.

Changes in reading accuracy when orthographic neighbourhood size (as measured by N) was manipulated, were also observed as reading ability improved. N is a more general measure of orthography as it only takes into account the number of neighbours of a letter string, and not their frequency, unlike the Orthographic Neighbour Ratio. It does not differentiate between regular and irregular words, as defined by grapheme-phoneme rules and words similar in body endings, as described by Patterson and Morton (1984). Results showed that there was an effect of the size of orthographic neighbourhood for children in a variety of tasks - words with large orthographic neighbourhoods (i.e. high N words) were easier to read than were low N words, which had very small orghographic neighbourhoods. Poor and average readers used this orthographic information, but the N effect was less pronounced in the lexical decision and reading performance of the good reader group, and was no longer present for skilled adult readers.

The final experimental chapter manipulated both probabilistic <u>and</u> rule based measures of orthographic regularity for low frequency words, as Seidenberg et al (1984) had observed regularity effects for low frequency words only; i.e. both neighbourhood size and body type were manipulated. The results showed that the 'regularity' effect was not just due to a regular-irregular dichotomy. An exploratory study showed that regularity is dependent upon the hostility and frequency of the word neighbour. When regular and irregular words had no hostile neighbours, then no irregularity effect was found. For low frequency words with many high frequency, hostile neighbours, there was no regularity effect. For words which existed in a mainly friendly neighbourhood, those with high frequency hostile neighbours were more easily read than words with low frequency hostile neighbours. The regularity effect was also dependent upon the frequency of the target word itself - it was only significant for low frequency words. In the final experiment which manipulated body type and orthographic neighourhood size - that covers neighbours not included in body type, both N effects and body effects were observed for low frequency words, N effects only being significant for the consistent body type. There was a change of effects of body type with change in age - body effects were most marked for yonger, poorer readers. It was not significant for nine year olds, and less marked for ten year olds. However an analysis of the error data showed that although body effects were most significant for younger readers, they made a great variety and number of errors that could be categorised in the "other" category; whereas there was a clear increase in regularisation and irregularisation of words according to body type, as age and reading ability increased.

8.2 IMPLICATIONS OF EXPERIMENTAL FINDINGS FOR DEVELOPMENTAL MODELS

These findings do not relate in a clear cut way to the developmental models of Marsh, Friedman, Welch and Desberg (1981), Frith (1985) and Seymour and MacGregor (1984). Marsh et al described a sequence of four stages which they suggested were part of the developmental procedure for most normal readers - linguistic guessing, discrimination net guessing, sequential decoding and hierarchical decoding. Frith (1985) adapted this to three stages, identified with three strategies, the logographic stage, the alphabetic stage and the orthographic stage.

She suggested that spelling development may proceed through similar stages, though somewhat out of phase. Seymour and MacGregor (1984) described these stages in an information processing model - the Dual Lexicon model. As implied it has a logographic lexicon and an orthographic lexicon. It is the orthographic lexicon which deals with orthographic information, either the whole word, subsets such as body units, bigrams or single graphemes.

The reader groups of varying ability in each experiment were located at a particular stage in the developmental model, described by Marsh, Friedman, Welch and Desberg (1981) and elaborated by Frith (1985) and Seymour (1984) in terms of their reading performance.

None of the reader groups seemed to be logographic because the pattern of performance did not suggest linguistic guessing or discrimination net guessing approach. Therefore it was assumed that they must either be in the alphabetic stage or the orthographic stage, or in stages of transition between these, or stages within the orthographic stage.

In this way, the applicability of the stages in the developmental model were assessed, in view of the experimental findings.

The effects of the first two word characteristics studied, Imagery and word Age of Acquisition, were most significant for the less skilled readers (CA 9:11, R.A. 7:11). These word characteristics became less important by the orthographic stage - the AOA effect decreased with age and the Imagery effect was significant only for poor readers. There was no AOA or Imagery effects in adults. The developmental models made no specific predictions about these two variables. The experimental results suggest that they are more important in the logographic than the orthographic stage.

It was predicted that orthographic variables such Initial Bigram Versatility and frequency would be used by readers in the alphabetic stage and more complex measures such as Orthographic Neighbour Ratio would be used in the Orthographic Stage. This prediction was supported by the data collected - as reading ability improved so did task accuracy when a word's Orthographic Neighbour Ratio was manipulated, at least for the lexical decision task. However it was noted that even poor readers (C.A 9:1, R.A 7:6) reading performance was affected when probabilistic measures of orthography were manipulated. There were very strong Imagery effects which tended to partial out effects of orthographic regularity.

When a model was fitted, excluding Imagery, the Orthographic Neighbour Ratio was a significant variable in the lexical decision, but not in the reading task for good readers (C.A 8:8, R.A 9:5) and average readers (C.A 9:2, R.A 8:5). This result is consistent with developmental predictions made - that in the orthographic stage (Marsh et al's Stage 4) orthographic variables assumed greater importance, and that an orthographic lexicon would have developed. However results also showed that poor readers used probabilistic measures of orthography such as Initial Bigram Frequency and First Order Approximation to English, indicating that even in the logographic stage, there was a certain amount of orthographic awareness.

Another probabilistic measure of orthographic regularity, neighbourhood size, showed that poor readers (C.A 9:6, R.A 7:6), and average readers (C.A 9:5, R.A 8:6) found it easier to read high N words than low N words, whereas good readers (C.A 9:7, R.A 10:1), assumed to be in the Orthographic stage were less affected.

Results obtained in a preliminary experiment using a rule-governed measure of orthographic regularity - body type, again showed that the reading accuracy of poor readers (C.A 8:3, R.A 8:2) was affected by manipulation of orthographic regularity as defined by body type, or low frequency words, but that this was dependent upon the frequency of the target word and hostility and frequency of word neighbours.

Evidence that performance on bodies would differ according to level of reading as described by Marsh et al (1981), with more skilled readers using body type information, was not obtained in an experiment which manipulated body type and neighbour hood size. Although there was a body effect for older readers (C.A 10:4, R.A 9:1), it was not present for younger readers (C.A 9:2, R.A 8:4) and most significant for the youngest readers (C.A 7:8, R.A 7:5).

It is suggested that the results are less contradictory than they appear. Effects of specific probabilistic measures of orthographic regularity are very robust and do tie in with developmental theories of reading. However more general measures of orthographic regularity such as body type may be susceptible to differences in task demands, and frequency of the target word itself. It may be that this type of orthographic information is mostly used for less frequent words and nonwords, in skilled readers and adults - a suggestion that should be followed up in future research.

In the following section 8.3, the impliction of the principal experimental findings are discussed and an assessment of theoretical models and their applicability to children's reading is made. The implications of these findings for teaching are set out in section 8.4 and finally suggestions for future research based upon the experimental work reported are discussed in Section 8.5.

8.3 IMPLICATIONS FOR THEORIES OF ADULT READING AND THEIR APPLICABILITY TO CHILDREN'S READING

A number of theories of adult reading have been described - the standard Dual Route model (e.g. Coltheart, 1982; Morton & Patterson, 1980), analogy models (Glushko, 1979; Marcel, 1980), McClelland and Rumelhart's (1981) activation-synthesis approach and Seidenberg's time-course model (1984), Henderson's (1982) lexical pooling model, the Multiple Levels position (Shallice, Warrington & McCarthy, 1983) and the Modified Standard Model (Patterson & Morton, 1984). They all attempt to explain the features of orthographic structure which are used in the reading process. These theories can be regarded as being on a continuum for several reasons. Firstly some theories describe units in the reading process in terms of abstract grapheme-phoneme correspondences. Others describe whole word units; and yet other models describe a combination of these e.g. Shallice et al's multiple levels model.

Secondly word analogies, considered important by analogy theories require only one examplar, where rules are derived from many examplars.

The critical features for these models also overlap. The standard dual route model considers that word frequency is an important variable in the reading process, as does the modified dual route model and Seidenberg's model. However the modified dual route model considers that inconsistency of bodies causes difficulty for readers, and is not specific about whether size of body membership is an important factor. Henderson's lexical pooling model suggests that the essential features are the size of orthographic neighbourhood, the remoteness of neighbours and the size of shared segments. Analogy theory predicts that the size of body membership and inconsistency of bodies are critical features for ease or difficulty of reading.

The experimental results showed that regularity of spelling-to-sound correspondence appears to be a continuum rather than a dichotomy, and that reading difficulty varies according to where on the continuum a word is. This implies that studies whose results are based on effects averaged over irregular words in general may therefore be misleading. Theories which consider a more complex form of orthographic regularity e.g. analogy theory, Multiple-Levels Model, lexical pooling models and the Modified Standard Model will therefore be more appropriate for describing children's reading.

Word frequency is a very important variable in children's reading and appears to interact in its influence on reading accuracy with other relevant dimensions such as regularity of orthography, size of neighbourhood etc. As Seidenberg et al (1983) concluded, high frequency words are insensitive to most other dimensions. The interesting finding is that regularity effects are much more marked in low frequency words. This implies that children are using their knowledge of orthography-orthographic neighbourhood size, frequency and hostility of neighbours in reading low frequency, unknown words. The models which made specific predictions about frequency and regularity are Seidenberg's Time Course Model (1984) and the Modified Standard Model (1981). These are therefore considered the more appropriate for describing frequency and regularity effects in children's reading.

It was shown that the ease with which a word is read is influenced not only by its own position on the regularity-irregularity continuum (although some experimental results were contradictory i.e. the ambiguous body type were harder than the irregular gang body type) but by the extent to which its neighbours agree or conflict in pronunciation. Glushko brought attention to this consistency effect in 1979. However, like regularity, it has been shown in these experiments that different patterns and degrees of consistency influence performance in children. The theories of direct relevance in this instance are Shallice and Warrington's Multiple Levels position with its concept of typicality of divegence, and Patterson and Morton's (1984) modified standard model. There was clear experimental evidence that children were affected by body type and frequency and hostility of neighbours. It is therefore considered that the modified standard model is of great relevance to children's reading performance.

It has been suggested that the distinction between the small GPC units of the dual route theory and the large orthographic units of the lexical analogy models is no longer clear cut. The translation unit which has been theoretically favoured by modified dual route theory and analogy theory is the body sub-unit, and experimental evidence of its use by children in a reading task has been presented. However the experiments results lead to the same conclusion as that of Henderson:-

> "... the recent accumulation of evidence points incontrovertibly to a broad range of size."
> (1985a, p495)

8.4 IMPLICATIONS FOR TEACHING

The teaching of reading in schools today involves both the phonics approach and the "look and say" method. In the phonics approach, the units typically stressed are the simple vowel patterns (i.e. single letter), the vowel digraph (ee, ea, ow, etc), sequences such as WH-, QU-, -DGE, and -TCH, the common (and not so common) initial and final consonant clusters, and the common prefixes and suffixes - all of which play a role in rule-governed regularity. The way in which patterns are introduced in the phonics approach is also important. A spelling such as EE is usually introduced alone (with its most common pronunciation) and then in a group of words divided by position. Thus SEE, FREE and BEE might be grouped, then SEEK, BEET, SEED, and so on, with EE emphasised by underlining or colour coding. Another method of presentation that tends to involve orthographic regularity centres on introducing common vowel-consonant or consonant vowel sequences that are produced for word building e.g. the -AN family is introduced: FAN, TAN, MAN, VAN and so on. In addition to emphasising particular letter-sound patterns, this practice encourages children to think in terms of families of words i.e. word neighbours, and encourages them to use measures of regularity other than that defined by grapheme-phoneme correspondence rules. Henderson (1982) pointed out that in the emphasis of grapheme-phoneme correspondences (or correspondeces of larger segments) the visual constitutents of word are also drawn attention to. He considered that it is important because the construction of a defined orthographic entry in the lexicon is required by the analogy model advanced by Glushko (1979) and Marcel (1980), at least implicitly.

The proponents of the "look-and-say" method of teaching reading suggest that the child's first attempts at acquiring a sight vocabulary of words often resemble paired-associates learning. They assume that the word is treated as a unitary symbol without attention to its constituent elements. Critics of the approach consider that it is not always effective, leading to an early acquisition of a small sight vocabulary and then little progress beyond this.

However as early as (1967) Chall and others were advocating a more eclectic approach and nowadays most teachers, at least at junior level tend to use both methods. Bradley and Bryant (1985) recommended two practical teaching methods. The first of the two methods is a more explicit version of the one used in the phonics approach. The child puts words into categories on the basis of their common sounds, and to relate these categories to particular spelling sequences using alphabetic plastic letters. Their second method shows the reader how the different strategies which he uses when he reads and when he spells are linked and often interchangeable, using "multi-sensory" teaching - asking children to trace round each letter of a letter string with a finger. The technique involves visual information, auditory/orthographic information (spelling out the letters) and writing movements. It had been shown to be particularly helpful with backward readers (Hulme, 1981).

The experimental findings reported in this thesis suggest that aspects of look and say or phonics teaching that emphasise grouping words in ways which call attention to "regularities" in the orthographic structure other than that defined by grapheme-phoneme-correspondence rules will be helpul because children appear to use variables other than the whole word or grapheme-phoneme correspondences. It is therefore important to call attention to regularities other than phonic; i.e. larger units of phonics e.g. bodies and to be aware of the effects of Imagery and Age of Acquisition.

8.5 SUGGESTIONS FOR FUTURE RESEARCH

Imagery has been shown to affect reading accuracy in poor young readers and AOA has been shown to affect reading accuracy for all reading ability levels tested. However neither of these variables has much explanatory power, and even Jone's (1985) concept of ease-of-prediction in which a high imagery word may be represented by a complex cognitive network does not make clear why it is the poor readers that are affected. It is possible that poorer readers may have fewer low imagery words represented in semantic memory. Thus low imagery words may lack entries into the phonological lexicon or they may lack entries the cognitive system which represents word meanings. Alternatively they may be represented in both systems, by indequately represented in the cognitive system. If the acquisition of word recognition units is dependent on the prior existence of an adequate entry in the cognitive system, then poor readers would be less likely to acquire word recognition units or low imagery words, and they cannot read words they don't comprehend by applying grapheme-phoneme correspondence rules, because they find this non-lexical procedure difficult. (Perfetti and Hogaboam, 1975). It is assumed that the acquisition of word recognition units for beginning readers is determined by the pre-existence of entries in the cognitive system. Future experiments would have to provide evidence to the contrary - that poorer readers can read irregular words that they do not understand.

The experiment on orthographic neighbourhood size suggested a developmental trend in which knowledge about orthographic neighbourhoods is used by poor and average readers, less important in good readers and not used by adult readers. An experiment could be conducted to investigte if this orthographic information is only required by adults for dealing with unfamiliar nonwords.

An exploratory study investigating the effects of regularity (as defined by Patterson & Morton (1984) body sub-units) showed that regularity is dependent upon the hostility and frequency of a target word's neighbours. The frequency of word neighbours was not investigated in Experiment 8 because it was not relevant to all the word sets used in the word list. However, it is considered that it is an important variable in the reading process - most of the adult models of word recognition have not made any specific predictions about frequency of neighbours, although Henderson (1982) considers that remote neighbours are unhelpful; though it is unclear whether this is syonymous with low frequency.

The final experiment showed that poor readers were affected by consistency as measured by Patterson and Morton's body types, but that the effect was less marked for skilled readers. In order to strengthen these results it would be necessary to look at the naming latency data, and to perform an error analysis.

We can conclude our look at adult models of word recognition by summarising our opinion of their relative merits. Even if a firm choice among the categories of models could be made, there are many points on which uncertainty would still exist and research undertaken on the details. Each of the major approaches has both strengths and weaknesses relative to the other approaches. Nevertheless, some general constraints have been identified and it is considered that further research within the framework provided by the Modified Standard Model (Patterson and Morton, 1984) would prove fruitful for a more detailed description of children's reading.

APPENDICES

WORD & NONWORD LIST - EXPERIMENT 1

Length	No.	WORDS	NONWORDS
4 LETTERS	1	body	doby
	2	city	ticy
	3	girl	lirg
	4	idea	eadi
	5	time	miet
	6	army	ramy
	7	bowl	wolb
	8	fact	caft
	9	form	morf
	10	rock	korc
	11	cane	neca
	12	cord	droc
	13	fork	korf
	14	frog	grof
	15	oven	enov
5 LETTERS	16	child	dilch
	17	ocean	enoca
	18	paper	arpep
	19	sugar	grusa
	20	table	balte
	21	chair	rachi
	22	month	nomth
	23	truck	kruct
	24	water	retaw
	25	woman	wanom
	26	hound	hunod
	27	peach	epach
6 LETTERS	28	answer	wensar
	29	friend	nefrid
	30	method	thomed
	31	moment	temmon
	32	square	quares
	33	bottle	totleb
	34	cattle	cletta
	35	flower	werlof
	36	insect	cestin
	37	ticket	tiktec
	38	beaver	veerab
	39	butter	turteb
	40	kettle	tetlek
	41	master	strame
	42	string	grinst
7 LETTERS	43	freedom	dreemof
	44	history	stohiry
	45	machine	hinecam
	46	picture	turpice
	47	thought	tohguth
	48	bravery	ravbery
	49	goddess	gosdeds
	50	product	prudoct
	51	robbery	rebobry
	52	scarlet	scerlta
	53	mastery	stamery
	54	settler	stetler
	55	slipper	perplis
	56	steamer	mestara
	57	thicket	teckith
8 LETTERS	58	building	idulbing
	59	industry	distunry
	60	interest	risteent
	61	position	stipooni
	62	railroad	drolaria
	63	contents	nectonts
	64	contract	trancoct
	65	friction	fictirno
	66	aptitude	duaptite
	67	vocation	vionatoc
	68	christmas	shractsim
	69	direction	crenidito
9 LETTERS	70	knowledge	geldknowe
	71	newspaper	prawspene
	72	vegetable	gateveble
	73	amplifier	flempiari
	74	fisherman	sherminaf
	75	gentleman	telnamgen
	76	infection	feniction
	77	policeman	mepicolan
	78	blandness	slendnabs
	79	performer	morferper

EXP1. DAT ITEM ANALYSIS APPENDIX A4.2

WORD	FREQUENCY			correction to 2 decimal places			INITIAL LETTER	
	K+F	A.H	I	FOA	SOA	ONR	f	v
body	0276	1783	6.40	-26.49	-22.67	0.9883	45529	2446
city	0393	1843	6.43	-24.19	-20.11	0.9470	47141	3849
girl	0220	1084	6.87	-24.24	-26.21	0.9865	16,755	1349
idea	0195	1397	2.20	-21.25	-22.13	1.0000	61,892	1659
time	1599	8441	4.13	-21.63	-16.82	0.9552	58,774	1970
army	0132	0412	6.53	-24.97	-10.92	0.5197	91,711	2555
bowl	0023	0325	6.30	-26.31	-25.01	0.5000	45529	2446
fact	0447	0925	2.20	-23.59	-21.30	0.4928	40,348	1694
form	0370	2720	4.30	-24.57	-21.03	0.4901	40,348	1694
rock	0075	0925	6.37	-26.21	-23.62	0.5597	25,769	2273
cane	0012	0118	6.43	-21.56	-17.36	0.0090	47141	3849
cord	0006	0115	6.03	-23.54	-18.37	0.0086	47,141	3849
fork	0014	0122	6.57	-26.66	-20.69	0.0088	40,348	1694
frog	0001	0171	6.73	-25.00	-22.76	0.0002	40345	1694
oven	0007	0128	6.40	-23.29	-17.95	0.0025	71,249	0934
child	0213	0730	6.50	-28.71	-21.96	0.9103	47,141	3849
ocean	0034	0843	6.77	-25.55	-23.66	1.0000	71,249	0934
paper	0157	2372	6.30	-28.25	-23.35	0.9290	38,525	3073
sugar	0034	0574	6.57	-28.92	-25.44	1.0000	67,966	4504
table	0198	1502	6.50	-27.13	-22.37	0.9340	58,774	1970
chair	0066	0421	6.63	-29.05	-24.51	0.5238	47,141	3849
month	0130	0803	4.37	-26.92	-20.96	0.5817	38,759	2160
truck	0057	0410	6.60	-31.42	-24.24	0.4634	58,774	1970
water	0442	7194	6.60	-26.30	-20.32	0.5182	60,167	1083
woman	0224	0750	6.70	-28.68	-20.81	0.4686	60,167	1083
hound	0007	0044	6.27	-28.12	-20.66	0.0075	53,423	1511
peach	0003	0076	6.60	-27.87	-24.55	0.0073	38,525	3073
answer	0152	2002	2.77	-30.87	-28.31	1.0000	45,527	2446
friend	0133	0923	6.57	-31.39	-26.26	1.0000	40,548	1694
method	0142	0626	2.63	-31.07	-24.33	1.0000	38,759	2160
moment	0246	0834	2.50	-31.23	-25.40	1.0000	38,751	2160
square	0143	0965	6.37	-36.28	-27.35	1.0000	67,766	4504
bottle	0076	0346	6.57	-30.91	-31.65	0.4606	45,527	2446
cattle	0097	0658	6.40	-29.76	-29.47	0.4947	47,141	3849
flower	0023	0283	6.57	-33.24	-29.16	0.5475	40,348	1694
insect	0014	0202	6.10	-29.59	-26.47	0.4118	61,892	1659
ticket	0016	0100	6.20	-32.63	-24.02	0.5517	58,774	1990
beaver	0003	0096	6.50	-33.00	-24.62	0.2000	45,529	2446
butter	0027	0275	6.57	-31.82	-27.70	0.0525	45,527	2446
kettle	0003	0088	6.23	-31.54	-33.90	0.1071	04,834	0467
master	0072	0335	4.97	-29.98	-24.62	0.1702	38,759	2160
string	0019	0483	6.20	-31.35	-24.41	0.0538	67,966	4504

% OCCURRENCE

APPENDIX A4.2 cont.

WORD	TERMINAL LETTER f	TERMINAL LETTER v	INITIAL BIGRAM f	INITIAL BIGRAM v	TERMINAL BIGRAM f	TERMINAL BIGRAM ✓
body	058,623	3647	09.15	15.49	00.89	02.62
city	058,623	3647	02.76	02.47	00.51	01.05
girl	031,146	1549	09.95	06.97	00.00	00.00*
idea	005,190	0955	01.27	03.25	08.20	08.74
time	200,915	4938	01.79	08.24	02.34	00.39
army	058,623	3647	08.63	10.33	01.07	05.24
bowl	031,146	1549	09.15	15.49	00.00	00.00*
fact	091,728	2432	12.53	17.06	00.09	00.04
form	216,422	0695	37.08	18.77	05.45	00.18
rock	009,032	0520	11.16	12.67	00.03	00.21
cane	200,915	4938	19.22	17.51	11.28	04.61
cord	103,836	4779	46.59	42.96	01.52	00.24
fork	009,032	0520	37.08	18.77	00.00	00.00
frog	030,801	3111	17.96	15.41	09.64	04.58
oven	087,212	3264	02.36	16.17	07.74	07.28
child	103,836	4779	12.19	12.47	00.02	00.12
ocean	087,212	3264	00.91	05.67	00.49	01.99
paper	058,650	2110	17.22	19.07	42.09	14.01
sugar	058,650	2110	10.25	11.52	00.92	03.53
table	200,915	4938	02.05	12.91	01.43	02.04
chair	058,650	2110	12.19	12.47	00.19	01.19
month	025,885	0718	21.80	19.68	04.79	25.02
truck	009,032	0520	03.10	23.42	00.03	00.21
water	058,650	2110	25.53	25.67	42.09	14.01
woman	087,212	3264	11.70	15.33	00.49	01.99
hound	103,836	4779	10.88	23.30	00.04	00.57
peach	025,885	0718	13.12	13.93	00.00	00.00*
answer	058,650	2110	40.92	12.26	42.09	14.01
friend	103,836	4779	17.96	15.41	00.04	00.97
method	103,836	4779	17.50	17.36	11.75	06.10
moment	091,728	2432	21.80	19.68	00.00	00.00*
square	200,915	4938	00.46	01.18	21.97	16.07
bottle	200,915	4938	09.15	15.49	01.43	02.04
cattle	200,915	4938	19.22	17.51	01.43	02.04
flower	058,650	2110	03.81	15.35	42.09	14.01
insect	091,728	2432	54.43	68.12	00.09	00.04
ticket	091,728	2432	01.79	08.24	10.82	23.14
beaver	058,650	2110	39.79	18.77	42.09	14.01
butter	058,650	2110	15.21	13.45	42.09	14.01
kettle	200,915	4938	22.20	19.49	01.43	02.04
master	058,650	2110	32.44	53.06	42.09	14.01
string	030,801	3111	17.52	16.67	04.89	01.94

* [illegible]

APPENDIX A4.2 contd

WORD	FREQUENCY K-F	FREQUENCY A.H	I	FOA (corrected to 2 decimals)	SOA (corrected to 2 decimals)	ONR	INITIAL LETTER F	INITIAL LETTER V
freedom	0128	0386	3.83	-34.10	-33.69	1.0000	40348	1694
history	0286	0726	3.47	-35.95	-27.65	1.0000	53483	1511
machine	0103	0861	6.00	-35.69	-27.15	1.0000	58759	2160
picture	0162	2500	6.20	-34.95	-30.51	1.0000	38.525	3073
thought	0515	2835	2.77	-36.99	-29.70	1.0000	58974	1990
bravery	0004	0024	4.40	-40.17	-29.39	0.5000	45529	2446
goddess	0003	0053	5.23	-36.64	-36.89	0.6000	16255	1349
product	0087	0860	4.20	-38.60	-33.87	0.5148	38.525	3073
robbery	0010	0021	5.00	-39.66	-35.19	0.5882	25769	2273
scarlet	0003	0055	6.37	-34.54	-34.84	0.6000	67966	4504
mastery	0010	0018	2.77	-36.14	-27.78	0.1370	58759	2160
settler	0003	0020	5.40	-32.36	-35.47	0.0405	67966	4504
slipper	0003	0023	6.47	-37.45	-35.21	0.0750	67966	4504
steamer	0001	0034	6.53	-33.26	-29.48	0.1429	67966	4504
thicket	0001	0031	5.60	-37.19	-27.50	0.1429	58974	1990
building	0160	0886	6.40	-44.87	-33.86	1.0000	45529	2446
industry	0171	0429	5.77	-41.98	-32.50	1.0000	61892	1657
interest	0330	0560	3.13	-35.61	-28.63	1.0000	61892	1657
position	0241	0540	2.97	-38.76	-33.56	1.0000	38.525	3073
railroad	0058	0335	6.27	-39.33	-45.27	1.0000	25.769	2273
contents	0016	0067	3.57	-37.38	-31.06	0.5161	47141	3849
contract	0060	0094	4.50	-39.21	-33.40	0.4478	47141	3849
friction	0017	0091	4.33	-40.11	-32.82	0.4250	40348	1694
aptitude	0003	0004	2.60	-40.02	-37.79	0.0263	45529	2446
vocation	0003	0010	3.80	-40.70	-37.01	0.0265	06395	0648
christmas	0027	0469	6.70	-44.70	-41.02	1.0000	47141	3849
direction	0134	0651	3.83	-48.68	-37.72	1.0000	28.862	2378
knowledge	0145	0424	2.97	-49.25	-43.51	1.0000	04834	0467
newspaper	0065	0308	6.57	-45.88	-46.77	1.0000	21.242	0815
vegetable	0010	0112	5.83	-46.57	-42.57	1.0000	06395	0648
amplifier	0006	0015	5.80	-46.59	-39.97	0.5000	45529	2446
fisherman	0005	0077	6.50	-44.61	-34.64	0.4167	40348	1694
gentleman	0028	0121	6.20	-43.83	-39.35	0.5714	16255	1349
infection	0008	0037	4.87	-43.19	-39.15	0.5333	61892	1657
policeman	0019	0155	6.70	-45.90	-36.54	0.6588	38.525	3073
blandness	0001	0000	2.83	-44.96	-39.43	0.0769	45529	2446
performer	0007	0044	5.43	-45.55	-38.83	0.1667	38.525	3073

% OCCURRENCE APPENDIX A4.2 contd

WORD	TERMINAL LETTER f	TERMINAL LETTER v	INITIAL BIGRAM f	INITIAL BIGRAM v	TERMINAL BIGRAM f	TERMINAL BIGRAM v
freedom	016.422	0695	17.96	15.41	00.21	03.48
history	058.623	3647	22.98	10.06	00.61	01.57
machine	200.915	4938	32.44	59.06	11.28	04.61
picture	200.915	4938	04.37	07.19	21.97	16.07
thought	091.728	2432	68.97	17.64	17.87	06.86
bravery	058.623	3647	07.27	15.90	00.61	01.57
goddess	122.649	8866	25.51	10.90	15.59	29.01
product	091.728	2432	30.76	26.33	00.09	00.04
robbery	058.623	3647	11.16	12.67	00.61	01.57
scarlet	091.728	2432	09.82	07.06	10.82	23.14
mastery	058.623	3647	32.44	33.06	00.61	01.57
settler	058.650	2110	13.48	09.88	42.09	14.01
slipper	058.650	2110	01.82	03.56	42.09	14.01
steamer	058.650	2110	17.52	16.67	42.09	14.01
thicket	091.728	2432	68.97	17.64	10.82	23.14
building	039.801	3711	15.21	13.45	04.49	01.94
industry	058.623	3647	54.43	66.12	00.61	01.57
interest	091.728	2432	54.43	66.12	36.04	26.16
position	087.212	3264	15.06	13.11	02.27	00.73
railroad	103.836	4779	12.72	14.03	02.06	06.10
contents	122.649	8866	46.59	42.95	36.04	26.16
contract	091.728	2432	46.59	42.95	00.09	00.04
friction	087.212	3264	17.96	15.41	02.27	00.73
aptitude	200.915	4938	02.56	06.89	31.27	41.76
vocation	087.212	3264	15.76	15.43	02.27	00.73
christmas	122.649	8866	12.19	12.47	00.27	01.68
direction	087.212	3264	23.38	31.58	02.27	00.73
knowledge	200.915	4938	40.57	12.42	00.09	00.12
newspaper	058.650	2110	27.98	26.84	42.09	14.01
vegetable	200.915	4938	24.21	26.85	01.45	02.04
amplifier	058.650	2110	02.41	04.85	42.09	14.01
fisherman	087.212	3264	15.57	14.88	00.49	01.99
gentleman	087.212	3264	16.50	11.56	00.49	01.99
infection	087.212	3264	54.43	66.12	00.49	01.99
policeman	087.212	3264	15.06	13.11	00.49	01.99
blandness	122.649	8866	03.41	09.65	15.59	29.01
performer	058.650	2110	12.72	13.93	42.09	[illegible]

APPENDIX A4.3

WORD	GOOD READERS n = 10 READING	GOOD READERS n = 10 L.D	AVERAGE READERS n = 19 READING	AVERAGE READERS n = 19 L.D	POOR READERS n = 9 READING	POOR READERS n = 9 L.D
body	09	10	16	18	05	08
city	10	10	18	19	07	08
girl	10	10	19	19	09	09
idea	09	08	16	16	02	05
time	10	10	19	19	09	09
army	09	10	14	16	04	08
bowl	10	10	13	16	06	06
fact	10	08	15	16	04	07
form	09	10	14	19	03	08
rock	10	09	19	18	09	09
cane	10	08	10	13	03	04
cord	09	08	15	15	04	07
fork	10	10	18	19	05	09
frog	10	09	17	18	06	08
oven	10	10	12	19	06	09
child	10	10	18	19	07	09
ocean	09	09	14	15	04	05
paper	10	10	19	18	08	09
sugar	10	10	17	17	07	08
table	10	10	19	18	08	08
chair	10	10	17	18	08	09
month	10	10	12	17	04	08
truck	10	10	11	17	01	08
water	10	10	19	18	08	09
woman	09	10	19	17	09	09
hound	10	10	12	15	03	07
peach	10	10	17	17	05	08
answer	10	10	16	16	05	09
friend	10	10	18	19	08	08
method	07	06	06	07	02	03
moment	09	08	15	15	04	08
square	10	10	16	16	05	06
bottle	09	10	18	19	08	08
cattle	10	09	14	16	03	07
flower	10	10	19	18	08	09
insect	09	09	15	17	06	06
ticket	09	10	16	18	04	09
beaver	10	10	12	14	04	04
butter	10	10	18	18	06	08
kettle	09	10	17	18	06	09
master	10	10	17	19	06	08
string	07	10	11	16	07	06

RAW DATA

APPENDIX ~~A4, 3~~ contd

WORD	GOOD READERS n=10		AVERAGE READERS n=19		POOR READERS n=9	
	READING	L.D	READING	L.D	READING	L.D
freedom	09	10	17	17	07	07
history	09	10	16	16	06	07
machine	10	10	13	16	03	08
picture	10	10	17	19	07	09
thought	10	09	16	16	07	08
bravery	07	08	10	15	02	06
goddess	07	06	06	12	01	04
product	04	10	04	13	02	03
robbery	08	09	12	13	02	07
scarlet	07	06	10	07	01	05
mastery	07	07	09	15	04	07
settler	07	07	09	13	02	06
slipper	09	09	16	16	01	07
steamer	08	06	12	12	03	03
thicket	09	02	09	17	01	07
building	10	10	17	17	07	07
industry	04	05	06	08	00	05
interest	06	09	10	16	05	04
position	05	07	06	13	00	05
railroad	10	07	16	11	08	05
contents	07	10	08	12	02	05
contract	08	10	05	13	01	04
friction	08	07	07	06	01	03
aptitude	02	03	05	07	01	04
vocation	06	03	06	05	01	04
christmas	10	10	19	18	08	07
direction	09	09	09	13	01	04
knowledge	06	08	05	10	00	04
newspaper	10	10	17	16	06	07
vegetable	09	09	15	15	05	07
amplifier	03	03	01	01	00	02
fisherman	10	10	16	17	08	09
gentleman	10	10	16	13	04	04
infection	08	07	06	10	03	06
policeman	10	10	18	18	09	08
blandness	05	00	09	09	01	05
performer	08	09	14	11	02	05

RAW DATA USED IN ITEM ANALYSIS – EXPERIMENT 1

RAW DATA

KEY TO THE ABBREVIATIONS OF VARIABLES IN THE TABLES

I	IMAGERY
FOA	FIRST ORDER APPROXIMATION TO ENGLISH
SOA	SECOND ORDER APPROXIMATION TO ENGLISH
ONR	ORTHOGRAPHIC NEIGHBOUR RATIO
ILF	INITIAL LETTER FREQUENCY
ILV	INITIAL LETTER VERSATILITY
TLF	TERMINAL LETTER FREQUENCY
TLV	TERMINAL LETTER VERSATILITY
IBF	INITIAL BIGRAM FREQUENCY
IBV	INITIAL BIGRAM VERSATILITY
TBF	TERMINAL BIGRAM FREQUENCY
TBV	TERMINAL BIGRAM VERSATILITY

PEARSON CORRELATION COEFFICIENTS OF THE INDEPENDENT VARIABLES

	KF	AH	I	FOA	SOA	ONR	ILF	ILV	TLF	TLV	IBF	IBV	TBF	TBV	WL
KF	1.00														
AH	.85**	1.00													
I	-.26	-.07	1.00												
FOA	.31*	.34*	.18	1.00											
SOA	.33*	.36**	.16	.90**	1.00										
ONR	.40**	.34*	-.16	-.13	-.14	1.00									
ILF	.13	.14	.15	.35**	.45**	-.09	1.00								
ILV	-.12	-.10	.18	.07	.09	-.12	.41**	1.00							
TLF	.15	.11	-.04	-.17	-.20	.13	-.16	-.01	1.00						
TLV	-.01	-.04	-.08	-.27	-.26*	.10	-.14	.07	.68**	1.00					
IBF	-.01	-.03	-.24	-.18	-.09	-.01	-.06	-.16	-.01	-.04	1.00				
IBV	-.06	-.10	-.09	-.12	-.04	-.01	.07	-.15	.00	-.00	.72**	1.00			
TBF	-.08	.08	-.01	-.12	-.13	-.18	.13	.18	-.07	-.10	.07	.01	1.00		
TBV	-.17	-.09	-.15	-.16	-.18	-.18	.14	.13	.12	.19	-.01	-.06	.66**	1.00	
WL	-.28*	-.32*	-.24	-.97**	-.89**	.11	-.29*	-.05	.22	.30*	.25	.20	.18	.24	1.00

* SIGNIFICANT AT $p < .01$ ** SIGNIFICANT AT $p < .001$

APPENDIX 4.5

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

GOOD READERS, LEXICAL DECISION

Predictor	Estimate	t(73)	p	partial r
IMAGERY	0.286	3.16	<.01	0.187
FIRST ORDER OF APPROXIMATION	-0.157	2.37	<.05	0.144
ORTHOGRAPHIC NEIGHBOUR RATIO	-2.096	.97	<.01	0.182
INITIAL BIGRAM FREQUENCY	-0.013	1.64	NS	0.097
WORD LENGTH	-0.537	1.82	NS	0.110
WORD FREQUENCY	1.058	6.06	<.001	0.401

Total variance explained i.e. $R^2 = 0.565$

R^2 shows the proportion of the variance accounted for by the significant variables.
A low R^2 indicates a significant but low predictor variable.

APPENDIX A4.6

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

GOOD READERS, READING TASK

Predictor	Estimate	t(76)	p	partial r
IMAGERY	0.411	5.34	$<.001$	0.360
WORD LENGTH	-0.283	3.38	$<.01$	0.233
WORD FREQUENCY	0.283	4.13	$<.001$	0.288

Total Variance explained i.e. R^2 = 0.462

APPENDIX 4.7

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

AVERAGE READERS, LEXICAL DECISION

Predictor	Estimate	t(75)	p	partial r
IMAGERY	0.200	4.18	<.001	0.216
ORTHOGRAPHIC NEIGHBOUR RATIO	-0.669	2.01	<.05	0.105
WORD LENGTH	-0.192	3.28	<.01	0.172
WORD FREQUENCY	0.394	5.12	<.001	0.274

Total Variance explained i.e. $R^2 = 0.485$

APPENDIX A4.8

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

AVERAGE READERS, READING TASK

Predictor	Estimate	t(74)	p	partial r
IMAGERY	0.337	7.67	<.001	0.350
INITIAL BIGRAM FREQUENCY	-0.004	0.70	NS	0.032
INITIAL BIGRAM VERSATILITY	-0.014	2.33	<.05	0.107
WORD LENGTH	-0.097	2.25	<.05	0.103
WORD FREQUENCY	0.346	8.08	<.001	0.383

Total variance explained i.e. R^2 = 0.482

APPENDIX A4.9

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

POOR READERS, LEXICAL DECISION

Predictor	Estimate	t(75)	p	partial r
IMAGERY	0.183	2.92	<.01	0.207
FIRST ORDER APPROXIMATION	-0.078	1.62	NS	0.116
WORD LENGTH	-0.567	2.65	<.02	0.191
WORD FREQUENCY	0.165	2.86	< .01	0.206

Total variance explained i.e. $R^2 = 0.330$

APPENDIX A4.10

MODEL 2: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES

POOR READERS, READING TASK

Predictor	Estimate	t(75)	p	partial r
IMAGERY	0.388	6.13	<.001	0.362
INITIAL BIGRAM FREQUENCY	-0.005	1.00	NS	0.057
WORD LENGTH	-0.007	0.13	NS	0.006
WORD FREQUENCY	0.462	7.3	<.001	0.448

Total Variance explained i.e. R^2 = 0.417

APPENDIX A4.11

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

GOOD READERS, LEXICAL DECISION

Predictor	Estimate	t(74)	p	partial r
FIRST ORDER OF APPROXIMATION	-0.196	2.94	$<.01$	0.178
ORTHOGRAPHIC NEIGHBOUR RATIO	-2.688	4.00	$<.001$	0.249
INITIAL BIGRAM FREQUENCY	-0.015	1.91	NS	0.113
WORD LENGTH	-0.676	2.32	$<.05$	0.140
WORD FREQUENCY	1.194	7.08	$<.001$	0.486

Total variance explained i.e. $R^2 = 0.530$

APPENDIX 4.12

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

AVERAGE READERS, LEXICAL DECISION

Predictor	Estimate	t(76)	p	partial r
ORTHOGRAPHIC NEIGHBOUR RATIO	-0.948	2.94	<.01	0.155
WORD LENGTH	-0.197	3.4	<.01	0.179
WORD FREQUENCY	0.453	6.0	<.001	0.324

Total variance explained i.e. $R^2 = 0.438$

APPENDIX A4.13

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

AVERAGE READERS, READING TASK

Predictor	Estimate	t(75)	p	partial r
INITIAL BIGRAM NEIGHBOUR RATIO	-0.012	2.29	<.05	0.104
INITIAL BIGRAM VERSATILITY	-0.008	1.42	NS	0.065
WORD LENGTH	-0.159	3.90	<.001	0.179
WORD FREQUENCY	0.334	8.20	<.001	0.386

Total variance explained = 0.359

APPENDIX A4.14

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

POOR READERS, LEXICAL DECISION

Predictor	Estimate	t(76)	p	partial r
FIRST ORDER OF APPROXIMATION	-0.103	2.17	$<.05$	0.156
WORD LENGTH	-0.707	3.40	$<.01$	0.246
WORD FREQUENCY	0.165	2.88	$<.01$	0.207

Total variance explained i.e. $R^2 = 0.287$

APPENDIX A4.15

MODEL 3: PARTIAL CORRELATIONS FOR SIGNIFICANT INDEPENDENT VARIABLES (AFTER REMOVING IMAGERY)

POOR READERS, READING TASK

Predictor	Estimate	t(76)	p	partial r
INITIAL BIGRAM FREQUENCY	-0.014	2.66	$<.01$	0.153
WORD LENGTH	-0.069	1.30	NS	0.074
WORD FREQUENCY	0.400	7.00	$<.001$	0.419

Total variance explained i.e. $R^2 = 0.286$

APPENDIX A5.1

CHARACTERISTICS OF WORDS USED IN TASKS 1

EARLY AGE-OF-ACQUISITION

	LETTER LENGTH	AGE-OF -ACQUISITION	IMAGERY	KUCERA & FRANCIS FREQUENCY
hurt	4	2.19	4.72	37
rest	4	2.22	4.41	163
block	5	2.44	4.89	66
hide	4	2.56	4.65	22
minute	6	2.64	4.68	53
cover	5	2.89	4.49	88
fifteen	7	2.89	4.97	56
luck	4	2.92	4.05	47
height	6	2.94	4.78	35
finish	6	3.00	4.43	39
goodness	8	3.14	4.22	16
addition	8	3.14	3.41	142
pattern	7	3.19	4.59	113
learning	8	3.22	3.76	60
blame	5	3.28	3.62	34
breath	6	3.31	4.59	53
opening	7	3.36	4.68	83
object	5	3.39	4.14	65
wonder	6	3.39	4.08	67
search	6	3.44	4.32	66
answer	6	2.94	3.72	152
bother	6	3.22	3.75	22
cousin	6	2.78	4.84	51
distance	8	3.44	4.38	108
favourite	9	3.03	3.84	41
hate	4	2.78	4.86	42
heaven	6	2.72	4.73	43
hold	4	2.67	4.24	169
lesson	6	2.72	4.44	29
primary	7	2.97	3.73	96
prize	5	2.32	4.98	28
shape	5	3.03	4.56	85
sold	4	3.08	3.97	47
spell	5	2.92	4.35	19
stranger	8	3.22	4.60	40
surprise	8	3.22	4.57	51
throw	5	2.64	4.73	42
touch	5	2.69	4.62	87
trouble	7	3.22	4.11	134
wrong	5	2.44	3.57	129
$\bar{x}$	5.9	2.94	4.35	68

APPENDIX A5.2

CHARACTERISTICS OF WORDS USED IN TASKS 1

LATE AGE-OF-ACQUISITION

	LETTER LENGTH	AGE-OF -ACQUISITION	IMAGERY	KUCERA & FRANCIS FREQUENCY
agency	6	5.53	3.72	56
amateur	7	5.22	4.03	25
chlorine	8	5.86	4.86	33
clerk	5	5.02	4.82	34
conflict	8	4.97	4.38	52
debate	6	5.64	4.03	32
denial	6	5.03	3.70	18
device	6	4.72	3.65	55
graph	5	4.70	5.84	17
illusion	8	5.42	4.14	37
income	6	5.06	4.81	109
interior	8	4.81	4.30	74
justice	7	4.00	4.22	114
labour	6	5.06	4.30	149
loyalty	7	4.97	4.03	22
lumber	6	5.19	4.65	35
majority	8	5.31	3.95	57
maturity	8	5.53	4.35	39
review	6	5.47	3.51	56
venture	7	4.92	4.05	19
congress	8	5.75	3.62	152
salary	6	5.58	4.35	43
buffer	6	5.53	4.65	16
freight	7	5.42	3.97	28
grant	5	5.40	4.00	47
liberty	7	5.28	4.08	46
dignity	7	5.22	4.19	35
dispute	7	5.22	4.27	34
professor	9	5.19	5.49	57
rice	4	5.17	4.35	41
career	6	5.14	4.24	67
security	8	5.08	3.97	91
solution	8	5.06	3.97	59
reaction	8	5.06	3.89	124
lover	5	5.05	5.44	19
pioneer	7	4.97	4.22	20
circuit	7	4.94	4.54	23
index	5	4.75	4.28	81
volume	6	4.61	4.73	135
permit	6	4.58	4.16	77
$\bar{x}$	6.65	5.16	4.29	55.7

CHARACTERISTICS OF WORDS USED IN TASK 2: HIGH IMAGERY

	letters	age	KF	I
brother	(7)	2.19	73	4.89
money	(5)	2.47	265	6.04
window	(6)	2.31	119	6.02
hospital	(8)	3.19	110	6.02
woman	(5)	2.58	224	6.26
body	(4)	2.67	276	6.14
island	(6)	2.89	167	6.43
oil	(3)	3.03	93	5.75
building	(8)	3.00	160	5.78
uniform	(7)	3.31	51	5.91
radio	(5)	3.17	120	6.13
market	(6)	3.28	155	5.83
valley	(6)	3.39	73	6.00
grave	(5)	3.39	33	6.19
minister	(8)	3.42	61	5.84
liquid	(6)	3.43	48	5.89
clay	(4)	3.58	100	5.75
cellar	(6)	3.61	26	5.72
wound	(5)	3.72	28	5.70
moonlight	(9)	3.78	13	6.13
wine	(4)	4.03	72	6.24
planet	(6)	4.14	21	5.78
furnace	(7)	4.28	11	5.86
athlete	(7)	4.28	9	5.91
menu	(4)	4.33	5	6.13
highway	(7)	4.36	40	5.81
disc	(4)	4.47	6	5.75
olive	(5)	4.58	5	5.78
president	(9)	4.58	382	5.72
examination	(11)	4.58	29	5.75
dormitory	(9)	4.66	2	5.75
crucifix	(8)	4.75	3	5.76
automobile	(10)	4.83	50	6.28
bosom	(5)	4.89	8	5.93
embrace	(7)	4.92	13	5.97
cologne	(7)	4.94	9	5.86
fountain	(8)	3.89	18	6.02
saxophone	(9)	5.19	4	6.02
inferno	(7)	5.25	2	5.72
physician	(9)	5.61	14	5.72
$\bar{X}$	6.55	3.88	72.5	5.85

CHARACTERISTICS OF WORDS USED IN TASK 2: LOW IMAGERY

	letters	age	KF	I
try	(3)	2.29	140	3.18
wrong	(5)	2.44	129	3.44
guess	(5)	2.92	56	3.30
anybody	(7)	3.06	42	3.29
addition	(8)	3.14	142	3.47
part	(4)	3.14	500	3.40
extra	(5)	3.31	50	3.37
nonsense	(8)	3.36	13	3.49
thought	(7)	3.47	515	3.48
moment	(6)	3.50	246	3.34
usual	(5)	3.53	96	2.35
excuse	(6)	3.56	27	3.10
impossible	(10)	3.58	84	3.02
main	(4)	3.64	119	3.09
spare	(5)	3.64	23	3.16
worth	(5)	3.69	94	2.75
normal	(6)	3.75	136	2.94
position	(8)	3.75	241	3.46
difference	(10)	3.78	148	2.93
reason	(6)	3.83	241	2.85
pause	(5)	4.14	21	3.47
hint	(4)	4.08	9	3.47
ease	(4)	4.28	42	3.27
respect	(7)	4.33	125	3.43
belief	(6)	4.47	64	3.28
proof	(5)	4.56	40	3.39
extreme	(7)	4.58	62	3.32
incident	(8)	4.58	49	3.44
scheme	(6)	4.75	33	3.19
indication	(10)	4.81	20	3.40
inquiry	(7)	4.83	17	3.21
transfer	(8)	4.89	38	3.13
origin	(6)	4.92	44	3.06
judgement	(9)	4.97	1	3.33
pretence	(8)	5.00	4	3.07
outset	(6)	5.09	13	2.70
deceit	(6)	5.19	2	3.38
namesake	(8)	5.19	2	3.48
forfeit	(7)	5.33	3	3.10
alias	(5)	5.61	1	2.94
$\bar{X}$	6.38	4.08	90.8	3.21

APPENDIX A5.5

AGE-OF-ACQUISITION/IMAGERY LIST

Instructions to child

We are going to play a word game on this BBC computer. When you press the bar gently using the hand you write with, you will see a word come up in the middle of the screen. I want you to say it out loud for me, as clearly and as quickly as you can, talking into this microphone. We'll go through a few examples first so as to get the hang of it. Now remember, I'm interested in the time it takes you to say the word - try and be as fast and accurate as you can. If you can't say the word just say PASS.

AGE OF ACQUISITION AND READING -- SUBJECT ANALYSIS

ANOVA TABLE

	SS	df	MS	f	p
READER GROUP	4923.08	2	2461.54	22.1	<.001
AOA	3553.55	1	3553.55	182.56	<.001
READER x AOA	234.47	2	117.24	6.02	<.001
MSE_b	856.45	44	19.47		

IMAGERY AND READING -- SUBJECT ANALYSIS

ANOVA TABLE

	SS	df	MS	f	p
READER GROUP	4999.02	2	2499.51	22.69	$<.001$
IMAGERY	100.50	1	100.50	14.99	$<.001$
READER x IMAGERY	80.54	2	40.27	6.01	$<.005$
MSE_b	295.02	44	6.71		

EFFECTS OF WORD IMAGERY AND AGE OF ACQUISITION ON ADULTS READING

MEAN REACTION TIMES (IN MILLISECONDS)
FOR TWENTY SUBJECTS TASK 1

	Early Age-of-Acquisition	Late Age-of-Acquisition
1	609.6	638.7
2	481.1	454.9
3	534.9	530.3
4	553.0	543.2
5	510.1	537.0
6	541.8	562.1
7	554.7	549.5
8	480.0	467.5
9	490.7	548.1
10	549.1	566.7
11	453.1	478.0
12	586.0	626.4
13	422.8	456.0
14	457.7	496.7
15	499.8	496.1
16	511.8	509.0
17	490.6	521.7
18	473.8	470.8
19	444.0	481.5
20	438.0	453.3
MEAN	504.2	519.4
MEDIAN	495.2	515.4
S.D	50.8	53.5
RANGE	422.8 - 609.6	453.3 - 638.7

EFFECTS OF WORD IMAGERY AND AGE OF ACQUISITION ON ADULT'S READING

MEAN REACTION TIMES (IN MILLISECONDS)
FOR TWENTY SUBJECTS TASK 2

	High Imagery	Low Imagery
1	555.4	558.3
2	557.5	613.9
3	556.9	541.7
4	611.4	567.8
5	550.5	522.9
6	583.4	615.9
7	544.4	513.5
8	510.9	475.3
9	525.7	500.5
10	594.3	614.5
11	476.0	466.7
12	657.3	646.6
13	428.1	451.0
14	436.2	471.9
15	562.2	562.9
16	572.7	609.6
17	454.7	445.9
18	525.3	493.7
19	425.5	437.4
20	460.3	434.7
MEAN	529.4	527.2
MEDIAN	547.5	518.2
S.D.	64.6	68.3
RANGE	425.6 - 657.3	434.7 - 646.6

CHARACTERISTICS OF WORDS & NONWORDS USED IN EXPERIMENT 4

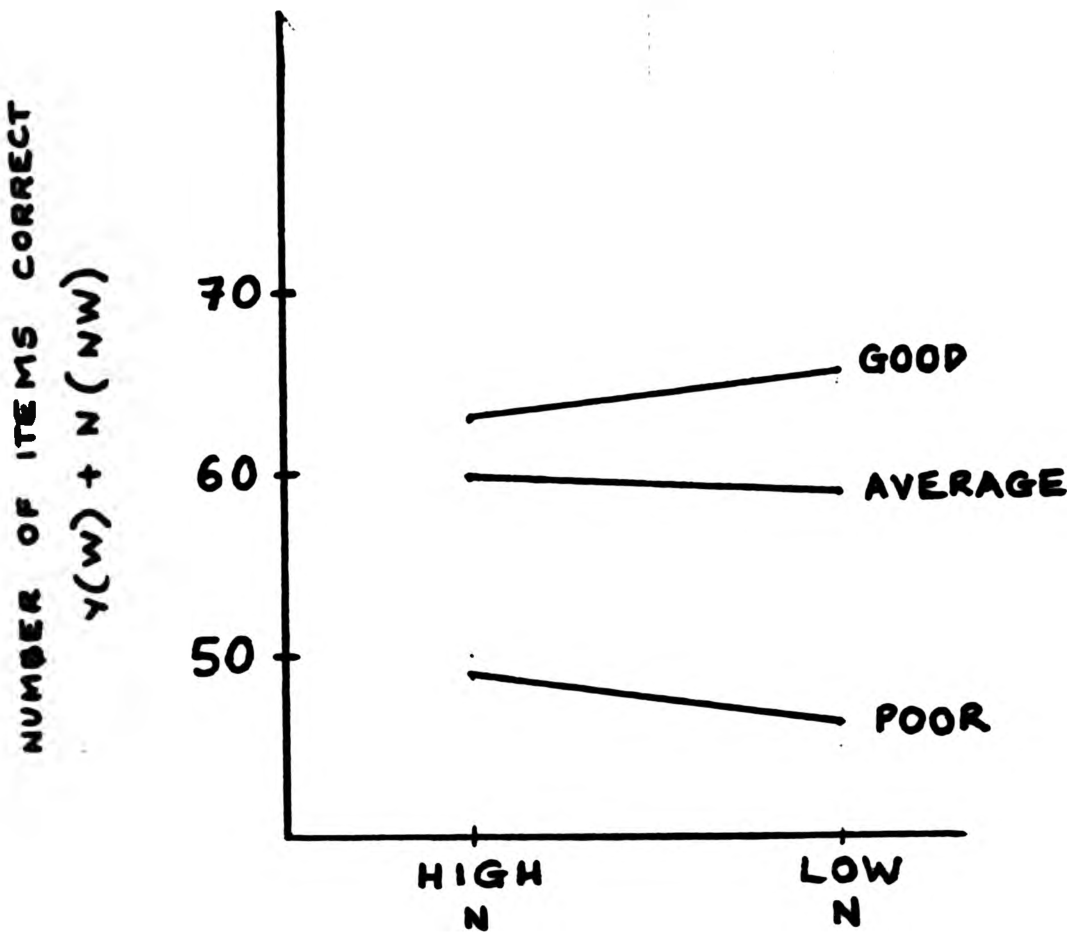
		HIGH-N WORDS	n	Kf	LOW-N WORDS	n	Kf	HIGH-N NONWORDS	n	LOW-N NONWORDS	n
FOUR LETTERS	1	same	17	686	know	4	683	bave	15	baft	3
	2	hand	9	431	once	0	499	brab	10	blid	1
	3	word	12	274	free	3	260	fand	11	buch	4
	4	full	14	230	else	2	176	filt	12	buke	4
	5	bill	18	143	club	3	145	flad	11	crem	2
	6	slow	12	60	neck	4	81	gark	10	dirn	3
	7	mine	22	59	jury	2	67	hage	14	frip	4
	8	seat	17	54	busy	3	58	mide	15	grun	2
	9	rare	18	41	poem	1	48	mook	13	jild	4
	10	barn	13	29	pond	3	25	nuck	11	plun	4
	11	bold	12	21	ugly	0	21	rame	21	saln	3
	12	sung	9	18	diet	4	21	sare	27	shub	4
	13	boot	13	13	gown	4	16	solt	15	slet	2
	14	cave	18	9	axle	1	5	tane	20	spog	4
	15	lash	13	6	itch	2	5	tink	11	swod	3
	16	suck	15	5	ajar	1	2	tord	12	tras	4
	17	hook	11	5	gulp	3	2	wace	13	trin	2
	18	cart	15	5	rasp	4	2	wull	13	vond	4
	19	mole	17	4	gnaw	1	1	yole	13	wege	2
	20	brag	9	2	snob	2	0	zale	12	woln	1
FIVE LETTERS	21	sound	7	204	major	1	247	bater	10	aspet	3
	22	space	6	184	below	0	145	blane	8	blice	2
	23	tried	6	170	month	1	130	brack	8	clune	3
	24	lower	10	123	wrong	3	129	chack	8	crich	1
	25	eight	10	104	piece	1	129	chone	7	darce	2
	26	share	12	98	scene	3	106	crade	9	flort	3
	27	watch	7	81	coast	3	61	cruss	6	glime	3
	28	grown	7	43	motor	1	56	flack	8	lepen	1
	29	grade	7	35	cross	3	55	gight	12	prent	2
	30	sheep	6	23	slept	2	27	glave	7	ronch	2
	31	pitch	8	22	stiff	3	21	grafe	7	rotan	0
	32	clock	8	20	flood	2	19	jated	9	shold	1
	33	trick	7	15	waist	2	11	mough	6	skarn	0
	34	spine	7	6	yacht	0	4	prain	6	sotch	2
	35	bully	9	4	dwarf	0	3	rared	16	stult	2
	36	greed	7	3	farce	1	3	shart	11	sturd	0
	37	fated	11	0	thorn	1	3	slare	14	tharn	1
	38	stank	9	0	clown	3	3	stabe	6	troct	2
	39	miner	8	1	daisy	2	0	stine	8	waral	0
								strat	6	wrolk	0
	$\bar{x}$		11.18	82.8		2.03	83.8		11.28		2.25
	RANGE		6-22	0-686		0-4			6-27		0-4

INSTRUCTIONS FOR READING TASK N LIST

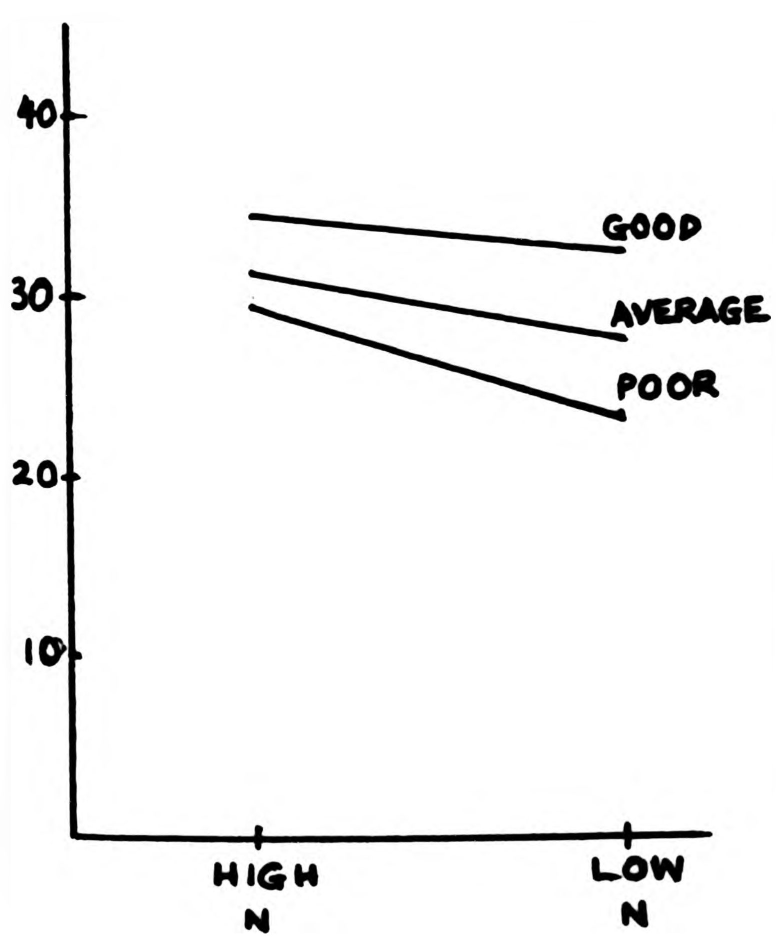
I am going to show you two lists approximately 80 items long each. One of those lists is a word list. The other is a nonword list. When you press the bar gently using the hand you write with, you will see a word/nonword come up in the middle of the screen. I want you to read it out as clearly and quickly as you can, talking into this microphone. We'll go through a few examples first so as to get the hang of it. Now remember, I'm interested in the time it takes you to say it, so don't rush and push yourself too much - try and be as fast and accurate as you can.

PLOT OF MEANS

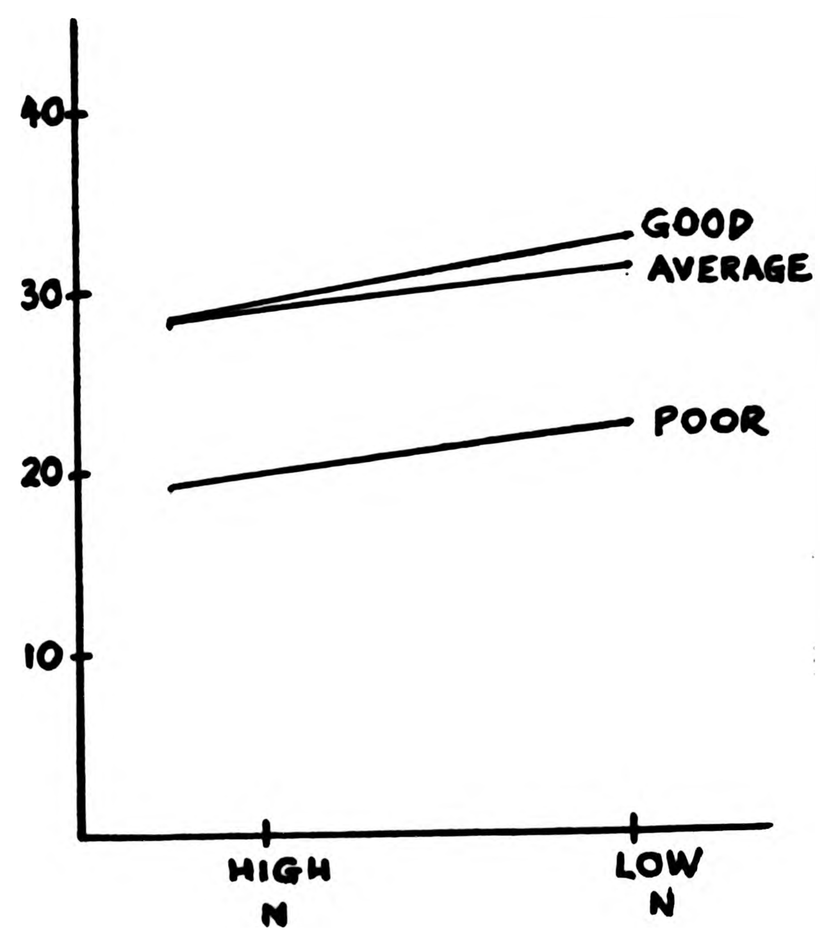
LEXICAL DECISION FOR WORDS & NONWORDS



LEXICAL DECISION FOR WORDS



LEXICAL DECISION FOR NONWORDS



APPENDIX A6.4

READING LATENCIES FOR SKILLED ADULT READERS EXPERIMENT 5

SUBJECT	HIGH-N WORDS	LOW-N WORDS
1	521.7	517.4
2	458.5	458.8
3	627.8	636.1
4	421.5	423.6
5	580.8	567.9
6	632.0	705.3
7	554.3	604.2
8	428.3	438.6
9	513.9	520.1
10	571.3	592.5
11	607.2	600.0
12	583.1	582.5
13	478.2	468.2
14	557.8	560.6
15	474.4	482.7
16	515.4	532.8
17	683.6	735.0
18	515.2	521.9
19	469.9	472.9
20	664.6	642.1
21	536.5	567.7
22	423.9	439.9
23	510.1	527.5
24	524.0	507.4
25	562.6	541.2
26	679.8	679.5
27	469.5	472.9
28	614.5	490.7
29	545.2	539.7
30	584.7	601.3
$\bar{x}$	543.7	547.7
S.D	73.6	79.6

APPENDIX A6.5

MEAN RATINGS OF AGE OF ACQUISITION MEASURES FOR HIGH-N AND LOW-N WORDS

HIGH N		LOW N	
same	2.15	know	2.5
hand	1.85	once	2.225
word	2.275	free	3.125
full	2.025	else	2.825
bill	3.975	club	3.5
slow	1.95	neck	2.35
mine	1.95	jury	4.825
seat	1.975	busy	3.225
rare	4.025	poem	3.2
barn	2.925	pond	2.575
bold	3.55	ugly	2.525
sung	2.675	diet	4.6
boot	2.25	gown	3.35
cave	2.95	axle	5.05
lash	4.15	itch	2.675
suck	2.05	ajar	4.8
hook	2.875	gulp	3.65
cart	2.8	rasp	5.25
mole	3.275	gnaw	4.575
brag	4.6	snob	4.775
sound	2.45	major	3.85
space	3.05	below	2.575
tried	2.775	month	2.675
lower	2.825	wrong	1.8
eight	2.225	piece	2.775
share	2.425	scene	3.975
watch	2.35	coast	3.6
grown	2.625	motor	3.2
grade	4.05	cross	2.425
sheep	2.275	slept	2.55
pitch	3.8	stiff	3.025
clock	2.225	flood	3.175
trick	3.1	waist	3.15
spine	3.8	yacht	3.45
bully	2.85	dwarf	2.8
greed	3.35	farce	6.1
fated	5.175	thorn	3.45
stank	4.075	clown	2.45
miner	4.00	daisy	2.4
$\bar{x}$	2.97	$\bar{x}$	3.36
S.D	0.84	S.D	0.98

APPENDIX A6.6

INSTRUCTIONS ON BOOKLET OR RATING SCALES FOR EXPERIMENT 6

AGE OF ACQUISITION

In connection with an experiment on certain properties of memory for words, we need your estimate of when in your life you probably first learned each of a series of words, i.e. first learned the word and its meaning either in spoken or written form.

This is done by using a 7-point rating scale. Use the rating scale numbers as follows:-

RATING	FOR A WORD LEARNED BETWEEN THE AGES OF:
1	1 and 2 years
2	3 and 4 years
3	5 and 6 years
4	7 and 8 years
5	9 and 10 years
6	11 and 12 years
7	13 years onwards

For example if you think you were between 1 and 2 years old when you learned the word "mum", give that word a rating of 1. If you think you were older than 13 when you learned the word "aardvark" give it rating of 7.

APPENDIX A7.1

STIMULUS MATERIALS FOR EXPERIMENT 7

HIGH FREQUENCY WORDS (CATF1)

NO HOSTILE NEIGHBOURS (HN1) (FN1)				The items in this cell are not included in the analyses because there are no matches. They are both LF & HF				
Consistent Regular R=1		Gang Irregular R=2		Consensus(h) Regular R=1		Heretics Irregular R=2		
				75	raid	37	broad	
				76	laid	38	break	
1	meet	12	high	77	stain	30	whose	
2	soil	13	tied	78	rave	40	foot	
3	dark	14	sight	79	gave	41	floor	
4	green	15	cried	80	dare	42	door	
5	check	16	walk	81	blare	43	both	
				82	care	44	great	
xf	162.8	xf	177.2	83	share	45	again	)
						46	said	) VHF
						47	have	)
						48	are	)

LOW FREQUENCY WORDS (CATF2)

NO HOSTILE NEIGHBOURS (HN1) (FN1)				MOST NEIGHBOURS HOSTILE (HN2) ONE AT LEAST IS HF (FN2)			
Consistent Regular R=1		Gang Irregular R=2		Hero Regular R=1		Heretic Irregular R=2	
				22	hast	27	soot
6	sage	17	nigh	23	mall	28	wand
7	jest	18	pied	24	pant	29	deaf
8	slay	19	bight	25	spook	30	steak
9	wince	20	stalk	26	scant	31	gross
10	sheen	21	balk	xf	1.3	xf	3.7
11	(jade)					32	(caste)
						33	(sew)
xf	1.4	xf	2.6			34	(sweat)
						35	(plaid)
						36	(sloth)

HIGH FREQUENCY WORDS (CATF 1)
MOST NEIGHBOURS FRIENDLY (HN3)

HOSTILE NEIGHBOUR IS HF (FN2)				HOSTILE NEIGHBOUR IS LF (FN3)			
Consensus(h) Regular R=1		Gang(h) Irregular R=2		Consensus(h) Regular R=1		Gang(h) Irregular R=2	
67	nose	94	kind	54	cross	90	plant
68	road	95	ball	55	taste	88	look
69	rain		xf 314.9	56	leaf	89	past
70	poor			57	few	91	salt
71	speak			58	land		
	xf 145.0				xf 228.7		xf 321.0
72	(root)						
73	(wheat)						
74	(cloth)						

LOW FREQUENCY WORDS (CATF 2)
MOST NEIGHBOURS FRIENDLY (HN3)

HOSTILE NEIGHBOUR IS HF (FN2)				HOSTILE NEIGHBOUR IS LF (FN3)			
Consensus(h) Regular R=1		Gang(h) Irregular R=2		Consensus(h) Regular R=1		Gang(h) Irregular R=2	
59	post	92	bind	49	toss	84	nook
60	toad	93	pall	50	baste	85	chant
61	fain		xf 2.2	51	sheaf	86	mast
62	moor			52	hew	87	malt
63	bleak			53	rand		xf 2.8
64	(loot) xf 3.4				xf 2.0		
65	(cheat)						
66	(broth)						

Key

CATF 1 = HIGH FREQUENCY WORDS
CATF 2 = LOW FREQUENCY WORDS
HN1 = NO HOSTILE NEIGHBOURS
HN2 = MOST HOSTILE NEIGHBOURS
HN3 = MOST NEIGHBOURS FRIENDLY
FN1 = NO HOSTILE NEIGHBOURS
FN2 = HOSTILE NEIGHBOUR IS HF
FN3 = HOSTILE NEIGHBOUR IS LF

( ) word not used in item analysis

WORD LIST USED IN EXPERIMENT 8

AMBIGUOUS LFLN	n	f	CONSISTENT 1 LFHN	n	f
moth	5	2	cosh	7	1
zone	10	20	made	12	16
rush	11	29	dish	5	26
tome	12	1	robe	8	5
glow	5	12	flew	8	23
sour	13	7	pair	7	59
moor	8	7	roar	8	9
mere	8	77	cure	13	20
$\bar{x}$	9	19.4	$\bar{x}$	8.5	19.9

CONSENSUS/ HERETIC LFLN	n	f	CONSISTENT 2 LFLN	n	f
load	9	44	coat	9	53
leaf	9	13	loaf	5	4
bare	12	71	wage	11	65
gait	6	2	coil	10	2
hint	9	19	gent	10	2
boss	12	6	mess	10	14
howl	5	3	lawn	7	15
pool	9	21	heel	7	25
$\bar{x}$	8.9	22.4	$\bar{x}$	8.6	22.5

GANG/HERO LFLN	n	f	CONSISTENT 3 LFLN	n	f
hasp	7	3	wisp	5	3
salt	5	40	belt	12	22
bath	8	34	mesh	4	4
mild	10	23	silk	7	32
nook	8	2	loop	10	13
vast	8	69	gust	9	7
gall	13	1	sell	11	47
hind	9	2	fond	10	26
$\bar{x}$	8.5	21.8	$\bar{x}$	8.5	19.3

GANG LFLN	n	f	CONSISTENT 4 LFLN		
calf	5	14	gulf	4	10
folk	4	47	bulk	8	31
sign	4	15	song	7	51
walk	5	84	sulk	6	1
palm	6	7	rank	11	40
colt	13	6	tilt	10	5
bold	12	21	weld	10	4
bask	12	1	desk	3	45
$\bar{x}$	7.6	24.4	$\bar{x}$	7.4	23.4

WORD LIST USED IN EXPERIMENT 8 (contd)

AMBIGUOUS LFHN	n	f	CONSISTENT 5 LFHN	n	f
cove	17	12	rope	16	12
cone	19	2	wine	20	70
dear	17	67	deer	14	7
lass	13	3	sack	13	9
dull	15	40	sill	17	4
bead	13	1	seam	12	4
dive	13	2	mile	20	34
mood	10	29	boom	11	10
$\bar{x}$	14.6	19.5	$\bar{x}$	15.4	18.8

CONSENSUS/ HERETIC LFHN	n	f	CONSISTENT 6 LFHN	n	f
cave	19	12	bake	19	3
coot	18	1	boon	14	3
dose	14	18	mole	19	11
beat	22	59	meal	13	51
fare	18	13	wire	16	33
beak	12	3	seed	18	41
rain	13	40	hail	14	1
lash	14	6	bark	15	1
$\bar{x}$	16.3	19	$\bar{x}$	16	18

AMBIGUOUS HFLN	n	f	CONSISTENT 7 HFLN	n	f
near	13	207	hair	6	131
food	11	223	feet	9	222
five	9	261	side	10	409
home	10	546	hope	12	207
cost	11	183	west	11	314
town	7	262	keep	8	265
hour	8	133	fire	11	125
poor	7	132	sure	7	208
$\bar{x}$	9.5	243.4	$\bar{x}$	9.3	235.1

CONSENSUS/ HERETIC HFLN	n	f	CONSISTENT 8 HFLN	n	f
goes	10	100	says	11	195
main	12	245	seem	11	246
paid	8	159	fair	8	116
road	10	205	real	12	214
note	9	111	face	12	408
lord	7	269	dark	11	145
mean	10	230	girl	3	307
form	11	394	soon	8	263
$\bar{x}$	9.6	214.1	$\bar{x}$	9.5	236.8

APPENDIX A7.3

RATIONALE FOR ITEMS INCLUDED IN THE n COUNT

1. Foreign words were excluded unless listed in Hofland & Johansson

    cafe, pate - included

    roue, moue, doge - excluded

2. Names excluded unless listed in Hofland & Johansson

    Dave, Mick etc in — Sean - out

    York, Rome, Nile etc - in — Copt - out

    but nicknames excluded

    Gert where Gertrude is listed

    Bess where Bessie is listed etc.

3. Only words listed in the Concise Oxford Dictionary were included

    except

    archaic words with no listing in Hofland & Johansson eg kine, wives

    dialect words with no listing

Thus certain slang words were included eg pooh, gosh, with no listing in Hofland & Johnsson, or with a derived listing - jelly - jell

Others were not - nosh, bonk

DERIVATION OF FREQUENCIES FOR NEIGHBOURS

Where these were not listed in Hofland & Johansson they were given a 0 unless a closely derived word was listed (-ing, -s, -ed, -ers and in one case -ery for rook); in this case the frequency of the derived word was used (usually this was very low frequency 1, or 2). Where there were two derived words, the highest frequency was used.

Where words were listed with capitals (Cape) only, this frequency was used.

For some words (eg. seas, ways) there were both listings for the word and its derivation. In this case the higher frequency was used.

SELECTION OF ITEMS: TARGET WORDS

1. **Frequency**

   Low frequency 100 range 6-84

   One item achieved its frequency through a derivation (bask)

   High frequency 100 range 100-546

2. **N**

   Low N range 3-14

   High N range 10-20

   The overlap was unfortunate but limited to one or two words in each set.

Every attempt was made to choose items yielding similar means for both frequency and N within the relevant comparisons.

3. **Endings used in target words**

An attempt was made to include one exemplar of each type of ending in the wordsets, excluding very small sized bodies, with LF members (eg. bomb, tomb, comb) or with words involving an apostrophe (eg. don't, can't)

Within the Consensus lists and Ambiguous lists it was necessary to include several endings of the same type e.g. poor, moor but never more than twice; and no two pairs of lists shared more than three endings.

4. **Definition of heretics and heroes**

Items were excluded if

- the heretic or hero had that status through foreign pronounciation

  e.g. -afe words for cafe

  -ate words for pate

  -ote was included because of (dove)-cote
- they did not have a listing in Hofland & Johansson thus wast and dost were not considered, whereas hast shalt and hath were deemed heroes

Items were included if -ook words where there were three "heros", none listed spook, stook, snoot

Where words existed with heroic pronounciations achieved in two syllable words (eg. gallon) these were not considered heros; i.e. only single syllable words were considered, letter length 3-6

-aft was not included in Gang/hero because the exception word waft was governed by the w rule

-asp was included because of asp despite the pronunciation (wasp), governed by the -w rule.

SELECTION OF CONSISTENT MATCHED ITEMS

Every attempt was made to match each target word with another of identical structure, so as to make the consistent body types as comparable as possible.

vowel + cons + e matched in most cases with the same vowel (eg zone-mode)

vowel digraphs were matched in most cases with others, oo with ee if an exact match was not possible or ea with oa

vowel dipthongs matched with others (ai with oi)

vowels + L + cons and +r + cons were matched with others (eg calf with gulf). Where this was not possible r was substituted for l or another consonant used, fork matched with curl where vowel digraphs + r resulted in tripthongs attempts were made to match this eg. moor paired with pair, sour with roar; in one case this wasn't possible and poor was matched with sure

words ending in double consonants with one silent were matched e.g. mess with boss, where this was not possible another double consonant was selected i.e. lass matched with sack

words ending in consonant digraphs were matched e.g. gush with dish, where this wasn't possible another consonant digraph was used e.g. bath with mesh

words ending in double consonants where both were sounded were matched gasp with wisp, gust with vast, desk with bark. In a few cases the match was not so close

sigh matched with song

palm matched with rank

These were selected within the range of both N values and frequency values, and every attempt was made to achieve very close means.

APPENDIX A7.4

EXPERIMENT 8

ANOVA TABLE FOR ANALYSIS 1: LOW FREQUENCY LOW N CONSISTENT BODIES

SOURCE OF VARIATION	ss	df	MS	F	P
AGE GROUP	583.36	2	291.68	21.59 (2,57)	<0.001
ERROR	770.13	57	13.51		
CONSIS BODY	2.58	3	0.86	0.81 (3,171)	NS
GROUP x CONSIS BODY	11.64	6	1.94	1.82 (6,171)	NS
ERROR	182.28	171	1.07		
TOTAL	1549.98	239			

APPENDIX A7.5

EXPERIMENT 8

ANOVA TABLE FOR ANALYSIS 2: THE N EFFECT IN LOW FREQUENCY WORDS

SOURCE OF VARIATION	ss	df	MS	F	P
AGE GROUP	893.51	2	446.76	31.99 (2,57)	< 0.001
ERROR	795.98	57	13.97		
N	3.6	1	3.6	5.3 (1,57)	< 0.024
GROUP X N	0.83	2	0.42	0.61 (2,57)	NS
ERROR	38.74	57	0.68		
BODY	28.04	2	14.02	18.37 (2,114)	< 0.001
GROUP X BODY	1.48	4	0.37	0.48 (4,114)	NS
ERROR	86.98	114	0.76		
N X BODY	23.45	2	11.73	17.56 (2,114)	< 0.001
AGE GROUP X N BODY	4.28	4	1.07	1.60 (4,114)	NS
ERROR	76.1	114	0.67		
TOTAL	1952.99	359			

APPENDIX A7.6

EXPERIMENT 8

ANOVA TABLE FOR ANALYSIS 3: THE FREQUENCY EFFECT IN LOW N WORDS

SOURCE OF VARIATION	SS	df	MS	F	P
AGE GROUP	754.62	2	377.31	39.77	<0.001 (2,57)
ERROR	540.73	57	9.49		
F	234.42	1	234.42	150.05	<0.001 (1,57)
GROUP X F	14.83	2	7.41	4.75	<0.012 (2,57)
ERROR	89.05	57	1.56		
BODY	7.42	2	3.71	4.72	<0.011 (2,114)
GROUP X BODY	12.14	4	3.04	3.87	<0.006 (4,114)
ERROR	89.52	114	0.79		
F X BODY	28.90	2	14.45	21.93	<0.001 (2,114)
AGE GROUP X F X BODY	2.06	4	0.52	0.78	NS
ERROR	75.12	114	0.66		
TOTAL	1848.80	359			

APPENDIX A7.7

EXPERIMENT 8

ANOVA TABLE FOR ANALYSIS 4: ANALYSIS OF LOW FREQUENCY, LOW N BODIES

SOURCE OF VARIATION	SS	df	MS	F	P
AGE GROUP	724.86	2	362.43	31.03 (2,57)	<0.001
ERROR	665.72	57	11.68		
BODY	78.49	4	19.62	20.05 (4,228)	<0.001
AGE GROUP X BODY	14.15	8	1.77	1.81 (8,228)	NS
ERROR	223.17	228	0.98		
TOTAL	1706.37	299			

BIBLIOGRAPHY

Andrews, S. (1982) Phonological recoding : Is the regularity effect consistent? Memory and Cognition, 10 (6), 565 - 575.

Backman, J., Bruck, M., Hebert, M. and Seidenberg, M.S. (1984) Acquisition and use of spelling - sound correspondances in Reading. Journal of Experimental Child Psychology, 38, 114 - 133.

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

Baker, R.J. and Nelder, J.A. (1978) The GLIM system Release 3 Generalized linear interactive modelling. Oxford : Numerical Algorithms Group.

Balota, D.A. and Chumbley, J.I. (1984) Are Lexical Decisions a Good Measure of Lexical Access? The Role of Word Frequency in the Neglected Decision Stage. Journal of Experimental Psychology : Human Perception and Performance, 10, 340 - 357.

Balota, D.A. and Chumbley, J.I. (1985) The Locus of word - Frequency Effects in the Pronunciation Task : Lexical Access and/or Production? Journal of Memory and Language, 24, 89 - 106.

Baron, J. and Strawson, C. (1976) Use of orthographic and word - specific knowledge in reading words aloud. Journal of Experimental Psychology : Human Perception and Performance, 4, 207 - 214.

Baron, J. (1979) Orthographic and word - specific Mechanisms in childrens Reading of words Child Development, 50, 60 - 72.

Barron, R.W. (1981) Reading skill and Reading strategies. In Lesgold, A. and Perfetti, C.A. (Eds.), Interactive Processes in Reading. Lawrence Erlbaum Associates : Hillsdale, New Jersey.

Barron, R.W. (1981) Development of Visual Word Recognition: A Review. In MacKinnon, G.E. and Waller T.G. (Eds.) Reading Research: Advances in Theory and Practice. Volume 3.

Barron, R.W. and Pittinger, J.G. (1974) The effect of orthographic structure and lexical meaning on same-different judgements. Quarterly Journal of Experimental Psychology, 26, 566-581.

Bauer, D.W. and Stanovich, K.E. (1980) Lexical access and the spelling - to - sound regularity effect. Memory and Cognition, 8 (5), 424 - 432.

Bradley, L. and Bryant, P.E. (1983) Categorizing sounds and learning to read - a casual connection. Nature, 301, 419 - 421.

Brown, D.A., and Watson, F.L. First In, First Out: Word Learning Age and Spoken Word Frequency as Predictors of Word Familiarity and Word Naming Latency. Memory and Cognition (in press).

Bryant, P. and Bradley, L. (1985) Childrens Reading Problems - Psychology and Education. Blackwell Press.

Carr, T.H. and Pollatsek, A. (1985) "Recognising printed words: a look at current models" In D. Besner, T.G. Waller and G.E. MacKinnon (Eds.) Reading Research: Advances in Theory and Practice. Volume V.

Carroll, J.B., Davis, P. and Richmond, B. (1971) The American Heritage Word Frequency Book. Boston : Houghton, Mifflin.

Carroll, J.B. and White, M.N. (1973) Age - of - Acquisition Norms for 220 Picturable Nouns. Journal of Verbal Learning and Verbal Behaviour, 12, 563 - 576.

Carroll, J.B. and White, M.N. (1973) Word Frequency and Age of Acquisition as determiners of picture - naming latency. Quarterly Journal of Experimental Psychology, 25, 85 - 95.

Carr, T.H. (1981) Building theories of reading ability : on the relation between individual differences in cognitive skills and reading comprehension. Cognition, 9, 73 - 114.

Chumbley, J.I. and Balota, D.A. (1984) A word's meaning affects the decision in lexical decision. Memory and Cognition, 12 (6), 590 - 606.

Chumbley, J.I. and Balota, D.A. (1984) A word's meaning affects the decision in lexical decision. Memory and Cognition, 12 (6), 590 - 606.

Coltheart,M, Davelaar, E., Jonasson, J.T. and Besner, D, (1977) Access to the Internal Lexicon. In S.Dornic (Ed.), Attention and Performance. Volume VI. New York : Academic Press.

Coltheart, V., Laxon, V.J., Keating, G.C. and Pool, M.M. (1986) Direct Access and Phonological Encoding Processes in Children's reading: Effects of Word Characteristics. British Journal of Educational Psychology, 56, 255-268.

Coltheart,V. and Winograd, E. (1986) Word imagery but not age of acquisition affects episodic memory. Memory and Cognition, 14 (2), 174 - 180.

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

Forster, K. I.and Chambers, S.M. (1973) Lexical Access and naming time. Journalof Verbal Learning and Verbal Behaviour, 18, 381 - 398.

Francis, H. (1984) Children's knowledge of orthography in learning to read. British Journal of Educational Psychology, 54, 8 - 23.

Friendly, M., Franklin,. P.E., Hoffman, D. and Rubin, D.C. (1982 The Toronto Word Pool : Norms for imagery, concreteness and orthographic variables and grammatical usage of 1080 words. Behaviour Research Methods and Instrumentation, 14, 375 - 399.

Frith, U. (1985) Beneath the surface of developmental dyslexia. In K.E. Patterson,J.C. Marshall and M. Coltheart (Eds.), Surface Dyslexia : Cognitive Neuropsychological studies of phonological reading. London : Lawrence Erlbaum Associates.

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

Gibson, E.J., Tenney, Y.J., Barron, R.W. and Zaslow, M. (1972) The effect of orthographic structure on letter search. Perception and Psychophysics, 11, 183 - 186.

Gilhooly, K.J. and Gilhooly, M.L. (1979) Age - of - acquisition effects in lexical and episodic memory tasks. Memory and Cognition, 7 (3) 214 - 223.

Gilhooly, K.J. and Hay, D. (1977) Imagery, concreteness, age - of - acquisition, familiarity, and meaningfulness values for 205 five - letter words having single - solution anagrams. Behaviour Methods and Instrumentation, 9 (1), 12 - 17.

Gilhooly, K.J. and Logie, R.H. (1980) Age - of - acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. Behaviour Research Methods and Instrumentation, 12 (4), 395 - 427.

Gilhooly, K.J, and Logie, R.H. (1982) Word of Age - of - acquisition andlexical decision making. Acta Psychologia, 50 (1), 21 - 34.

Gilhooly, K.J. and Watson, F.L. (1981) Word Age-of-acquisition Effects: a Review. Current Psychological Reviews, 1, 269-286.

Glushko, R.J. (1979) The Organisation and Activation of Orthographic Knowledge inreading Aloud. Journal of Experimental psychology. Human Perception and Performance, 5 (4), 674 - 691.

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

Henderson, L. (1985) Issues in the modelling of pronunciation assembly in normal reading in K.E. Patterson, J.C. Marshall and M.Coltheart (Eds.) Surface Dyslexia - Neuropsychological and Cognitive Studies of Phonological Reading. London : Lawrence Erlbaum.

Henderson, L. and Chard, D. (1980) Knowledge of Orthographic structure. In V. Frith (Ed.), Cognitive Processes in Spelling. London : Academic Press.

Hofland, K. and Johansson, S. (1982) Word Frequencies in British and American English. Bergen : The Norwegian Computing Centre for theHumanities.

Hume, C. (1981) The effects of manual tracing on memory in normal and retarded readers: some implications for multi-sensory teaching. Psychological Research, 43, 179-191. Developmental Dyslexia Issue.

Jones, G.V. (1985) Deep Dyslexia, Imageability, and Ease of Predication. Brain and Language, 24, 1 - 19.

Jorm, A.F. (1977) Effect of Word Imagery on Reading Performance as a function of Reader Nobility. Journal of Educational Psychology, 69, 46 - 64.

Juola, J.F., Schadler, M.,Chabot, R.J. and McCaughey, M.W. (1978) The development of visual information processing skills related to reading. Journal of Experimental Child Psychology, 25, 459-476.

Kamil, M.L. (1984) Current Traditions of Reading Research. In Handbook of Reading Research D.P. Pearson (Ed.) Longmans.

Kay, J. (1985) Mechanisms of Oral Reading : A Critical Appraisal of Cognitive Models. In Ellis, A. (Ed.), Progress in the Psychology of Language. Volume II. London : Lawrence Erlbaum.

Kay, J. and Bishop, D. (1986) Anatomical Differences between Nose, Palm and Foot, or The Body in Question. To appear in Attention and Performance XII : The Psychology of Reading. London : Lawrence Erlbaum.

Kay, J. and Lesser, R. (1985) The Nature of Phonological Processing in Oral Reading : Evidence from surface Dyslexia. The Quarterly Journal of Experimental Psychology, 37A, 39 - 81.

Kay, J. and Marcel, A. (1981) One process, not two in reading aloud : lexical analogies do the work of non - lexical rules. Quarterly Journal of Experimental Psychology, 33A, 397 - 413.

Klose, A.E., Schwartz, S. and Brown, J.W.M. (1983) The imageability effect in good and poor readers. Bulletin of the Psychonomic Society, 21 (6), 446 - 448.

Kroll, J.F. and Merves, J.S. (1986) Lexical Access for Concrete and Abstract Words. Journal of Experimental Psychology: Learning, Memory and Cognition, 12 (1), 92 - 107.

Kucera, H. and Francis, W.N. (1967) Computational analysis of present - day American English. Providence, R.I : Brown University Press.

Lachman, R., Shaffer, J.P, and Hennrikus, D. (1974) Language and Cognition : Effects of Stimulus Codability, Name - Word Frequency, and Age of Acquisition on Lexical Reaction Time. Journal of Verbal Learning and Verbal Behaviour, 13, 613 - 625.

Lefton, L.A., Spragins, A.B., and Byrnes, J. (1973) English orthography: relation to reading experience. Bulletin of the Psychonomic Society, 2, (5a), 281-282.

Marsh, G., Friedman, M., Welch, V. and Desberg, P. (1981) A cognitive - developmental theory of reading acquisition. In G.E. Mackinnon and T.G. Waller (Eds.) Reading Research: Advances in Theory and Practice. New York : Academic Press.

Mason, M. (1975) Reading ability and letter search time : effects of orthographic structure defined by single - letter positional frequency. Journal of Experimental Psychology, 104, 146 - 166.

Mason, M. (1978a) From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. Memory and Cognition, 6, 568 - 581.

Mason, M. (1978b) The role of spatial redundancy in grapheme recognition : perception or inference? Journal of Experimental Psychology : Human Perception and Performance, 4, 662 - 673.

Kroll, J.F. and Merves, J.S. (1986) Lexical Access for Concrete and Abstract Words. Journal of Experimental Psychology: Learning, Memory and Cognition, 12 (1), 92 - 107.

Kucera, H. and Francis, W.N. (1967) Computational analysis of present - day American English. Providence, R.I : Brown University Press.

Lachman, R., Shaffer, J.P, and Hennrikus, D. (1974) Language and Cognition : Effects of Stimulus Codability, Name - Word Frequency, and Age of Acquisition on Lexical Reaction Time. Journal of Verbal Learning and Verbal Behaviour, 13, 613 - 625.

Lefton, L.A., Spragins, A.B., and Byrnes, J. (1973) English orthography: relation to reading experience. Bulletin of the Psychonomic Society, 2, (5a), 281-282.

Marsh, G., Friedman, M., Welch, V. and Desberg, P. (1981) A cognitive - developmental theory of reading acquisition. In G.E. Mackinnon and T.G. Waller (Eds.) Reading Research: Advances in Theory and Practice. New York : Academic Press.

Mason, M. (1975) Reading ability and letter search time : effects of orthographic structure defined by single - letter positional frequency. Journal of Experimental Psychology, 104, 146 - 166.

Mason, M. (1978a) From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. Memory and Cognition, 6, 568 - 581.

Mason, M. (1978b) The role of spatial redundancy in grapheme recognition : perception or inference? Journal of Experimental Psychology : Human Perception and Performance, 4, 662 - 673.

Massaro, D.W., Venezky, R.L. and Taylor, G.A. (1979) Orthographic Regularity, Positional Frequency, and Visual Processing of Letter Strings. Journal of Experimental Psychology : General, 108 (1), 107 - 124.

McCaughey, M.W., Juola, J.F., Schadler, M. and Ward, N.J. (1980) Whole-word units are used before orthographic knowledge in perceptual development. Journal of Experimental Child Psychology, 30, 411-421.

Morton, J. and Patterson, K. (1980) A new attempt at an interpretation, or an attempt at a new interpretation. In M. Coltheart, K. Patterson and J.C. Marshall (Eds.) Deep dyslexia. London : Routledge and Kegan Paul.

Paivio, A., Yuille, J.C. and Madigan, S.A. (1986) Concreteness, Imagery and Meaningfulness Values for 925 nouns. Journal of Experimental Psychology, Monograph Supplement, 76 (1), 2.

Parkin, A.J. (1982) Phonological recoding in lexical decision : Effects of spelling - to - sound regularity depend on how regularity is defined. Memory and Cognition, 10 (1) 43 - 53.

Parkin, A.J. (1983) Two processes not one in reading aloud. Paper presented to the Experimental Psychology Society. Oxford 1093.

Patterson, K. and Coltheart, V. (1986) Phonological processes in Reading : A Tutorial Review : To appear in Attention and Performance XII : The Psychology of Reading. London : Lawrence Erlbaum.

Patterson, K. and Morton, J. (1985) From Orthography to Phonology: An Attempt at an Old Interpretation. In Patterson, K.E., Marshall, J.C. and Coltheart, M. (Eds.) Surface Dyslexia: Cognitive Neuropsychological Studies of Phonological Reading. London : Lawrence Erlbaum Associates.

Rubin, D.C. (1981) First - order approximation to English, second - order approximation to English and orthographic neighbour ration norms for 925 nouns. Behavioural Research Methods and Instrumentation, 13 (6), 713 - 721.

Rumelhart, D.E. and McClelland, J.L. (1982) An Interactive Activation Model of Context Effects in Letter Perception : Part 2. The Contextual Enhancement Effect and some Tests and Extensions of the Model. Psychological Review, 89 (1), 60 - 94.

Santee, J.L. and Egeth, H.E. (1982) Do Reaction Time and Accuracy Measure the same Aspects of Letter Recognition? Journal of Experimental Psychology : Human Perception and Performance, 8, 489 - 501.

Seidenberg, M.S. (1985) Literal and Virtual Context in Word Recognition. In Besner, D., Waller, T.G. and MacKinnon G.E. (Eds.), Reading Research : Advances in Theory and Practice. Volume V. London : Academic Press.

Seidenberg, M.S., Waters, G.S., Barnes, M.A. and Tanenhaus M.K., Spelling-Sound Correspondences and Reading Paper presented to the Psychonomic Society, Minneapolis, November (1982).

Seidenberg, M.S., Waters, G.S., Barnes, M.A. and Tanenhaus. M.K. (1984a) When does irregular spelling or pronunciation influence word recognition? Journal of Verbal Learning and Verbal Behaviour, 23, 383 - 404.

Seidenberg, M.S., Waters, G.S., Sanders, M. and Langer, P. (1984b) Pre - and Post Lexical Loci of Contextual Effects on Word Recognition. Memory and Cognition, 12, 315 - 328.

Seymour, P.H.K., and Elder, L. (1986) Beginning Reading Without Phonology. Cognitive Neuropsychology, 3 (1), 1 - 36.

Seymour, P.H.K., and MacGregor, C.J. (1984) Developmental Dyslexia: A Cognitive Experimental Analysis of Phonological, Morphemic, and Visual Impairments Cognitive Neuropsychology, 1(1), 43-82

Shallice, T. and McCarthy, R. (1985) Phonological Reading : From Patterns of Impairment to Possible Procedures. In K. Patterson, J.C. Marshall and M. Coltheart (Eds.) Surface Dyslexia - Neuropsychological and Cognitive Studies of Phonological Reading. London : Lawrence Erlbaum.

Shoben, E.J. (1982) Semantic and Lexical Decisions, in C.R. Puff (Ed), Handbook of research methods in human memory and cognition - New York; London : Academic Press.

Solso, R.J., Juel, C. and Rubin, D.C. (1982) The Frequency and Versatility of Initial and Terminal Letters in English Words. Journal of Verbal Learning and Verbal Behaviour, 21, 220 - 235.

Spoehr, K. and Smith, E. (1975) The role of orthographic and phonotactic rules in perceiving letter patterns. Journal of Experimental Psychology : Human Perception and Performance, 1, 21 - 34.

Stanovich, K. and Bauer, D. (1978) Experiments on the spelling to sound regularity effect in word recognition. Memory and Cognition, 6, 410 - 415.

Stanovich, K.E., West, R.F., and Pollack, D. (1978) The effect of orthographic structure on word recognition in visual search. Journal of Experimental Child Psychology, 26, 137-146.

Thomson, M.T. (1978) A psycholinguistic analysis of reading errors made by dyslexics and normal readers. Journal of Research in Reading, 1 (1), 7 - 21.

Underwood, G. and Bargh, K. (1982) Word shape, orthographic regularity and contextual interactions in a reading task. Cognition, 12, 197-209.

Vellutino F.R. and Scanlon, D.M. (1985) Free Recall of Concrete and Abstract Words in Poor and Normal Readers. Journal of Experimental Child Psychology, 39, 363 - 380.

Venezky, R.L. (1970) The Structure of English Orthography. The Hague : Mouton.

Walker, J. (1984) Walker's Rhyming Dictionary of the English Language. London : Routledge and Kegan Paul.

Waters, G.S., Bruck, M. and Seidenberg, M.(1985) Do Children Use Similar Processes to Read and Spell Words. Journal of Experimental Child Psychology, 39, 511 - 530.

West, R.F. and Stanovich, K.E. (1982) Source of inhibition in experiments on the effects of sentence context on word recognition. Journal of Experimental Psychology : Learning, Memory, and Cognition, 8, 385 - 399.

Whaley, C.P. (1978) Word - nonword classification time. Journal of Verbal Learning and Verbal Behaviour, 17, 143 - 154.

Wijk, A. (1966) Rules of Pronunciation for the English Language. Oxford : Oxford University Press.

Young, D. (1976) Manual for the SPAR spelling and Reading Tests. London : Hodder and Stoughton.

Zechmeister, E.B. (1969) Orthographic distinctiveness. Journal of Verbal Behaviour, 8, 754 - 761.

ADDENDUM TO BIBLIOGRAPHY

Aitchison, J. and Straf, M. 1982. Lexical Storage & Retrieval: A Developing Skill? In A. Cutler (Ed.) Slips of the Tongue & Language Production. Amersterdam: Mouton.

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

Baker, R.G., & Smith, P.T. (1976) A psycholinguistic study of English stress assignment rules. Language and Speech, 19, 9-27.

Beauvois, M.F., & Derouesne, J. (1979). Phonological alexia: three dissociations. Journal of Neurology, Neurosurgery and Psychiatry, 42, 1115-1124.

Brooks, L. (1974). Visual pattern in fluent word identification. Paper presented at the Conference on Reading. Brooklyn, N.Y., described in Venezky & Massaro paper.

Campbell, R. (1983). Writing nonwords to dictation. Brain and Language, 8, 133-61.

Carr, T.H., and Pollatsek, A. (1985). Recognising printed words: A look at current models. In D. Besner, T.G. Waller & G.E. MacKinnon (Eds.), Reading Research: Advances in Theory and Practice. Vol.5, Orlando, Academic Press.

Chall, J. (1967) Learning to Read: The great debate. New York: McGraw-Hill.

Chomsky, N. & Halle, M. (1968) The Sound Pattern of English. New York: Harper and Row.

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

Ehri, L.C. (1980) The development of orthographic images. In U. Frith (Ed.) Cognitive Processes in Spelling. London: Academic Press.

Ehri, L.C. (1981). The role of orthographic images in learning printed words. In R.L. Venezky & J.F. Kavanagh (Eds.), Orthography, reading and dyslexia. Baltimore, Maryland: University Park Press.

Fay, D. and Cutler, A. (1977) Malapropisms and the structure of the mental lexicon. Linguistic Inquiry, 8, 505-520.

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

Gibson, E.J., Shurcliff, A., & Yonas, A. (1970) Utilization of spelling patterns by deaf and hearing subjects. In H. Levin & J.P. Williams (Eds.) Basic studies on reading. New York: Basic Books.

Goodman, K.S. (1969) Analysis of oral reading miscues: applied psycholinguistics. Reading Research Quarterly, 5, 9-30.

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

Gough, P.B., & Cosky, M. (1975) One second of reading again. In J. Castellan, D. Pisoni, & G. Potts (Eds.), Cognitive Theory (Vol.II). Hillsdale, N.J., Erlbaum.

Hirata, K., & Bryden, M. (1971) Tables of letter sequences varying in order of approximation to English. Psychonomic Science, 25, 322-324.

Humphreys, G.W., & Evett, L.J. (1985) Are there independent lexical and non-lexical routes in word processing? An evaluation of the dual route theory of reading The Behavioural and Brain Sciences, 8, 689-740.

James, C.T. (1975). The role of semantic information in lexical decisions. Journal of Experimental Psychology: Human Perception and Performance, 1, 130-136.

Kamil, M.L. (1984) in D. Pearson (Ed.) Handbook of Reading Reseasrch. Longmans.

Kerlinger, F.M. (1977) The influence of research on educational practice. Educational Researcher, 6, 5-12.

Kieras, D. (1978). Beyond pictures and words: Alternative information - processing models for imagery effects in verbal memory. Psychological Bulletin, 85, 532-554.

Lefton, L.A., & Spragins, A.B. (1974). Orthographic structure and reading experience affect the transfer from iconic to short-term memory. Journal of Experimental Psychology, 103, 775-781.

Leu, D.J. (1982). Oral reading error analysis: A critical review of research and application. Reading Research Quarterly, 17, 420-445.

Marcel, T. (1980). Surface dyslexia and beginning reading: A revised hypothesis of the pronounciation of print and its impairments. In M. Coltheart, K. Patterson & J.C. Marshall (Eds.), Deep Dyslexia. London: Academic Press.

Marshall, J.C., and Newcombe, F. (1973). Patterns of paralexia. Journal of Psycholinguistic Research, 2, 175-199.

Mayzner, M.S. & Tresselt, Mayzner, M.S., & Tresselt, M.E. (1965). Tables of single letter and digram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1, 13-32.

Mayzner, M.S., Tresselt, M.E., & Wolin, B.R. (1965). Tables of trigram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1, 3b.

McClelland, J.L., & Johnston, J.D. (1977). The role of familiar units in perception of words & nonwords. Perception nd Psychophysics, 22, 249-261.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. Psychological Review, 88, 375-407.

Miller, G.A., Bruner, J.S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. Journal of General Psychology, 50, 129-139.

Morris, P.E. (1981). Age-of-acquisition, imagery, recall and the limitations of multiple regression analysis. Memory and Cognition, 9, 277-282.

Ohala, J. (1974). Experimental historical phonology. In J. Anderson & C. Jones (Eds.). Historical Linguistics II: Theory and description in phonology. Amsterdam: North Holland, 1974.

Paivio, A., Yuille, J.C. & Rogers, T.B. (1969). Noun imagery and meaningfulness in free and serial recall. Journal of Experimental Psychology, 79, 509-514.

Perfetti, C.A., and Hoagaboam, T. (1975). The relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 74, 129-144.

Pick, A.D., (1978). Perception in the acquisition of reading. In F. Murray & J.J. Pikulski (Eds.), The acquisition of reading. Baltimore, Maryland: University Park Press.

Rosson, M.B. (1983). From SOFA to LOUCH: Lexical contributions to pseudoword pronunciation. Memory & Cognition, 11, 152-160.

rumelhart, D.E. (1977). Toward an interactive model of reading. In S. Dornic (ed.), Attention and Performance VI. Hillsdale, NJ, Erlbaum Associates.

Santa, J.L., Santa, C.M., & Smith, E.E. (1977). Units of word recognition: Evidence for the use of multiple units. Perception and Psychophysics, 22, 585-591.

Seymour, P.H.K., and May, G.P. (1981). Locus of format effects in word recognition. Paper presented to the Experimental Psychology Society, Oxford.

Shallice, T. (1981). Neurological impairment of cognitive processes. British Medical Bulletin, 37, 187-192.

Shallice, T., & Warrington, E.K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K.E. Patterson, & J.C. Marshall (Eds.), Deep Dyslexia. London: Routledge & Kegan Paul. 119-145.

Shallice, T., Warrington, E., & McCarthy, R. (1983). Reading without semantics Quarterly Journal of Psychology, 35A, 111-119?

Shannon, C.E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379-423; 622-656.

Shulman, H.G., Hornack, R., & Sanders, E. (1978). The effects of graphemic, phonetic and semantic relationships on access to lexical structures. Memory and Cognition, 6, 115-123.

Smith, P.T., and Baker, R. (1975). The influence of English spelling patterns on pronunciation. Journal of Verbal Learning and Verbal Behaviour, 15, 267-285.

Steinberg, D., & Krohn, r. (1975). The psychological validity of Chomsky and Halle's vowel shift rule. In E. Koerner, J. Odmark, & J. Shaw (eds.) The Transformational-Generative Paradigm and Mod n inguistic Theory. Amsterdam: John Benjamin.

Sternberg, R.J., (1977). Componential investigations of human intelligence. Presented at NATO Conference on Cognition and Instruction. Amsterdam.

Stoke, S.M. (1929). Memory for anomatopes. Journal of Genetic Psychology, 36, 594-596.

Thorndike, E.L., & Lorge, I. (1944). The teacher's wordbook of 30,000 words. New York: Bureau of Publications, Teacher's College, Columbia University.

Travers, J.R., & Olivier, D.C. (1978). Pronounceability and Statistical "Englishness" as determinants of letter identification. American Journal of Psychology, 91, 3, 523-538.

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

Wallach, M.A. (1963). Perceptual recognition of approximations to English in relation to spelling achievement. Journal of Educational Psychology, 54, 57-62.

Yarmey, A.D., & Paivio, A. (1965). Further evidence on the effects of word abstractness and meaningfulness in paired-associate learning. Psychonomic Science, 2, 307-308.

D747008 7

END