Using word frequency and parafoveal preview to determine the locus of contextual predictability and imageability effects: Evidence from eye movements during reading and lexical decision

Aisha Shahid, BA Hons.

School of Psychology

University of Glasgow

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy (PhD)

Institute of Neuroscience and Psychology

College of Science and Engineering

University of Glasgow

May, 2014

Abstract

The present thesis examines the time course of two semantic variables, contextual predictability and word imageability. Both variables can be said to reflect semantic aspects of meaning. For example the contextual predictability of a given target indicates the semantic context within which the target word occurs. The imageability of a given word reflects the meaning aspects of the word itself (Whaley, 1978). The word frequency effect (the faster response to commonly used high frequency words compared to low frequency words which occur less often) was taken to index the moment of lexical access (Balota, 1990; Pollatsek & Rayner, 1990; Sereno & Rayner, 2003) and by applying the logic of additive factors method (Sternberg, 1969a, 1969b), we determined whether the combined effect of each respective semantic variable was additive or interactive. This allowed us to examine whether there are semantic influences on lexical access. Previous research has been undecided and the question remains as to whether semantic variables operate during the lexical access processing stage, or alternatively after lexical access, for example in the post-lexical stage (e.g., Hand, Miellet, Sereno & O'Donnell, 2010; Sereno, O'Donnell & Rayner, 2006). Another aim of the thesis was to address the issue concerning the information presented to participants in the condition of 'invalid parafoveal preview of a target' (e.g., Sereno & Rayner, 2000). Several criteria were identified as being important in order to make the assumption that parafoveal processing was successfully inhibited on the pre-target fixation. Another aim of the thesis was to investigate whether word frequency and contextual predictability of the parafoveal word affected parafoveal preview benefit. Preview benefit was calculated by subtracting fixation durations in a condition of 'valid' preview of the target with an 'invalid' preview of the target.

Experiment 1 utilised a lexical decision task to investigate the relationship between word frequency and the imageability of the word. Experiment 2 investigated whether the orthogonal manipulation of word frequency and contextual predictability led to an additive or interactive relationship between these two variables. Two pre-tests, the rating and Cloze tasks, were used to determine the predictability of the target. Experiment 3 and a further cross comparison of Experiments 2 and 3 replicated and extended Experiment 2 by additionally using an eye movement-contingent boundary change paradigm (Rayner, 1975). Experiment 4 examined the joint and combined effects of frequency, predictability and preview in a within-subjects design. A separate pre-test Cloze task was used to determine predictability of targets in their low and high predictable contexts. This experiment used a larger set of materials than in the previous experiments to examine these

variables. Finally Chapter 6 was an overall discussion of the thesis. It was concluded that display screen presentations in our eye tracking experiments led to very fast reading times (as well as more skipping) compared to past studies which have used dot-matrix display presentations. It is possible that faster fixation durations led to floor effects in conditions where reading times are already fast because of preferential circumstances of high frequency targets, high predictable contexts and being given a parafoveal preview of the target. Possible ways to counteract this floor effect as well as alternative experimental methods of investigation were discussed.

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature ___________________________________

Printed name _________Aisha Shahid_____________

Acknowledgements

I would like to thank my two supervisors, Dr Sara C. Sereno and Professor Patrick O'Donnell for their assistance, patience and guidance over the years. I appreciate your help and encouragement in making this thesis possible. I would also like to thank my husband, Tariq and my two children, Amna and Haroon for their support in my academic pursuits. This research was supported by a BBSRC studentship.

Overview of Table of Contents

Table of Contents

List of Tables

List of Figures

Abbreviations

%	Percentage of times the target was skipped
%Error	Percentage Error in making a lexical decision
%Skip	Percentage of times the target was skipped
ANOVA	Analysis of Variance
AoA	Age of Acquisition
FFD	First fixation duration
GD	Gaze duration
HF	High frequency
HI	High imageability
HP	High predictable
LF	Low frequency
LI	Low imageability
LP	Low predictable
MP	Medium/Moderate predictable
ms	milliseconds
ns	nonsignificant
PDP	Parallel distributed processing
RT	Reaction Time
SD	Standard Deviation
SE	Standard Error
Sent 1	Sentence 1 first pass reading time
SFD	Single fixation duration
TT	Total time

Chapter 1
Introduction

1.1 Overview of thesis

The focus of this thesis was to investigate when semantic variables affect the lexical access for a given word. By utilising the techniques of lexical decision and eye movements, the effects of contextual predictability and word imageability were examined. This was by orthogonally manipulating these variables with word frequency, respectively. A particular focus was on the locus of contextual predictability. Word frequency refers to how a common a particular word is in written language, with frequency counts obtained from corpus counts (e.g., for English words, the British National Corpus; BNC, http://natcorp.ox.ac.uk). The 'word frequency effect' refers to the faster response to commonly used high frequency (HF) words compared to low frequency (LF) words which occur less often. The effects of word frequency have been well researched and there is a wide consensus that the frequency of a word affects, for example, fixation time on that word where readers spend less time fixating high frequency (HF) words than low frequency (LF) words (e.g., Inhoff & Rayner, 1986). Furthermore, this effect of word frequency is generally considered as an index of the moment of lexical access (e.g., Balota, 1990; Pollatsek & Rayner, 1990; Sereno & Rayner, 2003). Contextual predictability refers to how predictable a given word is in a developing discourse depending on how predictable the preceding sentence context up to reading a particular word. A word is said to be highly predictable (HP) in text when it is more constrained by the prior context versus a low predictable (LP) words in text which is a word which is not constrained by the prior context. The 'word predictability effect' refers to the faster response to highly predictable words compared to low predictable words.

As with word frequency, numerous studies have shown that contextual predictability affects fixation durations where readers spend less time fixating words that are highly predictable in text than word which are low predictable in text (e.g., Rayner & Well, 1996). The temporal locus of contextual predictability remains unknown. For example, it could be that contextual predictability exerts its effect during lexical access or alternatively only has its effect in post-lexical stages of semantic integration. Thus, the experiments in this thesis explored the time-course of contextual predictability effects. This was by examining whether the combined effect of frequency and predictability was additive or interactive on word identification time. A number of different techniques were employed to obtain converging evidence across different experimental paradigms (e.g., Posner & McCandliss, 1993).

The visual word recognition (or just word recognition) system has largely been described in the literature as composing of several components: orthographic, phonological and semantic. What word recognition is (i.e., the process of going from the printed text to the meaning of that word) varies depending on the particular theory of word recognition. The end point of word recognition, or lexical access, is generally said to be the moment when the reader recognises that the stimulus is a word and thus all the information associated with the word (orthographic, phonological and semantic) is made available to the reader rapidly, if not immediately (Pollatsek & Rayner, 1990). One distinction which can be made is that of the role of semantics in achieving this end goal of lexical access. In particular, most models of word recognition suggest that the word has to be recognised before the meaning of the word is determined. However, if there is empirical evidence which suggests that semantic variables can affect the speed of lexical access, then this indicates that some of the standard assumptions about the architecture of the word recognition system have to be changed. This thesis will explore this issues to do with how best (based on empirical evidence) to characterise the architecture of the word recognition system. Specifically, the effect of semantics in word recognition can be investigated via ‘contextual predictability’, as mentioned above, since this variable can be seen to assess the effect of the semantic context in which a word is processed. In addition, however, the effect of semantics can also be seen to include the effect of the semantic attributes of the word itself. To this extent, a further aim of this thesis was to examine the locus of word imageability. Word imageability refers to the ease with which a word gives rise to a sensory mental image. As with contextual predictability, the locus of this variable (i.e., manipulated with word frequency) can tell us about the locus of word imageability. This is important since this variable can said to be a ‘semantic’ variable because imageability taps into the semantic aspect of meaning which varies from word to word.

It was reasoned that determining the locus of contextual predictability and secondly, word imageability was important in order to inform as to when these variables operate in word recognition. This was in order to assess the impact of semantics on lexical access to allow for a discrimination to be made between the ‘early-processing’ and ‘late-processing’ models of word recognition. Broadly speaking, early processing models are those that allow semantics to operate at the time of lexical access whereas late processing models posit that semantics only has a role once lexical access is complete. Furthermore, this served to address the unresolved theoretical issue of whether processing in the language system should be characterised as autonomous or as interactive. In addition, whether the combined effect of word frequency and contextual predictability is additive or interactive also has implications for recent serial or parallel computational models of eye movement control during reading. This latter issue was not specifically addressed in the current thesis; however, one avenue for future work is to investigate predictions from computational models since newer word recognition research has seen the development of detailed computational models of eye movement control (Radach & Kennedy, 2013).

Another aim of the thesis was to investigate the effect of parafoveal processing to the frequency-predictability interaction. There appears to be a consensus that orthographic and phonological but not semantic information is extracted from the parafoveal region (e.g., Rayner, Balota & Pollatsek, 1986; Starr & Rayner, 2001). In this thesis, we explore whether word frequency and contextual predictability of the parafoveal word affected parafoveal preview benefit. That is, if readers are able to extract more information from the parafoveal word when it is high versus low frequency (e.g., Inhoff & Rayner, 1986) as well as when is it of high versus low contextual predictability (e.g., Balota, Pollatsek, & Rayner, 1985) then on a theoretical level there are implications for processing models of word recognition.

1.2 The word recognition system

The architecture of the word recognition system has been described as modular or autonomous by some (Fodor, 1983; Forster, 1979) and interactive (McClelland, 1987) by others. Moreover, the definition of lexical access is closely intertwined with the theoretical perspective of the model. One working definition is that lexical access is the moment we recognize a stimulus is a word and all the information associated with the word, such as its meaning, syntactic class, sound, and spelling are accessed immediately, hence the word is recognized (Pollatsek & Rayner, 1990). A distinction is often made with post-lexical integration which refers to the processes which take place when lexical access is complete. This includes integrating the meaning of the lexical representation with the prior context. For example, semantic integration is when the reader acquires the overall meaning of the sentence as a whole. This is likely to be putting together the semantic representations of the individual words in the sentence.

The early psycholinguistic theories, for example, Forster’s (1976) serial search model and Morton’s (1969) parallel access model, used the term ‘mental lexicon’ to describe the mental store for words: the orthography (spelling), phonology (pronunciation) and semantic (meaning) information about words. Usually, orthography and phonology refer to the ‘form’ of the word. These early theories posited that words were represented as lexical entries containing that words orthographic and phonological information (stored separately) in the ‘mental lexicon’. The mental lexicon was the mental store of all the words a reader knows where there was one lexical entry for every word known by the reader. In these models, word recognition was viewed as the process of going from the printed letter string to the selection of a single entry stored in the lexicon. This means that, within such models, word recognition was synonymous with lexical access. This definition therefore strongly suggests that words are represented as (separate) lexical entries in memory. However, the idea of lexical entries is strongly disputed in connectionist models of reading (e.g., Seidenberg and McClelland, 1989) which argue instead that representations are distributed across sets of subsymbolic processing units. Moreover, distributed models incorporate semantics (and like orthography and phonology) follow the same rules of activation (Balota, Ferrarro & Connor, 1991). The issue therefore is very much to do with the impact of semantics – both at the level of the context within which the word occurs and semantic attributes of the word itself.

The early models of word recognition (e.g., Forster, 1979; Morton, 1969) had two assumptions. The first was that human information processing comprised a series of processing stages which operated in ‘series’ that is, one at a time in a non-overlapping manner. This means that information flow through the language processing system was in one (forward) direction and that processing in one stage had to be completed before processing in the next stage could begin. The term ‘thresholded’ was used to refer to this completion of one stage prior to processing in another stage taking place. This meant that a stage was ready to pass information to the next stage only when the activation at the initial stage reached a threshold. The contrasting operation is that information passes between stages as soon as information at one stage begins to be activated; this is the suggestion in ‘cascaded’ models (e.g., McClelland, 1979). Forster’s (1976) model also had a second assumption which was that was that the language processing system is an autonomous system. This means that the system only works with the information stored within it, such as stored lexical information (Forster, 1981). It is this second assumption that will be explored further in this thesis.

1.2.1 The role of semantic information in the word recognition system

One defining feature between the autonomous models of word recognition and the interactive views is that of the role of semantic information (both the semantic context within which the word occurs or the meaning of the word itself). The early models of Forster (1979) and Morton (1969) proposed, broadly, that there is first a perceptually based process which leads to the activation of sublexical units (such as letter units). The activation of these sublexical units allows for some kind of 'prelexical' code to form. This prelexical code activates those words (that is, lexical) units which are more or less consistent with it. Therefore, at the end point, one of these units is selected or accessed. Only at this end point does meaning start to become activated. This assumption that meaning activation strictly follows lexical selection or lexical access is called the 'form-first' assumption (Forster & Hector, 2002). Therefore, these early models of word recognition based on the principles of autonomy and thresholded processing suggested that there were no influences of higher order processes on lexical access. In contrast, connectionist models involved the role of semantic information with this being represented no differently from the other types of information about the word (orthography and phonology). Hence, these models indicate that higher order processes can influence early lexical access.

1.2.2 Autonomous and interactive theoretical views of the language processing system

The architecture of the language processing system continues to be debated (Coltheart, 1999; Lucas, 1999). The sub-systems or components of the system that contribute to language processing are agreed by most (e.g., orthographic, phonological and semantic). However, the relationship of the various components continues to cause disagreement. Broadly, the view of the modular position is that the various components are functionally autonomous (i.e., unaffected by feedback) (Fodor, 1983; Forster, 1979). However, the position of the interactive view is that feedback from higher levels of processing are able to affect the operations of lower-level component processing (McClelland, 1987; Seidenberg, 1985).

1.2.2.1 The autonomous architecture of the language processing system

In line with the information processing approach, the language processing system can be seen to comprise a number of processing components: a lexical processor, a syntactic processor, a message processor, and a general purpose central processor. Modularists (e.g., Fodor, 1983; Forster, 1979) argue that the flow of information through the components of the language processing system is in a serial – sequential – fashion (i.e., bottom-up), in which the components are regarded to be autonomous and highly specialised input modules which feed into a more general purpose central processor (see Figure 1.1.). These input modules – lexical, syntactic and message processors – convert their particular inputs into representations that can be dealt with by the central processor. Moreover, these input modules have the properties of being 'informationaly encapsulated' (i.e., dedicated to its own particular task independent of influences from other sources) and 'domain-specific' (i.e., dedicated to processing specialised representations). Furthermore, Fodor (1983) argues that the central processor receives only the outputs of the input modules but not the full information that the module uses in its internal computations. In this conceptualisation, the input modules are responsible for the lower-level aspects of language processing (e.g., word recognition or lexical access) whilst the central processor is responsible for higher level aspects (e.g., aspects of discourse such as semantics and pragmatics). Thus, when we are reading, the purpose of the lexical module is to deal with lexical access and therefore, this module identifies the individual words we are reading without receiving any information from the syntactic or the message processor. The only way in which the lexical module interacts with the rest of the language processing system is that it sends the identified word on for further processing to the syntactic module. The purpose of this syntactic module is to parse individual words in the sentence into their constituent parts. Although the syntactic module has communicated with the lexical module (since the syntactic module has received the identified word from the lexical module), it does not receive any information from the message processor. Furthermore, the syntactic module operates independently of whatever processes are occurring in the other components. Finally, since the message processor constructs a discourse representation of the words that are read, it receives output from the other two processors but does not send feedback to either the lexical or syntactic processor (Fodor, 1983; Forster, 1981).

1.2.2.2 The interactive architecture of the language processing system

In contrast to the modular conceptualisation of the word recognition system, there are other visual word recognition models which have an incorporated an interactive approach (e.g., McClelland, 1987; Seidenberg, 1985). According to this view, information flows in both bottom-up and top-down ways through the language processing system in what McClelland calls the *interactive activation* framework: processing in any one component can impact on processing in any level above or below that component in the system (therefore contrasting strongly with the notion of sequential flow in a modular architecture). Figure 1.2. shows the 'triangle framework' which characterises the parallel-distributed processing models. This idea of interactivity has been taken up to varying degrees in what are called ‘activation’ models.

1.2.3 Word recognition models

The sections below describe some of the autonomous and interactive models of word recognition which exist in the literature. The attempt here is to describe how word recognition is thought to take place as well as well as what a given model indicates about how higher level semantic information is handled in the word recognition system.

1.2.3.1 Search models

Search (or serial search) models (Forster 1976; 1989; also Becker, 1979; Glanzer & Ehrenreich, 1979) are those based upon the assumptions of thresholded, autonomous processing (see Section 1.2 above). The conjecture of these models is that readers recognise a word by comparing a prelexical code to a set of lexical candidates, one at a time, until a match is found. The search process does not unfold over all of lexical memory but rather some process informs that a particular part of lexical memory is the best area for searching. In addition, it is assumed that the search set is organised in such a way that higher frequency words are checked before lower frequency words.

1.2.3.1.1 The bin model

In Forster’s (1976; 1979) autonomous search model, the lexical system has three peripheral access files and a master file. Altogether, these files contain information about all the words a reader has in their lexicon. The access files are those containing orthographic codes (for reading), as well as phonological codes (for speech perception) and semantic codes (for speech production). All these peripheral files are a way of accessing word entries held in the master file where all the information about the word is stored. In terms of the orthographic access file (which is relevant to reading) each word a reader knows has an entry and moreover each entry in this file has an ‘orthographic access code’ which are examples of the orthographic properties of the word and secondly, there is a pointer to that entry contained within the master file.

Forster’s model stressed the importance on autonomous word recognition which is not affected by attentional control. In particular this model is in agreement with modular views of lexical processing because in Forster’s model, the visual presentation of text is what drives the search through the orthographic bin regardless of attentional control.

Figure 1.1. The modular architecture of the language processing system

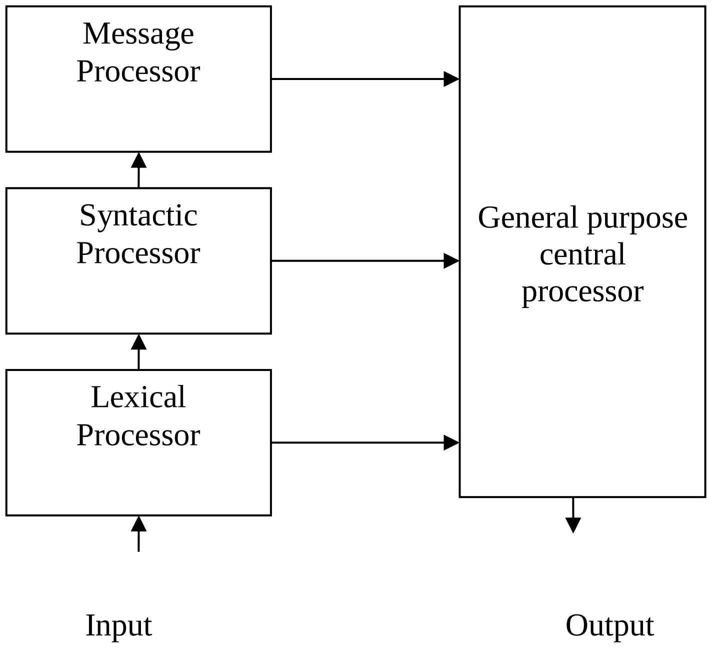
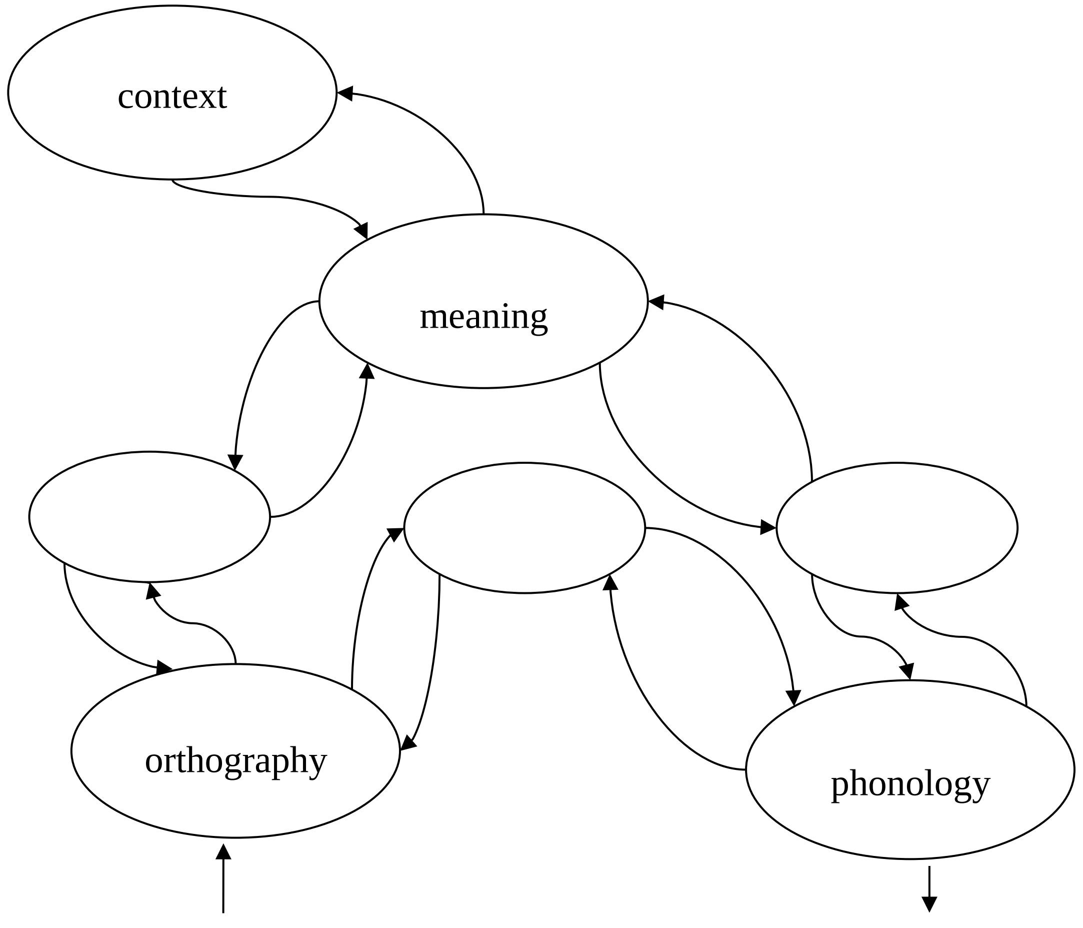


Figure 1.2. The interactive (parallel-distributed processing; PDP) architecture of the language processing system



Note: blank nodes represent hidden layers

According to this model, there is a limited impact of semantics on the word recognition system due to the structure of the master file. One implication of this proposal is that contextual predictability can only affect lexical access because of connections between individual lexical items which are in the sentence. That is, contextual predictability effects on lexical access occur because of word-to-word priming between words in the sentence rather than priming from the integrated sentence representation. Aside from this then, contextual predictability is envisaged to have a late time course in word recognition, for example, on later stage processes such as semantic integration.

1.2.3.1.2 The activation-verification model

Paap et al.'s (1982) activation-verification model is also an autonomous (modular) model. However, unlike Forster's (1976, 1979) model, the surmised architecture allowed for cascaded processing. That is, serial processing takes place in that first letter units and then word units are activated but processing is also cascaded in that information passes through the system even when initial processing is not complete. The activation of letters takes place in position-specific channels and it is suggested that this is a feature matching process. This means there is the possibility for an incorrect but featurally similar letter to be activated at each letter position. At the letter level, the activity constantly feeds into the lexicon so that when a lexical unit is activated, it is the result of the level of activity of each of the letters of that word.

Sets of word candidates which are chosen for additional processing is determined by the activity levels in the lexicon. Additional processing is the verification process which is when the sets of word candidates are serially verified against the perceptual representation i.e., the search process. In the case of there being enough of a match between a candidate and the perceptual representation, that candidate is then accepted and the verification process ceases.

As in Forster's model (1976; 1979), the architecture proposed in the activation-verification model does not allow for the impact of higher order semantics on the word recognition system. The suggestion then is that higher order contextual predictability has a later time course in word recognition, for example, in semantic integration stages.

1.2.3.2 Activation models

In contrast to the serial search process in search models, an alternative possibility is that the mind is more sophisticated than this because millions of neural connections make up the mind and therefore it could be that the mind is actually able to process words in parallel (i.e., more than simply one word at a time). These models are labelled parallel- (or direct) access models, of which the earliest version was the logogen model (Morton, 1969). Another form of parallel-access models is connectionist models (e.g., McClelland & Rumelhart, 1981) which is basically an adaptation of the logogen theory.

Other work has also included the basic interactive activation framework i.e., activation and inhibition processes as proposed by McClelland & Rumelhart (1981). For example Balota (1990) and Balota, Ferraro and Connor (1991) proposed a word recognition framework which incorporated the (early) impact of semantics on lexical access; Balota's model is a small adjustment of the basic interactive activation framework. Also Coltheart et al., (2001) proposed the dual route cascaded model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) (of reading aloud i.e. how the pronunciation of a printed word is generated). These two latter models embrace the interactivity proposed in the original McClelland and Rumelhart's (1981) model to varying amounts.

Finally, in contrast to representations being 'local', another set of models argue that representations are 'distributed' with these latter models being known as 'parallel distributed processing' (PDP) models (e.g., Seidenberg & McClelland, 1989). Basically, the idea of interactivity as originally proposed by McClelland and Rumelhart (1981) is embedded in 'activation' models to varying degrees – and the concept of interactivity contradicts the autonomy assumptions which characterise the search models.

1.2.3.2.1 The logogen model

The logogen model (Morton, 1969, 1979) used the activation principle and the concept of threshold. That is, where in the bin model words were accessed by locating their place in the lexicon, in the logogen model, word are activated to a certain threshold. The logogen model details how activation can be thought to take place as well as what influences the threshold of a particular word. In the logogen model, each word is conceived of having its own 'logogen' which can be thought of as a score sheet of the number of features that a lexical entry shares with the incoming perceptual stimuli. Each logogen has a threshold which can be thought of as encompassing the total 'energy' which will be needed to access that lexical entry. As input enters the system i.e., a reader reads a word, activation rises in logogens based on the incoming stimuli's orthographic, phonological and semantic information. All the information is accepted and totalled in parallel as the various logogens which have been affected by the incoming stimuli in a 'race'. Logogens are accessed when the activation reaches a certain threshold (which is determined on their being enough of a similarity to the input stimulus word). When a word is recognised, it is said that an entry has been activated to threshold and there is enough of a match between the input stimuli and the rise in information in the logogen. Many logogens are likely to be activated but the first logogen to reach a certain threshold of activation is recognised.

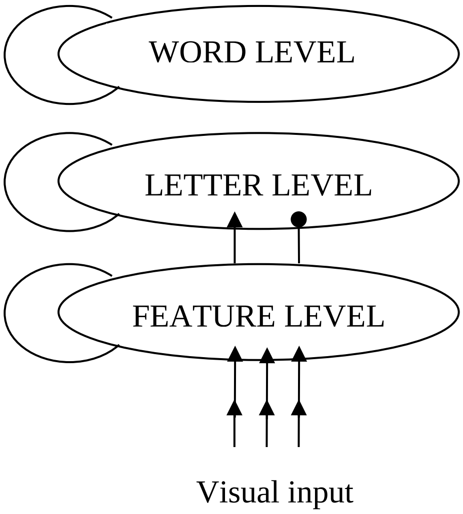
What is defining within this model is the supposition that readers use all of the available information (in order to unite on usually a single candidate in the lexicon). For example, readers will make use of a sentence context suggesting some meanings over others, presumably by narrowing the list of candidates which are activated. By creating a set of candidates, the lexical entries are accessed, that is, word meaning is made available. This means that this semantic information (information about the meanings of the word) is used to boost the activation of candidates that are semantically appropriate to the sentence context whilst lowering the activation of those which are not. In contrast, a modular view would argue that to bring in meaning at this stage serves as a huge complication which may distract from the task of assessing the evidence presented to our eyes.

1.2.3.2.2 The interactive-activation model

Whereas serial models of access (Forster, 1976; 1979) assume that we work through lexical entries in turn until we find a match in the stimulus, in parallel models of access (Morton, 1969, 1979) potential word matches are represented as being in competition with each other; evidence builds up for and against each candidate until a correct match is made. The basic features of the parallel-access logogen model are recognised as preceding the later formulation of McClelland and Rumelhart's (1981) interactive activation model. For example, in this latter model a basic tenet is the notion that a reader can use multiple cues to identify a word: the features of the letters, the sequence in which the letters appear, the word as a whole, the context within which a word occurs. Interactive models of word recognition suggest that all these cues can affect activation – and can do so simultaneously. In computer models of lexical access, nodes and connections between nodes make up 'neural nets'. There are three types of nodes: input nodes (which process the visual stimuli); output nodes (which decide responses) and hidden nodes (these perform the processing from when we see a word and our response to it). Thus, most of lexical processing is done by the hidden nodes.

McClelland & Rumelhart's (1981) model is a highly interactive language processing system (also exemplified in models of reading-aloud in the literature; Coltheart et al., 2001; Grainger & Jacobs, 1996). In particular, McClelland and Rumelhart proposed a computational model for visual letter recognition in which three levels of representations are assumed: a layer for the features of the word, a layer for the letters of the word and a layer for the representation of words (see Figure 1.3.). Hidden nodes represent different facets of words such as their visual, orthographic, phonological and semantic aspects. At the start of processing, there is a continuous flow of activation from feature-level to letter-level to word-level but also a downward flow of activation from word-level to letter-level to feature-level (that is, to lower-level representations; this is called feedback activation). The model allows for excitatory and inhibitory links between all the layers as well as the nodes in the same layer. For example, when a reader reads a particular word such as 'house', the letter 'h' in house gets some extra activation from the word level representation 'house'. Thus higher level lexical representations add to the lower level letter representations. This is because the processing system allows for information to flow in a cascadic manner, which means that information from one level is able to affect higher and lower levels in the system even though processing at any of the individual levels is not complete.

Figure 1.3. Interactive activation model (McClelland & Rumelhart, 1981)



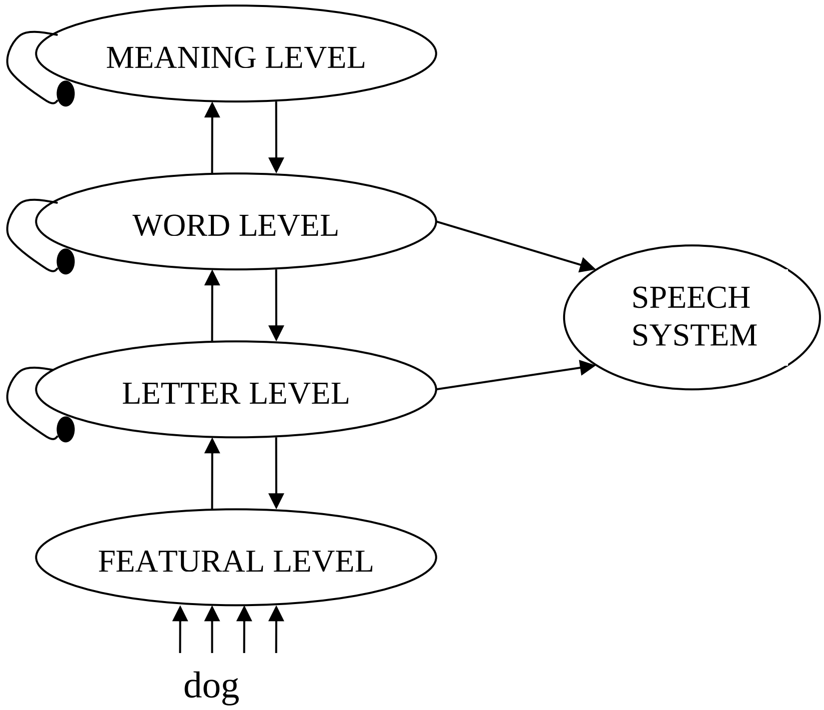
Unlike the preceding logogen model, in the interactive activation model, it is not necessary for information at one level of representation to reach some certain threshold before being passed onto another level of representation. Since information flows in a bidirectional way (that is, information travels from letters to words as well as from words to letters) between adjacent layers, a very defining aspect of this model is that the proposed architecture allows for the role of top-down information. That is, information from the higher cognitive levels i.e., information from word-level representations are able to affect earlier processes which take place at lower levels of representation (the letter-level and/or the feature-level). This therefore contrasts with the modular architecture exemplified in serial search models such as that by Forster (1979). Another difference between modular models and the interactive approach in McClelland and Rumelhart's model is that processes can take place in parallel, for example, several letters can be processed at the same time. This notion contrasts with an earlier model of letter recognition (Selfridge, 1959) in which it was suggested that only one letter is processed at a time in serial way.

The interactive-activation model (McClelland & Rumelhart, 1981) was originally formulated to explain the effects of higher level information on lower-level processing, in particular the word superiority effect (Reicher, 1969; Wheeler, 1970). The word superiority effect is the finding that participants are better able to identify letters (i.e., lower-level information) when they are presented in words rather than in non-words (e.g., Johnston & McClelland, 1973; Maris, 2002; Paap, Chun, & Vonnahme, 1999; Paap et al., 1982).

Therefore, McClelland and Rumelhart's (1981) model is able to explain the impact of semantics, both the context within which the word occurs (that is, the effect of contextual predictability) as well as the semantic aspects of the word itself (that is, the effect of word imageability). In particular, semantic effects arise because of higher level input, that is, semantic information, affecting word-level representations via feedback activation. The implication from this model regarding the locus of higher level semantic information is that contextual predictability (as well as word imageability) can affect the process of lexical access suggesting that semantic context is operating on a word the instant it is read.

In subsequent work, Balota (e.g., Balota, 1990; Balota et al., 1991) offered a small adjustment to the basic interactive activation framework proposed by McClelland and Rumelhart (1981). Balota's model is displayed in Figure 1.4. The extent of the (highly) interactive system proposed in McClelland and Rumelhart's (1981) model is taken up to a lesser degree in Balota's framework. That is, this model accepts only a certain extent of the interactivity proposed in the original McClelland and Rumelhart (1981) model; whereas McClelland and Rumelhart's model is that of a highly interactive language processing system in contrast, Balota does not commit the full extent of this interactivity.

Figure 1.4. Balota's model of possible interactive activation framework for word recognition with meaning-level influences



Balota's framework proposes several levels leading to the semantic access (the meaning features) of the word. There are the same three levels as in McClelland and Rumelhart's (1989) interactive activation model: a level representing the features of the word, a level representing the letters of the word and a level representing the meaning of the word. It should be recalled that in McClelland and Rumelhart's model, activation occurs in a bidirectional way from lower level feature and letter representations to higher level word representations as well from higher level word representations to lower level features and letter level representations with inhibitory pathways within a level. For example, because of the cascadic processing system, higher level lexical representations (such as the word 'house') contribute to the lower level letter representations (such as when reading the letter 'h' in the word 'house'). Thus the implication is that information from one level affects higher and lower levels in the system without the completion of processing at given level (McClelland, 1979). In addition to this, Balota's framework adds on a meaning level representation above the lexical level representation which has both bottom-up and top-down pathways operating from it; this framework is shown in Figure 1.4.

Balota's framework, shown in Figure 1.4., involves a serial flow of information. That is, the structure is seen as a bottom-up flow of information from features, to letters, to words, and to meaning. Word level and meaning level information have a facilitating effect at lower feature and letter levels subsequent to earlier processing in the system being sufficiently activated in order to pass on activation to these higher levels. This means that Balota's framework suggests that earlier aspects of lexical processing, for example identifying the features and letters of a given word have to reach a certain level before the meaning of the word becomes available. The implication is that higher level contextual predictability can speed up only later aspects of processing whereas in an architecture like McClelland and Rumelhart's, 1989 interactive activation framework suggests that higher level contextual predictability can speed up very early aspects of processing, i.e., contextual predictability speeds up the access to the form of the word such as the features and the letters as well as the meaning of the word being available to narrow down lexical candidates. Hence, the locus of contextual predictability within Balota's framework would be relatively later in lexical processing.

Another subsequent model which has incorporated the processing mechanisms of the interactive activation system (McClelland & Rumelhart, 1981) as well as the work of Morton (1979) is the relatively recent dual-route cascaded model (of word recognition as well as reading aloud). In particular, the basic interactive activation system is used to describe the dual-route cascaded model's 'lexical pathway'. Briefly, the dual-route model has three processing pathways, two lexical pathways and non-lexical pathway. Word-level representations are contained in the 'orthographic input lexicon' and this is connected to a 'phonological output lexicon' which contains the phonological codes of the words the reader knows; this phonological lexicon has a connection with the phoneme system to permit phonological codes to be changed into speech. In the non-lexical pathway, the printed letter string is changed to it sound because of the connection between letter-level representations to the phoneme system. There is also a semantic system which connects the orthographic input lexicon with the phonological output lexicon and through feedback operations can explain semantics.

Like the interactive activation model by McClelland and Rumelhart (1981), Coltheart et al.'s (2001) model can account the effects of contextual predictability and word imageability. This is because it is envisaged that information about the printed stimulus passes through all the pathways in a cascaded manner. In particular, for a particular printed stimulus, semantic (and phonological) representations can be activated prior to the activation of a node in the orthographic lexicon amassing to a critical recognition threshold. Specifically, bidirectional connections in place between orthographic (and phonological) systems with the semantic system mean that the semantic information can affect the increase in activation of nodes in the orthographic lexicon (though semantic information is not necessary in order to recognise a word) (e.g., Coltheart, 2004; Coltheart et al., 2001). For example, information about the printed stimulus passes through all the pathways in a cascaded manner so that, for example, when a reader reads the word 'house' this will activate the 'house' node in the orthographic lexicon because of feedforward activation coming from the letter nodes. At the same time, the orthographic node for 'house' provides further supporting activation from semantic nodes which have been activated from the phonological nodes and/or orthographic nodes.

According to the dual-route cascaded model (Coltheart et al., 2001), bidirectional connections are in place between orthographic (and phonological) and semantic information. This means that this model can account for semantic influences, such as those of contextual predictability and word imageability on word recognition. In particular, this model suggests that higher level semantic information such as contextual predictability and word imageability can occur early in lexical processing.

1.2.3.2.3 Distributed processing models

All the models discussed thus far vary in terms of the degree to which they allow parallel processing as well as how autonomous lexical processing is. However, there is one assumption which is inherent in these models and this is that the crucial word recognition process is in isolating (i.e., 'selecting' or 'accessing') the appropriate lexical unit. However, models which come under the description of 'distributed processing' models do no adhere to the notion of the lexical unit. That is, it is assumed that words are *not* represented as lexical entries in memory. Instead, representations are viewed to be *distributed* across sets of subsymbolic processing units. This means that the lexical system is assumed to be made up of sets of distributed, sub-symbolic codes which serve to represent all the aspects of the words we know in our language. In terms of the word recognition process, this is seen as a process of activating the correct set of the sub-symbolic codes.

The first parallel distributed processing model was theorised by Seidenberg and McClelland (1989). The basic conjecture of this model was that the word recognition system comprised three kinds of mental representations relating to aspects of words (therefore, units of orthographic, phonological and semantic representations). Figure 1.2. in Section 1.2.2.2. exemplifies such a 'triangle' structure. Each of these units of representations is connected to the other representations. The correct connections between sets of units are obtained via learning. For example, when a young child learns to read, they learn the rules for sounds of each letter, how individual letters combine, illegal combinations. According to Seidenberg and McClelland’s (1989) model, learning is viewed as an error correction process. When the system is presented with a word, units at all levels activate and inhibit other units which results in activation over all units. Such activation patterns are at first somewhat incorrect but by being compared to correct patterns, weights between units are altered which means that the next time, processing becomes more accurate than the initial time. Every time the system experiences the stimuli, this adjustment process takes place. This means that over time, the correct activation in units is produced. For example, when a reader is presented with the word ‘apple’, orthographic processing of the visually presented stimulus means that the phonological units for the phoneme sequence of ‘apple’ are also activated. Within this model, hidden units are also assumed. The exact nature as well as amount of orthographic, phonological and semantic representations depends on the particular parallel processing model (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996).

Seidenberg and McClelland’s (1989) model (as well as other models within parallel processing) include the semantics into the word recognition system in the same way that orthography and phonology are, that is, the representation of semantics is not any different to the other representations. This means that all these representations are assumed to follow the same rules of activation. The interactivity in these models is due to feedback processes which are assumed to occur between the semantic units and the lower-level orthographic and phonological units. Thus, the implication of these models is that higher level information such as semantic information affects lexical access.

1.2.4 Models of eye movement control

The word recognition field has also seen the implementation of detailed computational models of eye movement control. These can be seen to be more advanced models of reading or word recognition. It is beyond the scope of this thesis to discriminate between these computational models of which there are now several (Reichle, Rayner & Pollatsek, 2003). The starting point of computational models was the largely descriptive models such as the Strategy-Tactics model (O’Regan, 1990, 1992). Previously, there were models of reading proposed by Morrison (1984) and Just and Carpenter (1980). The more recent trend appears to be detailed quantitative models, of which the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003) was one of the first. In fact, the E-Z Reader model is essentially an extension of the earlier descriptive model of reading by Morrison (1984). There are now many other models such as SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Kliegl & Engbert, 2003) and GLENMORE (Reilly & Radach, 2003).

Since the focus of the current thesis is on empirical data from experimental manipulations to increase our understanding of reading processes, only a brief overview of two computational models is provided here: the E-Z Reader model and the SWIFT. The reason for discussing these two particular models is they are sometimes said to be the two main competitors. Basically, these models have differing explanations of the allocation of attention during reading. 'Serial attention shift' models (of which the E-Z Reader model is one example) agree on the assumptions that attention is allocated sequentially such that lexical processing of one word at a time takes place and that it is lexical processing that causes the eyes to move from one word to the next. 'Processing gradient' models (of which the SWIFT model is one example) instead suggests processing which is spatially distributed according to an attentional gradient.

The E-Z Reader model (e.g., Reichle et al., 2003) postulates that two stages of lexical processing commence when the reader fixates a word. The first stage ('L1') is an initial stage which is a familiarity check where only the orthographic representation of the word is activated. The saccade to the next unidentified word in the text takes place after L1. The second stage ('L2') can be said to be the lexical access of the word. That is, phonological and semantic activation takes place. When L2 completes, attention is shifted to the next word (i.e., n+1) and the two stages now start commencing for this word.

Whereas in the E-Z Reader model, serial lexical processing is envisaged, that is, where word n is identified, and then n+1 and then n+2 and so on, in the SWIFT model (e.g., Engbert et al., 2005) parallel lexical processing is permissible. That is, on a single eye fixation, more than one word can be processed. Many studies (e.g., Kennedy & Pynte, 2005; Rayner, Reichle, Drieghe, Slattery, & Pollatsek, 2007) have tried to discriminate between these models in terms of whether processing is serial or parallel and as yet a consensus has not been reached.

1.3 Inferring stages of processing

From our discussion thus far, it should be clear that autonomous models of word recognition (search models of word recognition which embody the principles of autonomy and thresholding) assume that a given word must be recognised before we can access the meaning of the word. Within such models then, word-level or lexical processing precedes processing at higher levels of analysis. The principles of autonomy and thresholded processing indicate that there are no influences of higher order processes on lexical access. Thus, semantic constraints such as contextual predictability as well as the effect of the semantic variable word imageability on early lexical access are controversial within search models.

When do semantic variables (contextual predictability and word imageability) affect lexical access? Is it only after lexical access or alternatively at the time of lexical access? It should be clear that the individual effects of each respective variable (contextual predictability and word imageability) are not disputed. The aim is to orthogonally vary each respective variable with *word frequency* (using the logic of additive factors method; Donders, 1868/1969; Sternberg, 1969a; 1969b) in order to infer processing stages within the language processing system.

Specifically, the word frequency effect represents the faster response to commonly-used high-frequency words compared to low-frequency words which occur much less often. That is, several eye movement studies have shown that readers spend more time fixating lower frequency (LF) words than high frequency (HF) words (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennedy, Pynte, Murray & Paul, 2012; Murray & Forster, 2004; Rayner, 1977; Rayner et al., 2004; Rayner & Duffy, 1986; Rayner & Raney, 1996; Rayner, Sereno, & Raney, 1996; Schilling, Rayner & Chumbley, 1998; Sereno, O'Donnell & Rayner, 2006; Sereno, Pacht & Rayner, 1992; Sereno & Rayner, 2000). In the ERP literature, several studies have shown frequency effects in ERP components with early time intervals (before 200 ms) (Braun, Hutzler, Ziegler, Dambacher, & Jacobs, 2009; Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006; Hauk & Pulvermüller, 2004; Penolazzi, Hauk, & Pulvermüller, 2007; Sereno, Brewer, & O'Donnell, 2003).

Therefore, the word frequency effect can be taken to index the moment of lexical access. The reasoning is that the word frequency variable can be examined with a number of other (semantic) variables which in the present thesis are contextual predictability and word imageability in order to determine the locus of these variables. A frequency by contextual predictability (or word imageability) interaction would suggest that these two variables share the same processing stage, specifically that they have an early joint effect on the lexical access stage. The implication would be that semantic information is made available in the early stages of word identification, presumably at the actual time we are matching the perceptual input to the stored word in the lexicon that is, during lexical access. If, however, frequency and contextual predictability (or word imageability) have an additive relationship, then this would suggest lower-level processes and semantics are occurring at different stages of processing. Such a finding would imply that semantic information is accessed after lexical access (so only orthographic and phonological information has been used to access the word) and that semantic information is only made available in later post-lexical stages such as semantic integration, for example when we begin to integrate the word into the meaning of the discourse. Furthermore if two variables are additive, then the suggestion of different stages of processing suggests modular or serial processing in the system; if two variables are interactive, the suggestion that they are occurring at the same processing stage supports cascaded processing in the system. These implications for the architecture of the language processing system can be deduced from the forthcoming experimental manipulations of word frequency and contextual predictability and of word frequency and word imageability: an additive result would be the ANOVA showing two significant main effects and a non-significant interaction whereas an interactive result in the ANOVA would be indicated by two significant main effects as well a significant interaction.

1.4 Measuring word recognition processes

To determine the locus of contextual predictability and word imageability effects, it is essential to measure the necessary processes. Since word recognition processes are fast (e.g., Sereno & Rayner, 2003), dependent variables are needed which can measure such processes across different experimental conditions to provide the essential millisecond-level information. Experimental tasks which have been employed by researchers include behavioural methods, namely the lexical decision task; measuring participants eye movements while they are reading text which has been embedded with a target word (which has been manipulated on some variable or variables of interest); and recording event-related potentials (ERPs) again while participants read text with an embedded target. Each method reveals important information about the time course of word recognition processes; however it is important to be aware of their respective limitations.

Since the present goal is in discovering how the normal brain constructs meaning and how it does so in real time, a pertinent distinction is made between on-line and off-line techniques. On-line techniques measures variables which tap into word processing as it happens, whereas the latter, off-line techniques, measures variables which are the subsequent outcome of processing. Research questions to do with time-course can be seen to be particular informed by the use of on-line methods since they tap into comprehension processes at precisely the moment they are occurring (Sereno & Rayner, 2003).

1.4.1 Measuring eye movements

There is overwhelming consensus that monitoring participants' eye movements while they are reading is sensitive to on-line perceptual and cognitive demands of lexical processing with the fixation duration data collected expressing moment-to-moment cognitive processes (Juhasz & Pollatsek, 2011; Kliegl et al., 2006; Rayner, 1998, 2009; Sereno, 1992; Sereno & Rayner, 2003). Such a technique is considered advantageous to other experimental paradigms such as the lexical decision task and naming aloud tasks described above. While both tasks provide useful information about aspects of word recognition, they require overt responses (e.g., yes/no response in the lexical decision task) which are do not typify normal silent reading. The possibility then is that participants could engage in additional mental processing which do not occur during natural silent reading. However, with the eye movement method, the experimenter takes advantage of what is a natural occurring phenomenon – to read, it is necessary to make eye movements. As such, there is no need to administer a secondary task to make inferences about some comprehension process which may have taken place. Rather, fixation durations are assumed to reflect as least partly aspects of lexical processing which therefore allows us to infer underlying cognitive processes (Kliegl et a., 2006, Rayner, 1998, 2009).

The typical methodology for researchers interested in investigating word recognition processes is to manipulate one word (the target word) on some variable(s) of interest – such as word frequency. This target word is then presented in a passage of text (on a computer screen), and while participants' are reading the passage, the duration of fixations made on the target word is recorded. A beam of light which is invisible to the naked eye is channelled to the cornea of any one eye; this light is reflected and recorded by a camera positioned on the desk. Both vertical and horizontal locations of the eyes are sampled up to 1000 times per second (depending on the exact specification of the eye movement recording apparatus); to record such positions of the eyes, it is necessary to eliminate head movements using either a bite bar or a chin rest (Just & Carpenter, 1980; McConkie, Hogaboam, Wolverton, Zola, & Lucas, 1979; Rayner & Sereno, 1994; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

1.4.1.1 Measures of processing time

When researchers are interested in a single target word, that target word is manipulated (whilst embedded in text) on for example, its length, frequency etc. Thus whilst participants read the passage, the duration of fixations on the target word are measured. The eye movement record yields a number of dependent measures (see Table 1.1 for these measures and their definitions). As can be seen in this Table, these measures are subject to a number of distinctions: ‘earlier’ processing measures (also called first-pass measures); ‘middle’ processing measures; and ‘later’ processing measures (e.g., Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

Table 1.1 Earlier, Middle and Late Dependent Variable Measures from the Eye Movement Record

‘Earlier’ processing measures	**Explanation**
First fixation duration (FFD) on the target word	The duration of the very first fixation on the target word, provided that the word was not skipped, regardless of whether it is the only fixation or the first of many fixations on that word
Single fixation duration (SFD) on the target word	The duration of the first fixation on the target word if it received only one fixation on the first pass
Percentage skipping (%skip) of the target word	The percentage of times when the target word was not fixated on the first pass. Also reported as the probability of fixating on the target word
‘Middle’ processing measure	
Gaze duration (GD) on the target word	The sum of all fixations on a target word on the first pass prior to making an eye movement to another word
‘Later’ processing measure	
Total time (TT) on the target word	The sum of all fixations on the target word including any re-reading of the target word (regressions back to it)

First and single fixation durations reflect the first-pass processing time for a word, that is, before any regressions are made back to the target word. It is these early measures which can perhaps be assumed to reflect lexical access processes since these measures reflect the difficulty of processing the word. Thus, the reasoning is that when frequency and contextual predictability effects appear may be able to give an indication of the underlying processes. For example, effects that appear in the early measures, first and single fixation measures, may indicate the initial access to representations of a word's orthography, phonology, or meaning. 'Middle measures' are likely to reflect the underlying processes that occur between initial lexical access and later processing activities, as reflected in the gaze duration measure. Later measures such as the total time spent on a target word can be informative as to reflecting later processing activities. The differnt measures of eye movements have at times been interpreted in terms of specific time underpinnings of cognitive processes. Inhoff (1984) suggested that FFD reflected lexical access processes and that gaze duration was likely to be post-lexical integration processes. It is now clear that this view is too simplified. However, now by convention, 'earlier' measures (see Table 1.1) are more likely to reflect earlier aspects of cognitive processes.

In dealing with single target words, there are other measures which can be computed so as to provide a more informative picture of the underlying processing activities. For example, the likelihood of skipping the target word as indicated by the percentage of skips tells us in how many cases the target word was not fixated at all in the first pass reading and this measure can perhaps then be a useful measure of early processing activities. What might influence a given word being skipped? It could be that when readers are currently fixated on word *n*, they are able to process some aspects of word n + 1 such that the reader gets enough information concerning the identity of n + 1. This would lead to a skip – the reader does not need to fixate the word on the next fixation since the word has already been identified on the previous fixation.

Recent display characteristics of eye tracker machines have blurred the assumptions of the fixation duration measures as detailed above. That is, as we discuss in detail in the first Experiment in this thesis, newer machines use a clearer font compared to older one which used a more pixelated font. The impact of this is that whereas previously researchers used first, gaze and total time to describe the data, nowadays it is common-place to analyse the single-fixation measure as well (previously not analysed because skipping rates were low). Clearer font also means that skipping rates tend to higher than those in past studies (see Chapter 2, Discussion for a detailed discussion of these issues).

1.5 Semantic variables

1.5.1 Contextual predictability

As mentioned in the sections above, the 'contextual predictability' of a given target word refers to the semantic context within which a given target word occurs. More specifically, the contextual predictability of a target word is a measure of the extent to which readers predict a particular word at a given point in a developing discourse based on the context they have read up to that point.

This variable is operationalised for the purpose of using ‘contextual predictability’ in an orthogonal experiment. The most common methods are to use ‘norming tasks’ in particular the Cloze task (Taylor, 1953). In addition, sometimes a word rating task (also called the predictability task). To prepare these tasks, the experimenter first constructs sentence contexts in the form of single sentences or paragraphs. The purpose of the sentence context is to build up context sufficiently so that a designated target word appearing in the sentence frame is able to be easily predicted for the condition of ‘high predictable’, or alternatively difficult to predict for the condition of ‘low predictable’. Of course, the next step must be for the experimenter to test the validity of the experimental items and the respective condition they have been assigned to. This is where the norming task is administered to a separate group of participants (to those who take part in the main experiment.

In the Cloze task, participants are given the sentence frame up to but not including the target word. They have been instructed to write the word they think most likely comes next. At the end, the experimenter is able to collate participants’ responses. In the eye movement literature on word frequency by contextual predictability studies, the consensus seems to be that if words are guessed by participants less than 10% of the time, then they are classified as low predictable and if words are guessed more than 60% of the time, then they are classified as high predictable (see Table 2.1 in Chapter 2). Figure 1.5. below gives an example of a passage used in a Cloze task (used in Experiment 4 of the current thesis).

Figure 1.5. Example passages measuring contextual predictability of a target word (in bold) presented to participants

Passage A:

The man we had hired to replace our slates had a fatal accident.
The fall from the **roof** had killed him instantly on impact.

Passage B:

The man’s family were consoled that he had not suffered in agony.
The fall from the **roof** had killed him instantly on impact’.

In the example passage presented in Figure 1.5., the target word (roof) was assigned a higher contextual predictability score when it was read in the context given in Passage A, whereas in Passage B the target word roof was given a lower predictability score (the 'neutral' passage). It is important to note that when sentence contexts are low predictable or neutral, they are not actually anomalous. This is because it could be more representative to compare a condition of high predictable with low predictable when they both make sense given the use of words in the language.

The second method used by experimenters to assess how predictable their experimental items are is to use a word rating (or predictability) task. In this task then, participants are presented with the sentence frame up to an including the target word (often the post-target context is included as well). The target word is underlined or in bold font and participants have to rate on a scale (typically from 1-low predictable to 7-high predictable) how well they think the target word fits in with the context they have read. Thus, the higher the rating, the more predictable a given target is considered to be. The rating task is used less than the Cloze task because the Cloze task is considered to have more validity than the rating task. This is because in the Cloze task, participants have to actually come up with a word themselves as opposed to rating a word already provided which can be argued to be speculation regarding its predictability. Nonetheless, the rating task is useful to confirm the results obtained from the Cloze task.

The *word predictability effect* refers to the faster response to a target word appearing in a highly predictable context versus a target word appearing in a low predictable context. This effect has been accounted for by two factors (Rayner & Balota, 1992). First, when a word is predictable from its previous context, a reader is able to better use parafoveal information compared to when a word is low predictable from its prior context. Thus, highly predictable words are more likely to be skipped than low predictable words because the word was identified on the previous fixation. A second explanation is that contextual predictability effects occur from top-down active processes whereby the reader constructs the meaning of the previous sentence context and using their top-down knowledge sources allowing them to predict what word is likely to appear next. This means that sentence context effects can be said to reflect higher order comprehension influences. That is, the sentence context facilitates recognition even though there is *no semantic relation* i.e. no lexical associates between 'roof' and other words in the sentence. Therefore, this effect of sentence context is considered a *sentential priming effect* (in contrast to lexical priming). Priming from sentential context is priming which is not due to the associative effects of individual words in the sentence and therefore suggests top-down effects from higher order levels of discourse representation. For example, imagine a reader is reading passage A in Figure 1.5. above, which is a text about a man replacing slates having an accident. Knowing that the text is about a man working on a roof, might cause the reader to create certain expectations at a rather conscious level i.e., 'slates appear on the roof of a house'; 'falling from a roof is likely to cause serious injury'. Thus the wider context (what the reader already knows) influences interpretation of a given word. In this example, when the reader comes to read 'roof', top-down influences of world knowledge generates expectations to the extent that it facilitates processing of upcoming words. As such, in eye movement studies, the typical finding is that readers spend less time on 'roof' when it is read in passage A than when 'roof' is read in passage B. Therefore, sentential context causes facilitation: the sentence context which is the beginning of the sentence over lines 1 and 2, 'The man we had hired to replace our slates had a fatal accident. The fall from the ______' facilitates the recognition of a word such as 'roof', which is a highly predictable continuation of the sentence.

Many studies have demonstrated that readers spend less time fixating words that are highly predictable (HP) in text (i.e., words which are more constrained by the prior context) than low predictable (LP) words in text (i.e., words which are not constrained by the prior context) (e.g., Balota, Pollatsek & Rayner, 1985; Carroll & Slowiaczek, 1986; Ehrlich & Rayner, 1981; Hand, Miellet, O'Donnell, & Sereno, 2009; Kliegl et al., 2004, 2006; Lavigne, Vitu, & d'Ydewalle, 2000; McDonald & Shillcock, 2003a, 2003b; Morris, 1994; O'Regan, 1979; Rayner et al., 2004; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987; Zola, 1984; but cf. Hyona, 1993). In addition, HP words are more likely to be skipped than LP words (e.g., Altarriba et al., 1996; Balota et al., 1985; Brysbaert & Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Ehrlich & Rayner, 1981; Hand et al., 2009; Rayner et al., 2004; Rayner & Well, 1996). In studies using ERPs, studies have shown predictability in a late time window, the N400 component (Kutas & Hillyard, 1980, 1984; for reviews see Kutas & Van Petten, 1994; Kutas et al., 2006; Barber & Kutas, 2007) (see the concluding Chapter 6 for a more detailed discussion of ERPs).

1.5.2 Word imageability

The *word imageability effect* refers to the finding that words with higher imageability are responded to faster than words with lower imageability (Clark & Paivio, 2004; Paivio, Yuille, & Madigan, 1968). Word imageability is a semantic variable since it reflects the meaning of the word itself (Whaley, 1978). That is, a word's imageability refers to the ease with which the given word gives rise to a sensory mental image. This term has often been confounded with word concreteness in the literature. The two are highly correlated but nonetheless refer to different concepts. Word concreteness has been defined as the ability to see, hear and touch something (Paivio, Yuille, & Madigan, 1968).

In terms of the current thesis, it was considered that since word imageability can be seen to represent the 'internal semantics' of a given word, this variable can be examined with word frequency (using the logic of additive factors method) to see if there are imageability effects in early lexical processing. That is, the combined effect might be additive or interactive in the ANOVA which can tell us something about additive versus interactive language processing (see earlier section 1.3).

For experimental purposes, imageability scores for given words can be obtained from recognised norms where researchers have typically presented participants with lists of words and have asked them to rate each on a seven-point scale ranging from low to high imagery (e.g., Clark & Paivio, 2004; Paivio et al., 1968). In order to use as much of the information as possible, these norms can be combined with others (e.g., Cortese & Fugett, 2004; Gilhooly & Logie, 1980) in order to obtain a mean imageability score for a given word.

1.6 Are there semantic constraints on lexical access? The locus of semantic constraints and variables: Orthogonal manipulation with word frequency

We now try to attempt to distinguish between the models of word recognition. As stated previously, one way to discriminate between models rooted in autonomous or interactive views is to examine when semantics (meaning-level influences) have their effect in achieving lexical access. This is because the two views have quite different predictions about the locus of semantics, either on lexical access (e.g., the interactive view) or on semantic integration (e.g., autonomous positions). It should be recalled that since word frequency indexes the moment of lexical access, the approach we are utilising in this thesis in order to infer stages of processing is Sternberg's additive factors method. Thus, if two variables are additive, then they are occurring at different stages of processing, suggesting modular or serial processing in the system. However, if two variables are interactive, then this suggests that they are occurring at the same processing stage, which suggests interactive processing.

According to the modular view of the language processing system, semantics can have one of two possible effects on word recognition. It should be recalled that the effect of semantics refers to both the semantic context within which a word is read (a constraint) and also the semantic attributes of the word itself (the 'internal' semantics of the word itself). Thus, the first way is that semantics can ease post-access processes (such as the integration of the lexical entry which has been accessed into the higher order sentence representation; semantic integration). That is, semantics have their effect on post-access processes such as semantic integration. The second way is that semantic constraints, specifically sentence context, can influence lexical access but only when the sentence context contains words that have lexical associates with the target word. This is because words appearing in the sentence frame prime one another. The mechanisms of spreading activation are invoked where words are represented as associations with direct links amongst the lexical entries so that if one entry is activated, the activation can spread to the related entries, thus facilitating lexical access. It should be clear then, that in the absence of lexical associates (that is, if care is taken when writing the materials so as to avoid intra-lexical primes within the sentence frame), according to the modular view, the only influence of higher order semantics is on later stages of word processing such as semantic integration. If indeed semantic constraints exert their influence only at the post-access stage, the modularity view is supported because it is within these models that the lexical processor is conceptualised as autonomous – so impervious to higher order constraints (e.g., Forster, 1979). This means that the contextual predictability effect is explained due to later processes such as those of semantic integration.

According to interactive views of the language processing system, semantics can facilitate the process of lexical access (through limiting the activation of the number of lexical candidates). This is because of the feedback mechanism in the system where processing in the lexical component of the system is influenced by processing in any of the other components. This is due to the flow of information being multidirectional among all the components of the lexical processing system. Context effects on lexical access are also accounted for via intralexical priming through the notion of associations between lexical items and the principles of spreading activation. Moreover, the contextual predictability effect, within such a framework, is explained due to the ease of identifying the word.

1.6.1 What is the locus of the contextual predictability effect and of the word imageability effect?

In order to investigate the architecture of the language processing system, we can investigate whether the effect of semantics, as considered by contextual predictability as well as word imageability, are on lexical access or on semantic integration. Thus, the locus (i.e., where in the system) these semantic variables are operating can be investigated. By determining the locus of contextual predictability and of word imageability, we can establish whether semantic information is made available at the moment of lexical access, or alternatively, if semantics is only made available in later post-lexical stages. As discussed previously, the individual effects of first word frequency and second contextual predictability are well established in the word recognition eye movement literature. Specifically, it can be investigated whether contextual predictability (as well as word imageability) affects early, lexical processing or only later post-lexical processing.

It may be the case that contextual predictability has a late time course where it affects those later stages such as semantic integration which take place once lexical access is complete. The argument here is that we need to have the meaning of the word in order to integrate that meaning into the ongoing discourse being built up by the reader. Alternatively, contextual predictability may have an early time course such that sentence context affects the early stages of word identification, such as in speeding up lexical access for example by limiting the number of word candidates that are chosen. There is a substantial disagreement as to how semantic information (contextual information as well as meaning attributes of a word) influences the construction of meaning. Table 1.2 below presents the two possibilities of the effect of semantic information.

It should be clear that within traditional models, limited semantic information is available during the early stages of word identification because anticipating the next word from a word context was viewed to be much too slow to affect early, fast stages of word identification such as lexical access. Stored meaning information was considered to be only provided in later stages (such as semantic integration) when we, for example, need to have the meaning in order to begin to integrate it into what we have previously read. Thus, contextual top-down effects were only predicted on post-lexical stages of semantic integration, whereas lexical access was assumed to be the outcome of bottom-up perceptual processes (e.g., Fischler & Bloom, 1979; Forster, 1981; Schuberth & Eimas, 1977).

Table 1.2 Summary of the Possible Views as to how Semantic Information Influences the Construction of Meaning Influences the Construction of Meaning

Possible views	**Explanation**
Autonomous//modular (bottom-up driven)	Bottom-up perceptual information is most important; semantic information is only used for checking and enriching it.
Interactive (both bottom-up and top-down)	Semantic information interacts with perceptual information at all stages of processing (both sources of evidence, top-down and bottom-up, are available throughout).

According to the more interactive positions, it was argued that the meaning of the word (stored semantic information as well as background knowledge stored via schemas) is available at the outset, at the moment of lexical access, to help in the task of word identification, for example to speed up the process of lexical access. Thus, lexical processing was seen to rely on both lower-level perceptual and higher level information (e.g., Stanovich & West, 1979, 1981, 1983; West & Stanovich, 1978) such that, for example, contextual top-down effects could be expected to occur jointly with word frequency bottom-up effects.

1.7 The effect of parafoveal preview on lexical access

In this section, we consider the role of parafoveal preview information during the processing of low and high frequency words as well as predictable and unpredictable words. *Parafoveal preview benefit* (or simply preview benefit) is the processing advantage obtained when a reader views a useful preview of an upcoming (target) word compared to when that preview was denied in some way. That is, orthographic, phonological or semantic similarity information is not viewed parafoveally hence useful information about the upcoming word has been denied. Typically, researchers deny the normal preview (i.e., obtained during the course of normal reading) of the upcoming word by using a string of x's or scrambled letters. This method is achieved through a special adaptation of the eye-movement recording technique and is called the boundary technique (first used by Rayner, 1975). As in a normal eye movement experiment, a number of dependent variables are measured. The parafoveal preview benefit is calculated by subtracting the fixation time in the condition where the preview was either identical or related to the target word from the condition where the preview was unrelated to the target. It comes as no surprise that the preview benefit is greatest when the preview and target are identical. However, studies have also shown that preview benefit is obtained from other kinds of preview (i.e., not identical to the target) which shows that partial (incomplete) information is obtained from the parafovea (Rayner, 1998; 2009; Rayner & Pollatsek, 1989). In other words, information obtained parafoveally on fixation n is integrated in some way with foveal information processed on fixation n + 1 (i.e., word information is partially processed on one fixation and is completed on the next). Before looking at what kinds of information this is likely to be, it is useful to offer some background information related to parafoveal processing of words.

1.7.1 The foveal and parafoveal regions in text

In the literature on the visual field and acuity limitations, it is common place to divide the visual field into three regions in terms of our point of fixation, which is the foveal region. The other regions are the parafoveal and peripheral areas. These definitions reflect the limitations in the physiology of the eyes, for example, these areas reflect our need to make saccades so frequently, that is, the further away we get from the point of fixation (the foveal region) the more acuity decreases (in the parafoveal area and even more so in the peripheral area).

In the visual field, for normal sized texts, 1° typically corresponds to 3-4 letter spaces (or characters). However, it is widely agreed that the most appropriate measure of how far readers move their eyes is letter spaces and not visual angle (e.g., Morrison & Rayner, 1981; O’Regan, Lévy-Schoen, & Jacobs, 1983). The foveal region corresponds to 3-4 letter spaces to either side of the fixation (so about 8 letter spaces). The parafoveal region is less defined and is about 15 to 20 letter spaces on either side of fixation (Rayner, Juhasz & Pollatsek, 2007). The peripheral region comprises those regions out with the foveal and parafoveal regions. Visual acuity is at its poorest here though readers can be aware of gross aspects of text such as the end of a line. Normally, information in peripheral vision is of no use to the reader compared to that in the foveal region.

In eye movement studies, it has become the norm to refer to the fixated word as the foveal (or foveated) word and the word adjacent to it as the parafoveal word. Mostly, this maps onto the physiological constraints of the visual system as highlighted above. However, in some cases, a very short foveated word and a very short parafoveal word would both fall into the foveal region. Nonetheless, the non-fixated word is still called the parafoveal word. Similarly, if the foveated word is very long, the letters at the end of that word would fall into parafoveal vision (because of the constraints of the visual system). As before, the terminology still applies – the fixated word is the foveal word.

1.7.2 Techniques used to study the use of parafoveal information in reading

The eye movement method has several adaptations, called *eye-contingent display techniques* (each of which has several variants). These are the boundary technique (Rayner, 1975); the *moving window technique* (McConkie & Rayner, 1975) and the *foveal mask technique* (Rayner & Bertera, 1979). The first two in particular have been used to research various questions about parafoveal processing so we will focus on these. All three techniques were originally used to answer questions about the perceptual span during reading, that is, how much information a reader is able to process in one fixation (more recently, Miellet, O'Donnell, & Sereno, 2009). In relation to parafoveal processing, the important point to note is that in all three techniques, changes in the text are made in order to disrupt normal reading, however the disruptions are not because of the changes in themselves but rather because certain information is withheld from the reader (which would have been available in the case of normal reading). This means with the use of the boundary technique, for example, the amount and type of information available to the reader can be changed in a particular target word, and with the use the use of the moving window technique, the total amount of information available to the reader form a particular region of the text can be changed. These changes can tell us what information about say the target word has been perceived and used by the language processing system prior to fixation. That is, the boundary technique can be used to investigate questions about how much readers benefit from a preview of words to the right of the fixated word (i.e., the parafoveal word), that is, do they get the orthography, phonology, morphology, semantics (meaning) of that parafoveal word?

In addition to the gaze-contingent display change paradigms outlined above, parafoveal preview benefit can also be indexed dependent on the distance of the fixation prior to fixating the target word and the beginning of the target word (launch site). For example, a recent study has included a manipulation of frequency by predictability by launch site (Hand, Miellet, O'Donnell, & Sereno, 2010).

1.7.2.1 The boundary technique

In the boundary technique, only one word in the text is changed (rather than a display change with every fixation as in the moving window technique). Typically, a target word will appear (embedded in text) in some disrupted form (known as an 'invalid' preview of the target). When the readers' eyes make a saccade over a pre-specified boundary, for example at the end of the pre-target word, the disrupted word changes to the actual target word. Since this display change occurs during a saccade i.e., when vision is suppressed, readers do not notice the display change. Thus, the reader fixates the target in the usual way but, crucially, has been inhibited from parafoveally processing that word (exactly what information was inhibited from being parafoveally processed depends on the nature of the invalid preview). Hence, this method allows the experimenter to determine which aspects of the target word have been perceived and used by the language processing system before actually being fixated. An example passage is shown in Figure 1.6 below.

Figure 1.6. Example of a passage of text presented using the boundary technique

He fondly remembered his| gessal who had been a great companion at sea.
*

He fondly remembered his| gessal who had been a great companion at sea.
*

He fondly remembered his| parrot who had been a great companion at sea.
*

Note: The * represents the location of the reader's fixation. The | represents the pre-specified invisible boundary

1.7.2.2 The moving window technique

In studies using the moving window technique, a pre-specified number of letters to the left and right of the fixation are presented normally; this is the 'window' of text. Outside this window, some other stimuli are presented. Hence the size of the readers viewing window and the stimuli outside the window (i.e., the parafoveal preview) are both determined by the experimenter. The window of text moves with the readers' fixation such that at the readers' fixation point there is always the normal text. Figure 1.7. below shows an example passage presented using the moving window technique. The manipulation of window size allows researchers to establish the minimum window size necessary for reading to ensue at a normal rate.

Figure 1.7. Example of a passage of text presented using the moving window technique

```
xe fondly remembered his pxxxxx xxx xxx xxxx x xxxxx xxxxxxxxx xx xxx.
    *
xx xxxxxx xxxxxxxxxx xxx parrot who had been a gxxxx xxxxxxxxx xx xxx.
                             *
xx xxxxxx xxxxxxxxxx xxx xxxxxx xxx had been a great compaxxxx xx xxx.
                                         *
```

Note: The * represents the location of the reader's fixation

1.7.3 The features of the text acquired from the parafoveal word

As mentioned previously, many studies have shown that preview benefit is obtained from various kinds of invalid preview of the target word. This result suggests that partial word information is acquired parafoveally. The implication is that (useful) partial or incomplete word information is integrated in some way with the foveal information from the next fixation. That is, word information is processed partially on one fixation but completed on the next subsequent fixation (the integration of information across consecutive saccades (Rayner, 1998, 2011; Starr & Rayner, 2001). How parafoveal and subsequent foveal information are integrated can tell us which features of the text are important in lexical access (as well as understanding the full extent of normal reading). That is, the question is which features of the text are being extracted from words in parafoveal vision? Is the information being extracted from the parafoveal word visual features, sound codes abstract letters, or lexical and semantic?

Many studies have investigated what information about the word in parafoveal vision is extracted whilst fixating the foveal word (e.g., McConkie & Zola, 1979; Rayner, McConkie, & Zola, 1980; Rayner et al., 1982; Pollatsek, Lesch, Morris, & Rayner, 1992). This research has demonstrated evidence that readers obtain low-level information about the parafoveal word in the form of abstract letter codes (McConkie & Zola, 1979; Rayner, McConkie, & Zola, 1980). In addition, it seems that readers obtain some sub-lexical information about the parafoveal word (orthographic information such as partial word information i.e., from the first three letters of the parafoveal word) (Inhoff, 1989; Rayner et al., 1982), as well as phonological (i.e., the sound of a word) information (Henderson, Dixon, Peterson, Twilley, & Ferreira, 1995; Pollatsek, Lesch, Morris, & Rayner, 1992).

Is the code semantic information? A number of studies have shown that readers do *not* gain semantic information from the parafoveal word during either first fixations or gaze durations (in alphabetic writing systems; Altarriba, Kambe, Pollatsek, & Rayner, 2001; Hyönä & Häikiö, 2005; Rayner, Balota, & Pollatsek, 1986; Rayner, McConkie, & Zola, 1980 but cf. Hohenstein, Laubrock, & Kliegl. 2010). The theoretical implication is that a late-processing account of word recognition is supported with orthographic and phonological processing taking place first i.e., being made available to the reader first with semantic information being available to the reader at a later stage (e.g., Balota, 1990; Balota et al., 1991; Forster, 1976; 1979).

Therefore, the studies mentioned above suggest that readers can obtain information about the parafoveal word in terms of its abstract letter codes as well as its sub-lexical and phonological codes (all these codes are relatively low-level) whilst there is no evidence for higher level semantic information (cf. Hohenstein, Laubrock & Kliegl, 2010).

Another way to research the level of parafoveal processing which takes place during reading is to examine whether or not lexical and semantic parafoveal processing is possible. Lexical parafoveal processing would be suggested when readers are able to use information from a high frequency versus a low frequency target presented in parafovea, whilst in the foveated region. Similarly, semantic parafoveal processing would be suggested when readers are able to use information from a high contextual predictability target compared to a low one presented in the parafovea, whilst in the foveated region.

1.7.3.1 Parafoveal lexical processing: word frequency effects

Before looking at parafoveal lexical processing we look briefly at a study which examined the effect of word-initial letters on parafoveal processing (Lima and Inhoff, 1985; also Hand et al., 2012 in a normal eye movement reading study). As mentioned previously studies have shown that readers acquire some sub-lexical information from the parafoveal word in the form of orthographic information such as partial word information such as from the first three letters of the parafoveal word (Inhoff, 1989; Rayner et al., 1982).

Lima and Inhoff (1985) examined the role of constraint (the effect of word-initial information) on foveal and parafoveal processing. In terms of the former, they investigated whether the constraint of a word-initial trigram affected reading behaviour. They presented target words in neutral single-line sentences which differed in the amount of constraint as given by the first three letters of the target word. That is, the word ‘dwarf’ was a high constraint word because the trigram ‘dwa’ does not produce many words in its candidate set (e.g., ‘dwarves’, ‘dwam’), however the word ‘clown’ is a low constraint word because the first three letters ‘clo’ produces lots of potential candidates (e.g., ‘close’, ‘clock’, ‘cloud’, ‘cloak’, ‘clothes’, ‘closet’). Lima and Inhoff hypothesised that if the constraint posed by the word-initial trigram affects lexical access then there would be shorter fixations on the high constraint words than on the low constraint words i.e., the recognition point in ‘dwarf’ is earlier than the recognition point in ‘clown’ because there are less candidates in the cohort set for ‘dwarf’. However, the results did not support this; there was a significant effect of constraint but it was in opposite direction to that hypothesised. Specifically, when words were low constraint, participants fixated on them for less time than when words were high constraint. Lima and Inhoff suggested that this was because the trigram of the low constraint words was more familiar to participants than the trigram of high constraint words. The possibility suggested by Lima and Inhoff is that when trigram familiarity is high, this confers an advantage to lexical access by way of decreasing the time to foveally process it.

Lima and Inhoff (1985) also examined the impact of constraint (of the word-initial trigram) on parafoveal processing. The reasoning behind this manipulation was that studies have shown that when readers’ parafoveally see the first three letters of the parafoveal word, this leads to a large parafoveal preview benefit (Inhoff, 1989; Rayner et al., 1982). Therefore, Lima and Inhoff (1985) hypothesised that participants may utilise the constraint from the trigram letters of the parafoveal word such that lexical access is influenced for that word when it is subsequently fixated. Specifically, lexical constraint having an effect on parafoveal processing would be suggested if the preview benefit for high constraint words such as ‘dwarf’ was more than that for ‘clown’. Lima and Inhoff (1985) used the moving window paradigm in their study; there was the one-word condition, the two-word condition where a string of x’s replaced the text outside the specified window, and a full-line condition (i.e., normal reading). In the one-word condition, parafoveal preview of the upcoming word was inhibited and in both the two-word and full-line conditions, the preview of the subsequent word was available. They had expected that results would show a greater preview benefit for high constraint word (i.e., 'dwarf') than for low constraint words (i.e., 'clown'), however the interaction between preview benefit and target constraint was nonsignificant, in addition there was the same preview benefit in the two conditions low and high constraint. This result suggests that lexical constraint does not impact on parafoveal information. However since participants fixated less time on low constraint words than on high constraint words, Lima and Inhoff suggested the familiarity of the trigram affects how quickly it is processed foveally. In addition, the interaction between constraint and window size was nonsignificant, suggesting that constraint as an effect of familiarity does not occur parafoveally (only foveally). This result suggests that constraint or familiarity does not affect parafoveal processing. Such a result supports serial models of word recognition. However, it is possible that the amount of constraint imposed by the word-initial trigrams was not enough to influence parafoveal processing particularly word-initial trigrams have small effects even in foveal processing. In addition, a subsequent later study (Inhoff & Rayner, 1986), in contrast to Lima & Inhoff (1985), showed that word frequency (which correlates with word familiarity) can affect parafoveal processing thus supporting interactive models of word recognition (e.g., Coltheart et al., 2001).

Inhoff and Rayner's (1986) examined whether word frequency of the parafoveal word affected parafoveal preview benefit. The frequency of the target word (embedded in a single line sentence and matched on word length and contextual predictability) was manipulated along with the viewing conditions (the moving window technique was used) where how much of the sentence to the right of fixation was manipulated (information to the left was available at all times). That is, low and high frequency target words were read in one of three viewing conditions, one-word window (only the fixated word was seen); two-word window (the fixated target and the word next to it was seen).

Inhoff and Rayner (1986) measured the first fixation duration and gaze duration on the low and high frequency target word in the three preview conditions, one-word, two-word and full-line. In the first fixation measure (i.e., reflecting lexical access processes) and in the gaze duration measure, there was a main effect of frequency where there were significantly shorter fixations on HF versus LF words; a main effect of window size (fixations in the one-word presentation were longer than those in both two-word and full-line). In the first fixation measure (but not in the gaze duration measure), there was also a significant interaction between word frequency and parafoveal preview which was that a parafoveal preview of HF words lead to larger preview benefits than a parafoveal preview of LF words (this was in all three parafoveal preview conditions). This finding that readers acquire more effective parafoveal previews from high frequency words than from low frequency words suggests that more information is extracted from a high frequency parafoveal word than a low frequency parafoveal word. These results suggest that frequency can modulate the use of parafoveal information. Moreover, this is in agreement with the interactive threshold model (Balota and Rayner, 1990). This is because within this model, it is envisaged that high frequency words have lower thresholds than low frequency words (e.g., Morton, 1969) which means that the interactive threshold for high frequency preview words is more easily exceeded than that for low frequency words. Therefore, lexical access takes more time (evidenced by longer fixation durations) the more ineffective the parafoveal information i.e., from lower frequency words.

Results also showed that when parafoveal previews were not available (i.e., in the one-word window), then there was frequency effect only in the gaze duration measure. However, when parafoveal preview was available (i.e., the two-word and full-line conditions) then there was a frequency effect in both first fixation and gaze duration measures. This result suggests that at least some of the effect of parafoveal preview is on early lexical processes as reflected in the first fixation duration measure.

1.7.3.2 Parafoveal semantic processing: contextual predictability effects

The relationship between contextual predictability and parafoveal processing was examined by Balota et al. (1985; also in earlier work by Balota & Rayner, 1983; McClelland & O'Regan, 1981). In Balota et al.'s (1985) study, sentence frames were constructed to accommodate either a predictable target word or an unpredictable target word, given the prior sentence context. An example sentence frame from this study is shown in Figure 1.8. In this example, sentence the target word 'cake' was the predictable target word given the past context and the target word was 'pies' given the preceding context. In addition, the availability of parafoveal information was also manipulated using the boundary technique (Rayner, 1975). There were five parafoveal preview conditions: in the example presented in Figure 1.8., for the predictable target word 'cake' the preview was either identical to the target (i.e., 'cake'), visually similar but a nonword to the target word (i.e., 'cahc'), semantically related and visually dissimilar (i.e., 'pies'), visually dissimilar nonwords (i.e., 'picz') and finally, semantically anomalous and dissimilar (i.e., 'bomb').

Figure 1.8. Example sentence from Balota, Pollatsek and Rayner (1985)

Since the wedding was today, the baker
rushed the wedding _____ to the reception.

There were two results which showed that contextual predictability affected the use of parafoveal information. The first was to do with skipping the target word. In particular, the high predictable target word was skipped significantly more than the low predictable word, as had been shown in an earlier study (Ehrlich & Rayner, 1981) as well as in subsequent studies (e.g., Rayner & Well, 1996). This finding indicates that contextual predictability affects the use of parafoveal information. In addition, when the target word was not skipped, Balota et al., (1985) examined the gaze duration on these target words. Results showed that gaze durations were significantly shorter when the target word was high predictable compared to low predictable, again this finding had also been previously shown (Ehrlich & Rayner, 1981). Another result was that the difference in gaze durations between the identical preview condition (i.e., 'cake' for 'cake') and the visually similar preview condition (i.e., 'cahc' for 'cake'). Specifically, there was a significant difference in gaze duration time on the target word between the two conditions (identical and visually similar) for the high predictable target words but not for the low predictable target words. This result indicates that more preview information was extracted when a target word was high predictable compared to when it was low predictable. For the present discussion, this is an important finding because it suggests extraction of parafoveal information is enhanced when it is by guided higher level sentence context.

The implication from both the above reported findings in Balota et al.'s (1985) experiment do not agree with predictions of a modular account of lexical access (e.g., Forster, 1976; 1979) but rather are in line with an interactive position of the language processing system (e.g., McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1985). In fact, other work by McClelland (McClelland & O'Regan, 1981) specifically addressed the relationship between contextual predictability and the use of parafoveal information.

According to McClelland & O'Regan's (1981) model because information flow is multidirectional among the component processors, lexical access is suggested to be influenced by contextual information as well as parafoveal information. Specifically, contextual information is used in conjunction with parafoveal information which allows for enough activation to accumulate leading to a logogen exceeding its threshold. Therefore, when a sentence context is highly predictable towards a given target word and when there exists useful parafoveal information, these information sources are combined so that enough activation accumulates so that a logogen exceeds its threshold. On the other hand, when a sentence context is low predictable towards a given target word, a greater number of logogens are activated which means that it takes more time for the right logogen to exceed its threshold. Hence this model explains the difference in time spent on the target word when it is read in a highly predictable context versus a low predictable context.

A subsequent model, the interactive threshold model, was put forward by Balota and Rayner (1991; also Rayner & Balota, 1989 and Balota & Rayner, 1983) and takes its starting point from McClelland and O'Regan's (1981) model in offering an account of the underlying mechanisms which could possibly explain the contextual predictability and parafoveal information superadditive interaction. Balota and Rayner (1991) suggested that the superadditive interaction between contextual predictability and parafoveal information arises because lexical-level representations obtain two joint sources of information: top-down sources via the sentence context and bottom-up sources of parafoveal information. In their framework, Balota and Rayner (1991) used the notions of 'interactive thresholds' and 'interlexical inhibitory processes'. The role of interactive thresholds is that when activation gets to a particular level at a logogen (Morton, 1969), there is then a different effect on performance. For example, it is assumed that there is a level of activation which is enough to engage attentional resources. With interlexical inhibitory processes, the idea is that when a particular lexical representation starts to accumulate activation, it will also inhibit activation of related lexical representations. For example, a participant may pick up 'ca' from the parafovea which can then partially activate lots of different lexical candidates. These partially activated candidates also mutually inhibit each other so that each particular candidate does not receive enough facilitation to be considered as being consistent with 'ca'. However, depending on the text, there can be instances where either contextual constraint or parafoveal information can be enough to permit one lexical candidate to supersede over others. For example, if the context pertains to 'cat', then the lexical representation for 'cat' is likely to stand out compared to the relatively lower levels of interlexical inhibition produced by the other possible candidates. In terms of parafoveal activation, if 'cak' is viewed parafoveally by the reader, then this might be enough for the lexical representation of 'cake' to stand out from the other neighbours and thus produce facilitation.

Balota and Rayner (1991) interactive activation model accounts for the superadditive effects of context and parafoveal information. For example, there may be cases where the context (e.g., for the target ‘cake’ if the context was about ‘pastry’) may not be enough to exceed the interactive threshold and affect performance. The contextual constraint posed by ‘pastry’ will of course lead to some activation, that is, a number of logogens would be activated but there is not enough activation for one candidate to dominate over other possible candidates. Similarly, in terms of parafoveal activation, in some cases, the parafoveal activation of reading ‘ca’ may not be enough to exceed the interactive threshold. Thus, the parafoveal information leads to some activation for many visually consistent logogens, again, these mutually inhibit each other. However, when both sources of information – contextual constraint and parafoveal information – are available, then there is enough activation for a single logogen to exceed the interactive threshold. That is, if the reader has the availability of ‘pastry’ from the context and ‘ca’ from the parafovea, then this should lead to enough activation for the logogen corresponding to ‘cake’ to exceed the interactive activation threshold and thus stand out from the inhibition produced by partially activated representations.

In conclusion, Balota et al.’s (1985) study suggests that parafoveal semantic activation is possible. This is because their results showed that higher level contextual predictability (via top-down sources operating from the sentence context) can modulate the use of parafoveal information. Specifically, parafoveal information was used more efficiently when the predictability of the target word was high (that is, the sentence context was constrained towards the target word) than when it was low (thus, when the sentence context was not constrained towards the target word). However the difference in time spent on the target word was in the gaze duration measure. Since the gaze duration measure is likely to reflect processes which occur between lexical access and post-lexical access stages, the interaction between contextual predictability and parafoveal information needs to be reported in those initial measures which are likely to reflect lexical access processes i.e. the first and single fixation duration measures.

1.7.4 Other approaches to tracking the time course of semantic information extraction

The studies reviewed so far which have examined parafoveal processing have done so with the use of either a moving window or a boundary technique. The first paradigm allows for researchers to manipulate how much of a preview is given to readers on a given fixation whilst the latter allows researchers to implement a condition of a valid and invalid preview of a target, with the assumption that invalid parafoveal previews has prevented the reader from processing the target parafoveally.

However, aside from moving window and boundary techniques, there are some other approaches to examining parafoveal preview benefit (along with frequency and/or contextual predictability). Two such studies were conducted by Hand et al. (2010) in which they used launch distance (the distance between the beginning of the target word and the location of the previous fixation) to index the extent of parafoveal processing. In some earlier work, Sereno and Rayner (1992) devised a new technique called the fast-priming paradigm to track the time course of semantic information extraction form the foveal word.

1.7.4.1 Launch distance to the target

Hand et al. (2010) conducted an eye movement study which examined the joint effect of word frequency, contextual predictability and preview information. Parafoveal preview benefit was indexed by launch distance rather than conditions of valid and invalid preview (through use of the boundary technique). Launch distance was measured as the distance between the beginning of the target word (that is, the space before the target) to the location of the previous fixation. There were three levels of launch distance (near: 1-3 characters; middle: 4-6 characters; and far: 7-9 characters). They suggested that the closer the prior fixation to the target word, the more preview of the target word the reader would have prior to fixating it. Thus, a lesser distance, ‘near’ (a distance of 1-3 characters from the target word) gives more of a preview than ‘middle’ (a distance of 4-6 characters from the target word) and both give more preview than ‘far’ (a distance of 7-9 characters from the target word).

Hand et al. (2010) hypothesised that results would show an effect of launch distance which would be longer target fixations the greater the launch distance. In first, single, and gaze duration measures, similar results were obtained in which in all these measures, there was a main effects of frequency (longer fixations on LF than on HF words), of predictability (longer fixations on LP than on HP words) as well as of preview. Follow-up tests to the latter effect showed that the closer the launch distance, the shorter the fixation time on the target word; all contrast were significant: near vs. middle; near vs. far; and middle v. far. This suggests that shorter launch distances gives way to greater parafoveal previews and thus there are shorter fixation times on the target.

In terms of the interaction, the two-way frequency by predictability was not significant; however the other two-way interactions were significant: frequency by preview as well as predictability by preview. Furthermore, the three-way interaction was also significant. To follow up this latter effect, Hand et al. (2010) conducted separate frequency by predictability ANOVAs at each level of preview (i.e., near, middle and far conditions). The results showed an interactive pattern in the 'near' and 'middle' conditions and an additive pattern in the 'far' condition. Specifically, in the 'near' condition, there were main effects of frequency and predictability and a significant interaction in which there was a larger predictability effect for LF words than for HF words. In the 'middle' condition, only the main effect of frequency was significant; the interaction was significant but there was an opposite pattern to that of the significant interaction in the 'near' condition: there was a larger predictability effect for HF words than LF words. In the 'far' condition, the only significant effect was that of the main effect of frequency.

Hand et al. argued that their initial additive results of frequency and predictability were the result of a combination of these results: the interaction in frequency by predictability and launch distance in the 'near' and 'middle' launch distances and an additive pattern of results in the 'far' condition. They suggested that the overall pattern of results showed that the greater the launch distance, the less the effect of parafoveal preview. In other words, Hand et al. (2010) showed additive effects of frequency and predictability, but there is the indication that these factors can exert interactive effects – but the frequency-predictability interaction is dependent on parafoveal preview.

1.7.4.2 The fast-priming paradigm

Sereno and Rayner (1992) developed the fast-priming paradigm. This method tracks the extraction of semantic information from the *foveal* word. In this paradigm, when the eyes are to the left of an invisible boundary (the last letter but one of the word before the target) a preview of random letters (e.g., 'gzsd') occupies the space of the target in order to prevent parafoveal processing of the given target. During the saccade that crosses the invisible boundary, the random letter string was replaced (for a particular length of time) with one of three possible prime words: semantically related to the target (e.g., 'love') or semantically unrelated to the target (e.g., 'rule') or finally an identical condition where the actual target was shown (e.g., 'hate'). The target ('hate) then replaces the prime and stays there until the participant had finished reading the sentence. Prime durations of 30, 45 and 60 ms showed that there was an effect of prime type at the 30 ms duration. That is, when the prime was presented for 30 ms, gaze durations were significantly faster for targets preceded by the semantically related primes compared to targets preceded by the unrelated primes (there was a 28 ms difference in gaze duration between the two prime conditions). In the second experiment, Sereno and Rayner aimed to investigate the 30 ms prime presentation by using prime duration just above and below 30 ms: the prime durations were either 21, 30, or 39 ms. Each target noun has three corresponding primes (as in Experiment 1): semantically related and semantically unrelated and instead of an identical prime (i.e., the word itself), they used a 'random letter string' condition (e.g., 'frxe'). Again, they showed a priming effect at the 30 ms duration level. Semantic priming effects at similar durations (35 ms) were shown in later work (Sereno, 1995).

The results from Sereno and Rayner (1992) and Sereno (1995) studies suggest that the extraction of semantic information occurs within a relatively narrow time frame of between 30 to 35 ms of fixating the foveal word. Therefore, these studies are important in that they shed light on the time course of foveal semantic processing. However, little is known about the time course of parafoveal semantic processing and one aim of the thesis was to investigate this.

1.8 Outline of experiments

The experiments in this thesis explore the time course of the effects of semantic variables, that is, of contextual predictability and word imageability. This is by using the logic of additive factors method (Donders, 1868/1969; Sternberg, 1969a; 1969b). Word frequency was taken to index the moment of lexical access and the two semantic variables are each orthogonally manipulated to determine whether the combined result in the ANOVA is additive or interactive. An additive result indicates that the two variables are occurring at different stages of processing indicating that bottom-up information is most important and semantic information is only used for checking and enriching the bottom-up information. On the other hand, an interactive result would indicate that the two variables share the same processing stage, in particular that they have a joint influence during lexical access. This would indicate that semantic information interacts with perceptual information at all stages, bottom-up and top-down, of processing and thus suggests the role of both bottom-up and top-down processing to achieve lexical access.

Experiment 1 (reported in Chapter 2) examines the focus to the time course of word imageability which encompasses at least some part of the semantic aspects of the word itself (Whaley, 1978). This first Experiment, using the lexical decision task, was considered a starting point for investigating the relationship between word frequency and word imageability on reaction times and percent error rates. Specifically the interest was in whether these factors demonstrate an additive or interactive pattern of results on the two dependent variables.

Experiment 2 (Chapter 3) and Experiment 3 (Chapter 4) explored the nature of the relationship between word frequency and contextual predictability, specifically whether the two variables are additive or interactive on a number of fixation duration measures as well as on the likelihood of skipping the target word. Chapter 4 also includes a cross experiment comparison of Experiments 2 and 3.

In Experiment 2, participants' eye movements were recorded while they read a two-line passage of text. Word frequency and predictability were manipulated (of a single target word, placed in line 2 of the passage of text) and the effects of this manipulation on a number of 'on-line' eye movement measures of reading, such as the first and single fixation durations of target word, were examined. These experiments were the first to use a very high level of contextual predictability than has been the case in past studies which have examined the joint effect of word frequency and contextual predictability.

Experiment 3 additionally used an invalid parafoveal preview of the target word. The boundary technique (Rayner, 1975) was used to manipulate parafoveal preview of the target word. This was to investigate the contribution of parafoveal processing to the interaction between word frequency and contextual predictability. Lexical access starts from parafoveally viewing the target word (Rayner, 1998), and past research has shown that more information is obtained parafoveally from a high versus low frequency word (Inhoff & Rayner, 1986) and from a high versus low predictable word (Balota et al., 1985). This means that it is possible that the processing advantage for high versus low frequency and for high versus low predictability target words is due in part to the parafoveal processing aspect of the respective words. A secondary aim was that in using the boundary technique, fixation durations would be slowed. That is, fixation duration measures observed in Experiment 2 were faster than in past eye movement frequency-predictability studies (e.g., Hand et al., 2010; Rayner et al., 2004). It was argued that faster reading times were most likely due to the clearer display used in the current experiment. It is difficult to make a font 'less clear' whilst preserving the actual text. Therefore another way to slow down reading times is to deny the parafoveal access of targets. From past studies it is known that when parafoveal information is withheld, reading rate can be slowed by as much as 40% (Rayner et al., 1982). By using an invalid parafoveal preview of the target word, we ensured foveal-only processing of the target word where no useful lexical information was available to the reader prior to actually fixating the target word in foveal vision.

Chapter 4 also reported additional data analyses using the data obtained in Experiments 2 and 3. The aim here was to examine whether word frequency and contextual predictability of the parafoveal word affected parafoveal preview benefit. Preview benefit was calculated by subtracting fixation durations in a condition of ‘valid’ preview of the target with an ‘invalid’ preview of the target. To do this, we conducted a between-groups design comparison in which we compared data from Experiment 2 where participants’ parafoveally viewed a normal i.e., valid preview from the target word with data from Experiment 3 where participants’ were inhibited from parafoveally processing the target word by viewing a nonword letter string in place of the target. The nonword letter string was better than the stimulus used in the invalid preview condition in past studies. That is, as well as maintaining word length between target and nonword, overall word shape was also maintained where ascenders (e.g., b, d, f), descenders (g, j, y) and all small letters (e.g., a, c, n, m) replaced ascenders, descenders and small letters respectively. Also the nonword letter string was orthographically legal and pronounceable given the rules of the English language. All of these procedures were in order to insure that the nonword letter string did not stand out as being ‘odd’ when occupying the target region in a given passage of text whereby it could command the reader’s attention.

Experiment 4 (reported in Chapter 5) investigated the single and joint effects of word frequency, contextual predictability and parafoveal preview. All three factors were manipulated in within-subjects design as this had not been done in the prior cross experiment comparison of Experiments 2 and 3.

Chapter 2

Frequency and imageability effects in lexical decision

2.1 Introduction

Autonomous models of how readers locate words in the mental lexicon (e.g., Forster, 1979; Morton, 1969) assume that a given word must be recognised before we can access the meaning of the word. Such models assume that word-level (lexical) processing precedes processing at higher levels of analysis. However, if it can be shown that there are interactive effects of frequency and imageability then this would suggest a role for semantic (or meaning) level information at the same time as early lexical processing (e.g., McClelland & Rumelhart, 1986). Such semantic effects on early lexical access are controversial and the aim of the current study was to examine if there are imageability effects in early lexical processing.

Word imageability refers to the ease with which a word gives rise to a sensory mental image. This is different from word concreteness which has been defined as the ability to see, hear, and touch something (Paivio, Yuille, & Madigan, 1968). Typically, imageabilty scores are obtained from a number of established norms which have been collected previously by researchers. In such cases, researchers have presented participants with lists of words and participants have been asked to rate each word on a seven-point scale ranging from low-imagery to high-imagery (e.g., Clark & Paivio, 2004; Pavio, Yuille, & Madigan, 1968). These and several other norms (e.g., Cortese & Fugett, 2004; Gilhooly & Logie, 1980) can be used together to obtain a mean imageabilty rating for a given word. Similarly, a word's concreteness[1] is determined via 7-point rating tasks, with 1 denoting least concrete and 7 most concrete (e.g., Spreen & Schulz, 1966). Words referring to objects are instructed to be given a high concreteness rating, and words referring to abstract concepts (which cannot be experienced by the senses) are instructed to be given a low concreteness rating. The 'concreteness effect' refers to concrete words (e.g.,

[1] Word imageabilty and word concreteness are highly correlated with each other. For example, *r*=0.83 was reported in Paivio, Yuille, and Madigan (1968). In most of the early literature, the two terms were often confounded. For example, studies which investigated concreteness effects in actual fact made a distinction between low and high imageability words to represent abstract and concrete words respectively. In the present study we made a distinction between concreteness and imageability, however previous findings researching either variable are relevant since both can be seen to represent a semantic (meaning) level variable.

aeroplane) being more easily recognised (processed more quickly and accurately) than abstract words (e.g., truth).

Typically, the concreteness effect has been investigated in isolated word recognition tasks such as lexical decision (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Bleasdale, 1987; Chiarello, Senehi, & Nuding, 1987; de Groot, 1989; Howell & Bryden, 1987; James, 1975; Kroll & Merves, 1986; Ransdell & Fischler, 1987; Rubin, 1980; Schwanenflugel & Shoben, 1983; Schwanenflugel, Harnishfeger, & Stowe, 1988; Whaley, 1978); and naming (Balota et al., 2004; Bleasdale, 1987; de Groot, 1989). Other tasks which have been employed were word association (de Groot, 1989); recall (Paivio, 1971; also see Paivio, 1986; Ransdell & Fischler, 1987, 1989; Rubin, 1980; Wattenmaker & Shoben, 1987; Marschark & Surian, 1992; Nelson & Schreiber, 1992; Paivio, Walsh, & Bons, 1994; Schwanenflugel, Akin, & Luh, 1992); sentence verification (Holmes & Langford, 1976); and sentence comprehension (Haberlandt & Graesser, 1985; Schwanenflugel & Shoben, 1983; Schwanenflugel & Stowe, 1989; Wattenmaker & Shoben, 1987).

In the literature, there are two prominent explanations to explain the concreteness effect that is the concrete word advantage over abstract words: the dual coding theory (Paivio, 1986; 1991; 2007) and the context availability model (Schwanenflugel & Shoben, 1983; Schwanenflugel, 1991; also see Bransford & Johnson, 1972). Both proposals include the assumption that representations for concrete words are richer than those of abstract words. The dual-coding theory holds that there are two structurally and functionally separate yet interconnected representational systems. One such system is verbal which is dedicated to the representation and processing of linguistic stimuli. Another system is nonverbal and is specialised for imagistic stimuli. The concrete word advantage is explained by concrete words having access to information from multiple systems, in particular being coded by both systems, compared to abstract words which rely on verbal code representations only. According to the second explanation, the context availability model, comprehension of a given word is dependent on verbal context which can be either from the preceding discourse or by the readers stored schemas in long-term semantic memory. According to this proposal the concrete-word advantage is not because concrete words have the additional benefit of a distinct nonverbal system, but rather because concrete words have greater and denser links to contextual information in semantic memory than do abstract words Thus even with little provision of contextual information, the concreteness advantage over abstract words can be explained. Much research has been conducted in the context of these two theories, and more recently researchers have employed cognitive neuroscience techniques such as event-related potentials (e.g., Kounios & Holcomb, 1994) and fMRI (e.g., Binder et al., 2005). Results have however been inconclusive. In addition, some researchers have offered modifications of these accounts. For example, the dual-coding theory has been extended to better account for ERP findings (see Holcomb, Kounios, Anderson, & West, 1999).

In contrast to the studies mentioned above which have shown the concreteness effect, other studies have shown that concrete words were not processed more quickly and accurately than abstract words (Feldman, Basnight-Brown, & Pastizzo, 2006: lexical decision; Kroll & Merves, 1986: Experiment 3, lexical decision; Rubenstein, Garfield, & Millikan, 1970: lexical decision; Richardson, 1976: naming and lexical decision; Samson & Pillon, 2004: lexical decision; Schwanenflugel, Harnishfeger, & Stowe, 1988: Experiment 1, lexical decision; Schwanenflugel & Shoben, 1983: Experiment 1, sentence reading task, Experiment 2, lexical decision; Van Hell & De Groot, 1998b: Experiment 3, lexical decision, 2008: Experiment 1, lexical decision) and yet other experiments where abstract words showed a processing advantage over concrete words (Kousta, Vigliocco, Andrews, & Del Campo, 2011, lexical decision; Paivio & O'Neill, 1970, tachistoscope; Tokowicz & Kroll, 2007: Experiments 2 and 3, translation task and lexical decision respectively; Van Hell & De Groot, 1998, word association; Zhang, Guo, Ding, & Wang, 2006, in Chinese using event-related potentials).

Studies monitoring eye movements during the course of natural reading have been minimal. Only one experiment has shown the concreteness effect on fixation durations (Juhasz & Rayner, 2003). These researchers carried out a multiple regression analysis to determine the contributions of five lexical variables which have been implicated in word recognition performance (variables which influence the ease with which a word is recognised) and also are naturally correlated in the English language: word frequency, subjective familiarity, word length, concreteness, and age-of-acquisition on reading time of a target word while participants' eye movements were measured. In a traditional factorial design, it is difficult to separate out the influence of correlated variables form other variables which also influence word recognition. By analysing the data using a multiple regression, the analyses can inform as to whether each variable investigated influences eye fixation durations over and above the other variables included in the regression equation. In Juhasz and Rayner's study, the intercorrelations between the five variables were kept to a minimum. The materials used in the study were 72 nouns used as targets presented in single line sentences so that they were not the first or the last two words in the sentence. Sentences were designed so that they were neutral of the upcoming target. To do this, a separate group of ten participants were presented with the sentence frame up to but not including the target word. They were instructed to write in the next word which follows on from what they have read. The targets were predicted by participants just 1.5% of the time and thus the sentence frame was considered neutral. Results from this study showed that concreteness significantly predicted first fixation duration, gaze duration, and total time.

In studies which have used the event-related potential (ERP) technique, the typical finding is that concrete nouns elicit a larger N400 (a late negative-going wave which peaks at around 400 ms post stimulus onset and is largely centro-parietally distributed) than do abstract nouns (Ferlazzo, Conte, & Gentilomo, 1993; Holcomb et al., 1999; Kounios & Holcomb, 1994; Lee & Federmeier, 2008; Nittono, Suehiro, & Hori, 2002; Paller, Kutas, Shimamura, & Squire, 1987; Tolentino & Tokowicz, 2009; Tsai, Yu, Lee, Tzeng, Hung, & Wu, 2009; West & Holcomb, 2000; Zhang et al., 2006). Other studies have used functional neuroimaging (e.g., Binder et al., 2005; Fiebach & Friederici, 2003); PET (e.g., Whatmough, Verret, Fung, Cherktow, 2004), and TMS (e.g., Papagno, Fogliata, Catricala, & Miniussi, 2009) to assess regions of brain activation when concrete and abstract words are processed.

In the present study, we investigated whether there are imageabilty effects in early lexical processing by examining the combined effect of word frequency and imageability since word frequency indexes early lexical processing (e.g., Sereno & Rayner, 2003). One previous lexical decision study (West & Stanovich, 1982) orthogonally manipulated word frequency and a semantic (meaning level) variable, contextual predictability, and obtained an interactive result of the two variables. Contextual predictability can be viewed as an extra-word variable which represents the top-down process of comprehension (Rayner & Sereno, 1994) and imageability can be viewed as a semantic variable. By orthogonally varying the imageability variable with word frequency, it is possible to determine whether the combined effect is additive (e.g., Fodor, 1983; Forster, 1979) or interactive (e.g., McClelland, 1987; Morton, 1969).

Furthermore, the present study used words which were not correlated on imageability and age-of-acquisition (AoA) such that any observed effects would be due to the effect of imageability alone. AoA is a lexical variable which affects word recognition and has been highlighted in the literature as far back as the early 1970s (Carroll & White, 1973). The typical finding is that words which are learned earlier in life are responded to faster and more accurately than words which are learned later in life (Alario et al., 2004; Johnston & Barry, 2006; Juhasz, 2005). An AoA rating for a given word is typically obtained from pre-existing norms (e.g., Gilhooly & Logie, 1980; Morrison, Chappell, & Ellis, 1997). These norms have been collected by asking adults to rate words on a seven-point scale, with each scale representing an age band, at what age they think they learned a particular word.

There is a natural negative correlation between imageability and AoA (Barca, Burani & Arduino, 2002; Bird, Franklin & Howard, 2001; Gilhooly & Logie, 1980; Reilly, Chrysikou & Ramey, 2007). Thus, words which are learned early in life (early AoA) also tend to be of higher imageability (e.g., zebra) than words which are learned later in life (late AoA) which are also likely to have lower imageability (e.g., dogma). The ramification for not controlling for AoA in studies investigating the concreteness effect is that AoA could account for at least some of the observed concreteness effect. Specifically, high imageabilty words are likely to have been words which were also of earlier AoA than the low imageabilty words (which are likely to have been of later AoA). Thus, some of the reported effects of imageability on reaction time and errors and some of the reported interaction with frequency may be due to AoA, rather than to imageability.

In the literature investigating concreteness (and/or imageability) and frequency via use of the lexical decision task, findings are very mixed (de Groot, 1989; James, 1975; Kroll & Merves, 1986; Rubenstein, Garfield, & Millikan, 1970). All studies have shown a main effect of frequency where high frequency words were responded to faster than low frequency words. The main effect of imageability and the interaction between frequency and imageability, however, differs between studies. Also inconsistent between studies is participant accuracy, as indicated by which experimental condition most errors were incurred in. Furthermore, in some studies where effects have been significant, they are so either by subjects or by items, not both (e.g., Kroll & Merves, 1986; Experiment 4). In terms of RT results, one lexical decision study has shown main effects of frequency only; there was no main effect of concreteness nor was there a significant interaction between frequency and concreteness (Rubenstein, Garfield, & Millikan, 1970). In terms of participant accuracy, participants' discrimination of words from nonwords was high across all conditions (Rubenstein et al., 1970).

Rubenstein et al., (1970) investigated the effect of concreteness along with two other variables, word frequency and polysemy (homography; words with identical spelling but with more than one meaning vs. nonhomograph; words with one meaning). There were three levels of frequency (low, moderate and high). Whether a word was concrete or abstract was a decision made by the authors using the criteria that a word can be classified as concrete if its referent can be perceived by the senses. Nonwords followed the rules of orthography and phonology of the English alphabet. Critical and nonwords were all 4-5 letters long and there were a total of 185 critical words. 39 participants took part in total. Results showed a main effect of word frequency only (as well as polysemy) and there was no interaction between word frequency and concreteness. Interestingly, even though there were no significant differences in RT between concrete and abstract nonhomographs, there was an interaction between concreteness and polysemy where the RT for concrete homograph words was less than for abstract homograph words, specifically, this was for homographs with two meanings where one meaning was concrete and the other abstract compared to homographs with both meanings concrete. The authors had expected concrete words to have a shorter RT based on the idea that word meanings which are richer and more dense in sensory information are more easily retrieved from our store of words. The authors could not explain this in terms of frequency, familiarity, word length differences or frequency of meanings. As an explanation, they suggested that the significant RT difference could be due to homographs with both concrete and abstract meanings have a more associates compared to homographs with just concrete meanings. However, when a later experiment was conducted to test this, the opposite result was found. In terms of participant accuracy, Rubenstein et al. reported accuracy of discrimination between a word and a nonword which was the percentage of correct responses made in each condition; this was high for all conditions, ranging from 92 to 100% (in the high frequency homograph concrete condition). The lowest percentage of correct response of 92% was made in the low frequency nonhomograph abstract condition; high frequency abstract nonhomographs condition had 98% correct responses. In terms of concrete words, low frequency words had 95% correct responses, and high frequency 99%. There were no statistical tests performed on this data.

One particular flaw of this study is the way in which these researchers determined concreteness. That is, it could be argued that this was very subjectively done and a more valid method would have been to collect norms from a small number of participants to validate their own intuitions regarding the concreteness of a word. Also there was a failure to report the error rate in each condition. In addition, a later re-examination of this study found that the effect of homography which had been significant in the subjects analysis was in fact, not significant in an items analysis (Clark, 1973).

In contrast to the above finding reporting a lack of a main effect of concreteness, a small group of lexical decision studies have shown a main effect of concreteness (as well as frequency) and a significant interaction between the two, whereby there were faster lexical decision times to concrete words compared to abstract words for low frequency but not high frequency words (de Groot, 1989, Experiment 4; James, 1975, Experiments 1 and 2; Kroll & Merves, Experiments 2, 1986; Experiment 1, significant main effects only). However, in terms of participant accuracy, the data are inconsistent between studies (see below). More problematic is the failure of all these studies to match words on AoA; therefore it is possible that the significant main effect of imageability and the significant interaction between frequency and imageability may be partly due to AoA, rather than to the effect of imageability.

James (1975) reported a series of four lexical decision experiments in which concreteness and frequency were orthogonally manipulated. In the first experiment, there were three levels of frequency (low, moderate and high) taken from Francis and Kučera (1967) norms. Concreteness of words was determined rather subjectively. It is stated that two people (it is unclear whether one of these was the author) judged words to be either concrete or abstract. There is no other information reported including if any rating scales were used and if so, cut-off's for judging a word to be either abstract or concrete. Furthermore, the data reported for all four experiments includes only subjects analysis; items analysis were not conducted.

In the first experiment, James used nonwords which followed the rules of English orthography. The RT data showed main effects of concreteness and frequency as well as a significant interaction where abstract low and moderate frequency words had a longer reaction time than concrete low and moderate frequency words. However, there were no significant effects of concreteness for high frequency words. In terms of participant accuracy, James reported the percentage error per condition that participants made in making a lexical decision. James did not report whether there were significant main effects of frequency or of concreteness; however the interaction was significant and there was one significant difference where the highest error rate (13.5%) was incurred in the low frequency abstract condition vs. the error rate in the low frequency concrete condition

(4.2%). There were no significant differences in high frequency words; abstract words incurred an error rate of 1% and concrete words 1.3%.

In the second and remaining experiments, there were two levels of frequency, high and low. A somewhat similar pattern of results to the first experiment was reported in this second experiment in which the nonwords now consisted of homophones with the same pronunciation to real English words (e.g., brane). As in experiment 1, for the RT data, there were main effects of concreteness and frequency as well as a significant interaction where low frequency abstract words had a longer reaction time than low frequency concrete words. However, there were no significant effects of concreteness for high frequency words. The percentage error data reported was that there were significant main effects of frequency (more errors were incurred in low frequency than in high frequency words) and concreteness (more errors were incurred in abstract words than in concrete words) but the interaction was only marginally significant.

In the third experiment, James used nonwords which were pronounceable homophones. The RT data showed a main effect of frequency but no main effect of concreteness or a concreteness and frequency interaction. In the percentage error data, there were no significant effects (though the frequency main effect approached significance). These results were expected and James suggested that unpronounceable words do not require lexical access processing i.e., participants assessed the correctness of the orthographic letter combinations rather than engaging in word identification strategies.

In the final and fourth experiment, James' aim was to eliminate the interaction seen in experiments 1 and 2 by increasing subjective familiarity of targets. To do this, participants were given a sentence generation task where they were able familiarise themselves with the targets which would appear in the subsequent lexical decision task. Thus, each target word appeared on a card and participants were given ten seconds to make up a sentence containing that word. If a participant did not know the meaning of a word, it was explained to them and they were given an additional ten seconds to create the sentence. After a three minute break, the lexical decision experiment was conducted, pronounceable nonwords were used. The RT data showed a main effect of frequency; neither the main effect of concreteness nor the interaction was significant. The percentage error data reported was that there were no significant differences in error scores between the conditions. James concluded that concreteness had no significant effect on lexical decision time because recency served to raise the subjective familiarity of targets and the suggestion is that participants did not access meaning level semantic information.

Kroll and Merves (1986) conducted two lexical decision experiments where they orthogonally manipulated concreteness and frequency. Critical words were nouns and pronounceable nonwords were used. Concreteness of critical words was determined by pre-existing ratings as well as a rating task. That is, the experimenters initially selected abstract nouns from norms provided by Paivio, Yuille, and Madigan (1968) in which the rating scale used was the 7-point scale ranging from '1' denoting 'highly abstract' and '7' denoting highly concrete. In these norms, words with a rating of less than or equal to 3.77 were considered to be abstract. Following this, Kroll and Merves selected concrete words from Kučera and Francis (1967) frequency norms. If these concrete words appeared in the Kučera and Francis norms, their rating was used; the rating for concrete words ranged from 5.60 to 6.87. Words which did not appear in the Paivio et al. norms (and thus did not have a concrete rating) were judged by Kroll and Merves as being concrete on the basis of how easily the word gave rise to an image and the extent to which the words referred to objects accessible to sensory experience. Following this, the abstract and concrete words (212 in total) were matched word by word on length and frequency (from Kučera and Francis norms). These words were subsequently rated by a group of participants on their concreteness/imageability on a scale of '1' denoting 'highly abstract' and '7' highly concrete'. That is, there was no distinction made between the two variables and instructions to participants were to use imageability as well as the availability of sensory experience when rating the words. The mean rating for abstract words was 2.7 and for concrete words 6.2.

In Kroll and Merves' (1986) Experiment 1, one group of participants made lexical decisions to the concrete words, and a separate group to the abstract ones. The RT results showed a main effect of frequency, a main effect of concreteness but the interaction between frequency and concreteness was not significant. Moreover, the main effect of concreteness where lexical decisions to concrete words were faster than to abstract ones was significant for items and not by subjects. In terms of participant accuracy, the authors reported percentage errors made by participants for making a lexical decision in each condition. The results were a significant main effect of frequency (significantly more errors were made on low frequency words than high frequency words); a significant main effect of concreteness where there were significantly more errors were made on abstract words (6.4%) than on concrete words (3.6%); however this was significant by items and only marginally significant by subjects. The interaction between frequency and concreteness was not significant.

In Experiment 2, Kroll and Merves, like James (1975), presented all the target words to one participant group. The RT results showed a significant main effect of frequency, a main effect of concreteness as well as an interaction significant by subjects only. Follow-up tests revealed that RTs to concrete words were significantly faster than those to abstract words for low frequency words, therefore replicating the result of James (1975). In the participant accuracy data the main effect of frequency was significant (significantly more errors were made on the lowest frequency words); both the main effect of concreteness and the interaction between frequency and concreteness were not significant.

In another lexical decision experiment, De Groot (1989; Experiment 4), presented participants with Dutch words which were nouns and nonwords which were pronounceable words. Critical words were selected from an existing Dutch corpus of words (van Loon-Vervoorn, 1985) which provided imageability ratings (on the 7-point scale). Words with a mean imageability rating of 3.5 or above were considered to be high imageability and words with a mean imageability rating of below 3.5 were considered to be low imageability. The mean imageability ratings of words in the four conditions, LF-LI; LF-HI; and HF-LI and HF-HI were 2.8 (SD=0.5); 6.3 (SD=0.5); and 2.7 (SD=0.5); 6.4 (SD=0.4). However, the authors also obtained concreteness ratings of the critical words in a norming task where participants were asked to rate words on a scale of 1 to 7 (low concreteness to high concreteness respectively). Instructions were taken from an earlier study by Spreen and Schulz (1966). The mean concreteness ratings were comparable to the imageability ratings and for the four conditions, LF-LI; LF-HI; and HF-LI and HF-HI were: 2.6 (SD=0.7); 6.5 (SD=0.5) and 2.7 (SD=0.6); 6.3 (SD=0.6). The mean imageability ratings were highly correlated with the mean concreteness ratings (r=0.96, p=0.0001)

In De Groot, the RT data showed a significant main effect of frequency, a significant main effect of imageability and an interaction between frequency and imageability which was significant by subjects and marginally significant by items. Specifically, there were no significant differences between words high and low imageable high frequency words, only for low frequency words. This result therefore replicates that of James (1975) and Kroll and Merves (1986, Experiment 2). In terms of participant accuracy, De Groot reported percentage error rates per condition: most errors were incurred in the low frequency low imagery condition (7.1%); followed by in the low frequency high imagery condition (3.5%). High frequency low imagery condition incurred an error rate of 2.4% and high frequency high imagery an error rate of 1.4%. There were no statistical tests performed between the different conditions.

2.1.1 Present study

The relationship between word frequency and imageabilty remains an open question and the aim of this initial experiment of the thesis was to examine this relationship. A factorial design and a lexical decision task were used as a first step towards investigating the relationship between these variables. To address the question of the time course of imageability effects, specifically whether there are imageability effects in early lexical processing, it was examined whether the combined effects of word frequency and imageability are additive or interactive. Sternberg's (1969a; 1969b) additive factors method was utilised where imageability was orthogonally varied with word frequency since this latter variable indexes early lexical access (e.g., Sereno & Rayner, 2003). Previous research which has examined the combined effect of these two variables is limited and with conflicting results. In addition, experimental findings which have suggested that imageability may have an early time course (interactive finding; main effects of frequency and imageability and an interaction between the two variables) are problematic since failure to control for a potentially confounding variable, AoA, suggests that at least some of the reported effects of imageability on reaction time and errors and some of the reported interaction with frequency may in fact be due to AoA rather than to imageability. Nonetheless, it remains possible that there is an interactive relationship between frequency and imageability based on findings in the present thesis from the subsequently reported Experiments 2 and 3 which showed that frequency can interact with contextual predictability. In addition, a previous lexical decision study showed an interaction between frequency and contextual predictability (Stanovich & West, 1982). Therefore, it is reasonable to speculate that frequency may interact with imageabilty which is also a semantic or meaning level variable. For the present study, words were selected which were not correlated on imageability and AoA. Additionally, unlike many past studies, concreteness and imageability were not confounded with each other. Since the two refer to different concepts, an attempt was made to differentiate between the two. The combined effect of the frequency and imageability were examined on two standard lexical decision dependent variables: participants' reaction time (RT) in discriminating between a word and a nonword and on accuracy of making this discrimination which was indicated by measuring the percentage of errors (%Err) made in each experimental condition. If the semantic quality of concreteness exerts an early lexical effect then interactive effects with frequency should be obtained for both RT and %Err. Therefore the hypotheses for RT was that there will be a main effect of frequency (RT will be faster to HF words than to LF words); main effect of imageability (reaction time will be faster to HI words than to LI words); and a significant interaction where imageability is more beneficial for LF words than HF words (the frequency effect, LF>HF, would be bigger in LI words than HI words). Similarly, the hypotheses for %Err was that there will be a main effect of frequency (there will be more errors in discriminating words from nonwords when they are LF than when they are HF); a main effect of imageability (there will be more errors in discriminating words from nonwords when they are LI than when they are HI); and a significant interaction where imageability is more beneficial for LF words than HF words (the frequency effect, LF>HF, would be bigger in LI words than HI words). Theoretically, parallel models (e.g., McClelland, 1987; Morton, 1969) would predict simultaneous early effects of both word frequency and imageability which indicates a shared processing stage. Such an interactive view would be supported by main effects of word frequency and imageability as well as an interaction between the two variables. In contrast, serial models (e.g., Forster, 1979; Fodor, 1983) would predict late imageability effects with word frequency occurring first. This additive view would be supported by a main effect of word frequency, a main effect of imageability and no interaction between the two factors.

2.2 Method

2.2.1 Participants

Thirty participants, 15 male and 15 female, participated. All were from the University of Glasgow undergraduate and postgraduate community. The mean age was 22 and the age ranged from 18 to 30 years. All participants were native English speakers and had not been diagnosed with any reading disorders and all had normal or corrected-to-normal vision. One participant was left handed. All 30 participants were presented with the complete set of 210 stimulus words (10 practice trials, 100 real words and 100 nonwords). The duration of the experiment was approximately 20 minutes. Participants were naïve as to the purpose of the study and informed consent was obtained before participation.

2.2.2 Apparatus

The experiment was administered via a Mac G4 (OS 9.0.4) using PsyScope 1.2.5 PPC software (Cohen, MacWhinney, Flatt, & Provost, 1993). Stimuli (black letters on a white background) appeared in 24-point Courier font on a Hansol 2100A 19-inch screen with a 120 Hz refresh rate and 1024 by 768 pixel resolution. Viewing distance was approximately 86 cm and three characters of text subtended 1 degree of visual angle. Participant responses were recorded using a PsyScope Button Box allowing reactions times (RTs) to be recorded with millisecond accuracy. To indicate a real word had been seen, the participant made a response via their right index finger and to indicate that a nonword had been seen, the left index finger was used.

2.2.3 Materials

The total material set were 100 critical words (25 words per condition) and 100 word-length and syllable matched nonwords. To generate words, we used an online database, the MRC Psycholinguistic Database (Coltheart, 1981; http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). From this website, we obtained the following information for each word: concreteness rating; imageability rating; AoA rating; and number of syllables. Word frequencies were subsequently determined from the British National Corpus (BNC), a database of 90 million written words (http://natcorp.ox.ac.uk), mean frequency values for LF were occurrences of 15 million or below (range: 0-15 occurrences per million) and HF were occurrences of 40 million or above (range: 40-728 occurrences per million). Across the four conditions, we controlled for word length (mean word length was 5.36-5.7 characters; range: 4-8 characters, per condition). Words were all nouns. Words generated from the MRC Psycholinguistic database were entered into another online database, the English Lexicon Project (ELP; http://elexicon.wustl.edu/). This allowed several kinds of word information to be obtained: orthographic neighbourhood (the number of orthographic neighbours a word has); phonemes (the number of phonemes in the main pronunciation); morphemes (the number of morphemes a word has). In addition, the ELP also has stored behavioural results from a lexical decision task for all their stored words collected across 1200 participants at six different American universities. Thus, for each of word, we also obtained its mean RT (in milliseconds) and standard deviation; observations (number of observations that were made of the mean RT); and mean accuracy. These data are based on all 1200 participants.

Imageability norms reflect participants' ratings of words on a scale of 1 (low imageability) to 7 (high imageability). Similarly, concreteness ratings refer to participants' ratings of words on a scale of 1 (low concreteness) to 7 (high concreteness). For our critical words, we obtained imageability *and* concreteness norms and obtained the mean of the two ratings in order to determine an imageability rating of a given target word. This is because there have been few attempts in the literature to separate out these two variables and we thought it more meaningful to utilise as much of the information as possible. Low imageability words had a mean rating of 442 (SD=52.6; range: 366-543) and high imageability words had a mean rating of 567 (SD=44.3; range: 475-621). AoA ratings refer to participants' ratings on a seven-point scale of how early in life they think they learned a given word. The seven-point band ranges from 0-2; 3-4; 5-6; 7-8; 9-10; 11-12; and 13 plus years. Therefore, the earlier in life a word is learned, the lower it's AoA. Since it is known that AoA is negatively correlated with imageability (the earlier in life a word was learned; low AoA, the higher its imageability and likewise, the higher the AoA, the lower the imageability) (e.g., Bird et al., 2001) we wanted to control for this in order to ensure any results were due to the effects of imageability and not due to AoA as a confounding variable. To do this, critical words were selected on the basis of there being an absence of the typical negative correlation between imageability and AoA. Figure 2.1. shows a plot of the imageability and AoA scores for each condition. Based on an eye-balling of the data, the indication is that imageability and AoA were not correlated.

The complete set of 200 stimuli words (100 critical words and 100 word length and number of syllables matched nonwords) arranged according to their four conditions are presented in Appendix A. The nonwords were generated to match the words in terms of number of letters (mean of 5.6) and number of syllables (mean of 1.7). Nonwords were always orthographically legal and pronounceable letter strings without having the same phonology of a real word. For example, the nonword 'roble' was matched the word 'metal'. The tables presented in Appendix B shows the word specifications (number of letters; frequency per million; number of syllables, phonemes, and morphemes, orthographic neighbours, imageability rating and age-of-acquisition rating) for each of the 100 critical words. The mean values and SD's are shown in Table 2.1 which presents a summary of the specifications given in Appendix B. Across all four conditions, words differed on frequency and imageability; all other variables were controlled.

Figure 2.1. Distribution of imageability and AoA in the four conditions

Figure 4.1.1. Distribution of LF-LI words

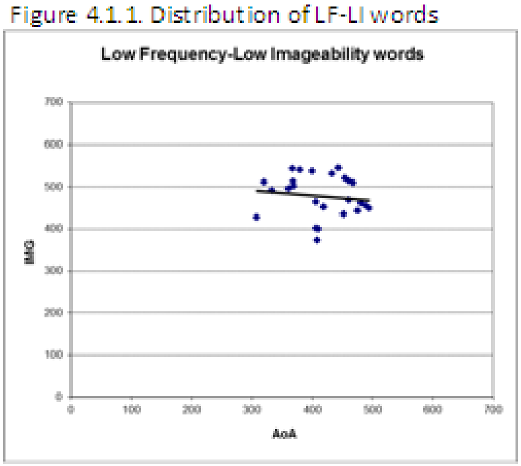


Figure 4.1.2. Distribution of LF-HI words

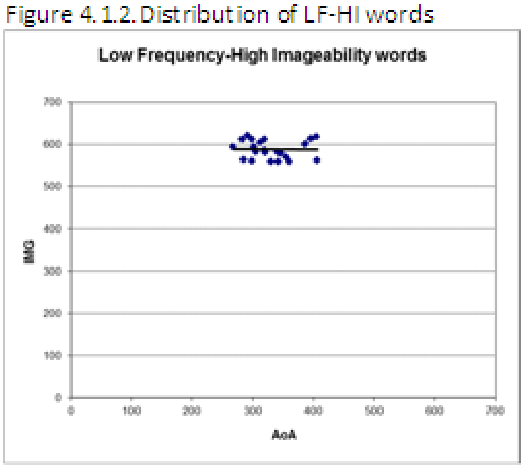


Figure 4.1.3. Distribution of HF-LI words

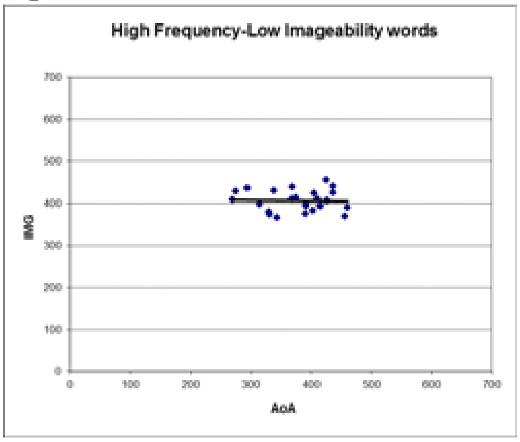
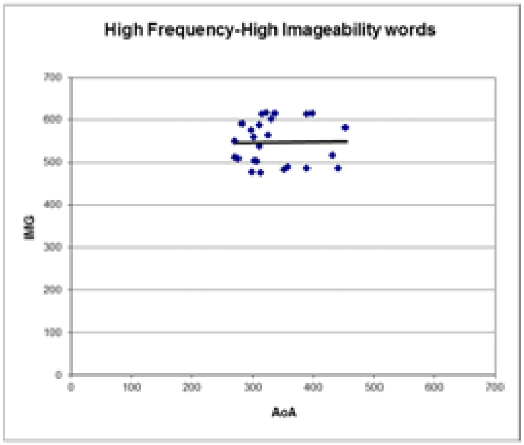


Figure 4.1.4. Distribution of HF-HI words



Note: LF = low frequency; HF=high frequency; IMG=imageability; AoA=Age of Acquisition; LI=low imageability; HI=high imageability

Table 2.1 Specifications of Critical Words in the Four Conditions

Condition	Mean lett. (SD)	Mean freq. per mill. (SD)	Mean no. syll. (SD)	Mean no. phon. (SD)	Mean no. morph. (SD)	Mean ortho. neigh. (SD)	Mean img. score(SD)	Mean AoA (SD)
LF-LI	5.76 (1.3)	7.1 (3.94)	1.84 (0.8)	4.72 (1.54)	1.24 (0.52)	3.2 (3.91)	477.68 (48.99)	414.76 (53.7)
LF-HI	5.76 (1.2)	5.35 (3.78)	1.76 (0.6)	4.72 (1.21)	1.04 (0.2)	3.04 (4)	587.29 (20.76)	333.91 (39.56)
HF-LI	5.64 (1.19)	212 (167.18)	1.52 (0.71)	4.32 (1.46)	1.2 (0.41)	2.64 (3.8)	406.47 (24.65)	379.03 (54.63)
HF-HI	5.36 (1.35)	108.74 (78.5)	1.76 (0.72)	4.32 (1.22)	1.2 (0.41)	5.2 (5.18)	546.33 (51.97)	336.63 (53.08)

Note: lett. = number of characters; SD = standard deviation; freq. per mill. = frequency per million; no. = number; syll. = syllables; phon. = phonemes; morph. = morphemes; ortho. neigh. = orthographic neighbours; img = imageability; AoA = age-of-acquistion; LF = low frequency; HF = high frequency; LI = low imageability; HI = high imageability. Imageability score range is 100 (low imageability) to 700 (high imageability); age-of-acquisition is from 100 (low AoA) to 700 (high AoA)

2.2.4 Design

A 2 (frequency: low frequency; LF, high frequency; HF) x 2 (imageability: low imageability; LI, high imageability; HI) within-subjects design was used. This lead to the following experimental conditions: LF-LI (low frequency-low imageability words); LF-HI (low frequency-high imageability words); HF-LI (high frequency-low imageability words), and HF-HI (high frequency-high imageability words). Each condition had 25 critical words and 25 nonwords matched on length and number of syllables. There were also 10 practice trials which were 5 real English words and 5 nonwords and these trials were the same for every participant. All participants took part in all four conditions. Stimuli were randomly divided into four blocks of 50 trials (a break was programmed in after each block) so that there was a different random order of presentation for every participant. Dependent variables were standard lexical decision measures of mean reaction time (RT) and mean percentage error (%Error) in discriminating the critical word from the nonword.

2.2.5 Procedure

Participants were tested individually in a booth housed within the School of Psychology at the University of Glasgow. Upon arrival, participants were given a consent form to read and sign and also written instructions pertaining to the experimental task. They were told that of the stimuli presented to them on the screen, half would be words and half nonwords and that their task was to respond quickly but as accurately as possible to indicate if the stimulus they had seen was a word or a nonword. They were instructed that their finger responses were to be made via a button box where a word was indicated on the right (green) button labelled ‘W’ with the index finger of the right hand. Similarly, a nonword was indicated on the left (red) button labelled ‘NW’ with the left index finger of the left hand. Participants were then presented with the 10 practice words followed by the 200 experimental words (100 critical words and 100 nonwords) arranged in four blocks of 50 letter strings.

All trials started with the presentation of a blank screen for 1000 ms followed by the appearance of a fixation cross (+) in the centre of the screen for 200 ms followed by another blank screen for 500 ms. A letter string was then presented in the centre of the screen which remained on the screen until the participant made a response via the button box. Experimental trials were presented in four blocks of 50 letter strings. After the presentation of each block, participants could have a short break. Following the break, the next block commenced with the participant indicating their intention to start the block by placing their hands back on the button box and the experimenter pressing the return key on the keyboard to resume the trials. Letter strings were presented in a different random order for each participant. At the end of the experiment, participants were debriefed. The total duration of the experiment was approximately 20 minutes in length.

2.3 Results

The mean reaction time (RT) in milliseconds (ms) and mean percentage error (%Error) data (with standard deviations) for the four conditions of low frequency-low imageability; low frequency-high imageability; high frequency-low imageability and high frequency-high imageability (LF-LI; LF-HI; HF-LI, and HF-HI) are presented in Table 2.2. In order to determine the effects of frequency and imageability, a 2 (frequency: LF, HF) x 2 (imageability: LI, HI) within-subjects repeated measures ANOVA was conducted both by subjects (F_1) and by items (F_2).

Table 2.2 Mean Reaction Time and Percentage Error for words across the Four Experimental Conditions

Condition	RT	%Error
LF-LI	568 (69)	7.6 (6.92)
LF-HI	567 (60)	8.93 (9.08)
HF-LI	503 (54)	2.93 (4.45)
HF-HI	498 (55)	1.87 (4.17)

Note: RT = mean reaction time, in milliseconds, %Error = mean percentage error; LF-LI = low frequency-low imageability; LF-HI = low frequency-high imageability; HF-LI = high frequency-low imageability; HF-HI=high frequency-high imageability

2.3.1 Reaction Time (RT)

With the reaction time data, prior to analysis, we removed individual data points for the following reasons: 1) participants who had made an incorrect response on the lexical decision task; 2) items which had RTs less than 250 ms or greater than 1500 ms (these were considered to be outliers) and 3) for each participant in each condition, items with RTs beyond two standard deviations of that mean were also excluded. These measures lead to a total data loss of 5.17%.

2.3.1.1 Main effects of frequency and imageability

HI words elicited faster responses than LI words (533 ms vs. 536 ms), however the small numerical difference of 3 ms is did no show as significant [all *F*s <1]. HF targets were identified much faster than LF targets (500 ms vs. 568 ms; $F_1(1,29)=188.85$, *MSE*=722, $p<.001$, and $F_2(1,24)=154.42$, *MSE*=826, $p<.001$).

2.3.1.2 Frequency x imageability interaction

Figure 2.2. shows the mean RT (with standard error bars) and is suggestive of no interaction being present since the lines are fairly parallel. Indeed statistical analyses confirmed this and in both subjects and items analyses, the frequency x imageability interaction was nonsignificant [all *F*s <1].

2.3.2 Percentage Error (%Err)

2.3.2.1 Main effects of frequency and imageability

It was hypothesised that abstract words would elicit greater errors than concrete words. Contrary to our prediction, HI words elicited slightly more errors than LI words (5.7 vs. 5.3 %Err); however, this numerical difference was nonsignificant in both subjects and items analyses [both *F*s <1]. The main effect of frequency occurred in the expected direction whereby participants made significantly more errors on LF words compared to HF words, both in subjects and items analyses (8.5 vs. 2.4 %Err; $F_1(1,29)=41.77$, *MSE*=27, $p<.001$, and $F_2(1,24)=15.01$, *MSE*=63, $p<.001$).

2.3.2.2 Frequency x imageability interaction in subjects analysis

Figure 2.3. presented below is suggestive of an interaction between frequency and imageability; results showed that this interaction was significant in subjects but not in items [$F_1(1,29)=4.26$, *MSE*=15, $p<.005$, and $F_2<1$]. To find out where the differences between groups lay, we carried out Bonferroni post-hoc tests by subjects which were four t-tests: we looked if means were different for LF-LI and LF-HI words; for HF-LI and HF-HI words; and then LI, LF and HF words and finally HI, LF and HF words.

The first t-test showed a marginally significant difference in %Err for LF-LI and LF-HI words [$F_1=3.45$, $p=.073$]; participants made marginally more errors on LF-HI words than on LF-LI words (8.93 vs. 7.6 %Err). For HF -LI and -HI words, there were no significant differences in %Err (2.93 vs. 1.87 respectively) between these two groups [$F_1=1.13$, $p>.15$]. Third, participants made more errors on LF-LI words than they did on HF-LI words [7.6 vs. 2.93 %Err; $F_1=21.56$, $p<.001$]. In the final t-test, HI, LF and HF words %Err means were compared; participants made significantly more errors on LF-HI words than on HF-HI words [8.93 vs. 1.87 %Err; F_1=57.17, p<.001].

Figure 2.2. Mean reaction times (with standard error bars) to low imageability and high imageability words by low and high frequency

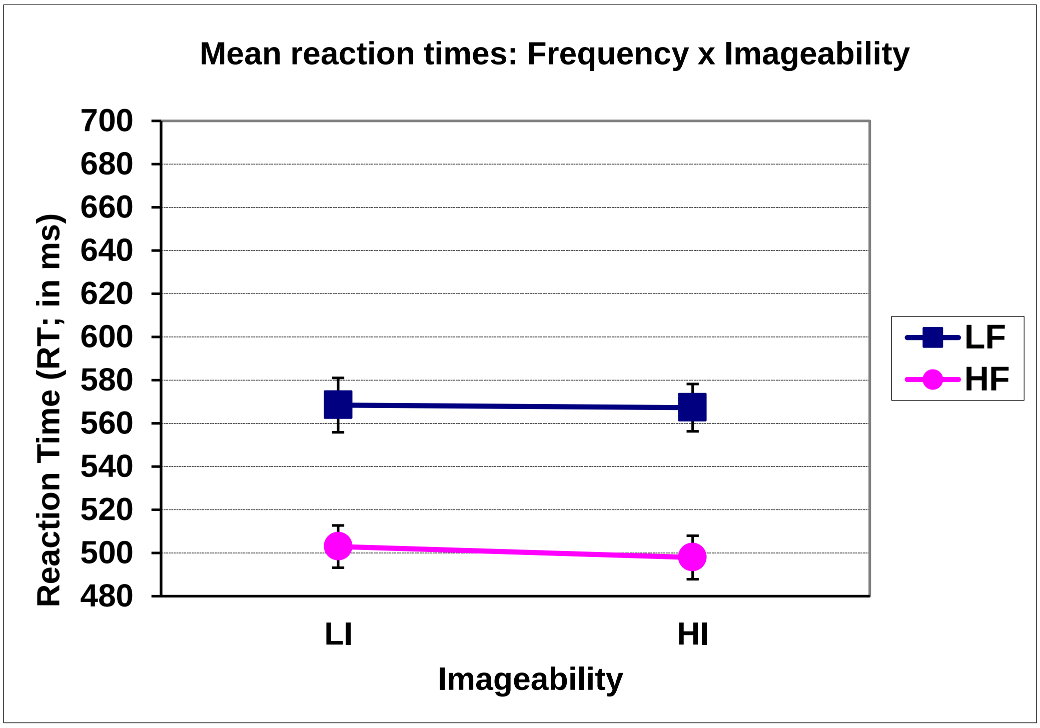
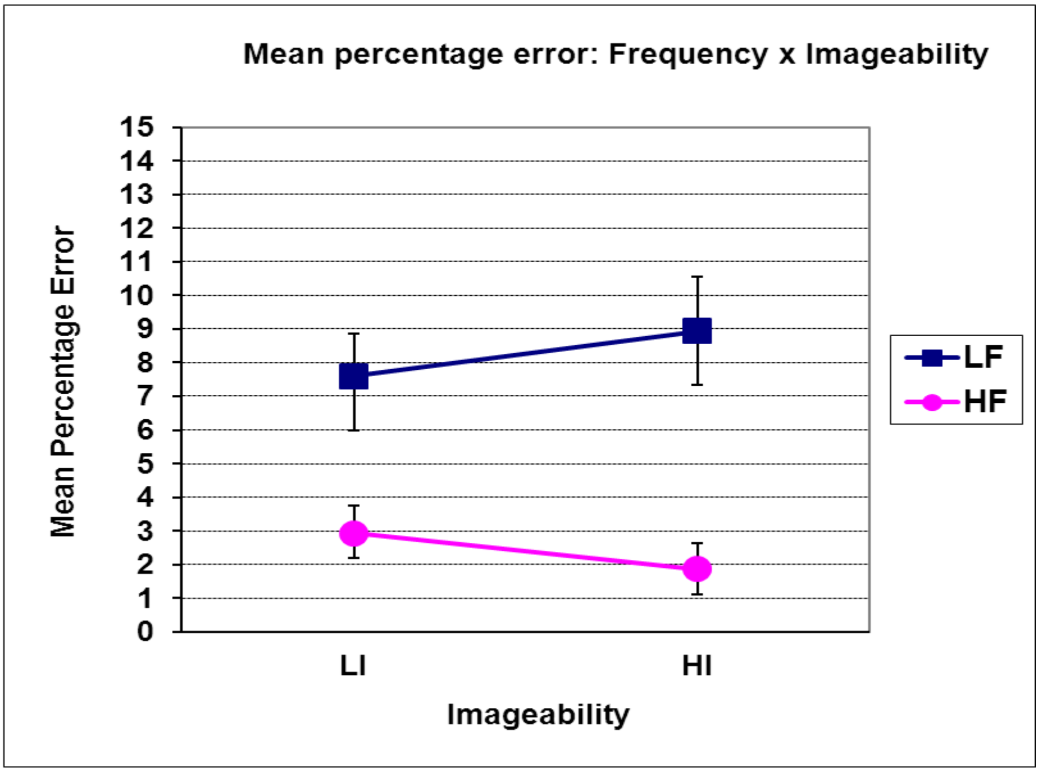


Figure 2.3. Percentage error rates (with standard error bars) to low imageability and high imageability words by frequency (low and high)



Figures 2.2. and 2.3. Notes : LF = low frequency; HF = high frequency; LI = low imageability; HI = high imageability

2.3.2.3 Frequency x imageability interaction in items analysis

As stated above, the interaction between frequency and imageability was significant in subjects but not in items. Closer inspection of each participant's %Err data revealed one item which incurred a much higher error rate than any other item: this was item number 13, 'muzzle' which appeared in condition LF-LI. Specifically, 53% of participants made an error on this word (for the condition LF-LI: mode=0; median=3.3). On the basis of this item representing an outlier in the data, we removed this item from the dataset. Table 2.3 shows the %Err of the condition LF-LI with the removal of this outlier item. A re-analysis of the data replicated the previous nonsignificance of the main effect imageability [F_2<1]. However, the imageability by frequency interaction (see Figure 2.4.) was significant $F_2(1,23)=4.33$, $MSE=39$, $p<.05$].

Table 2.3 Mean Reaction Time and Percentage Error for words across the Four Experimental Conditions with the Removal of One Outlier Item

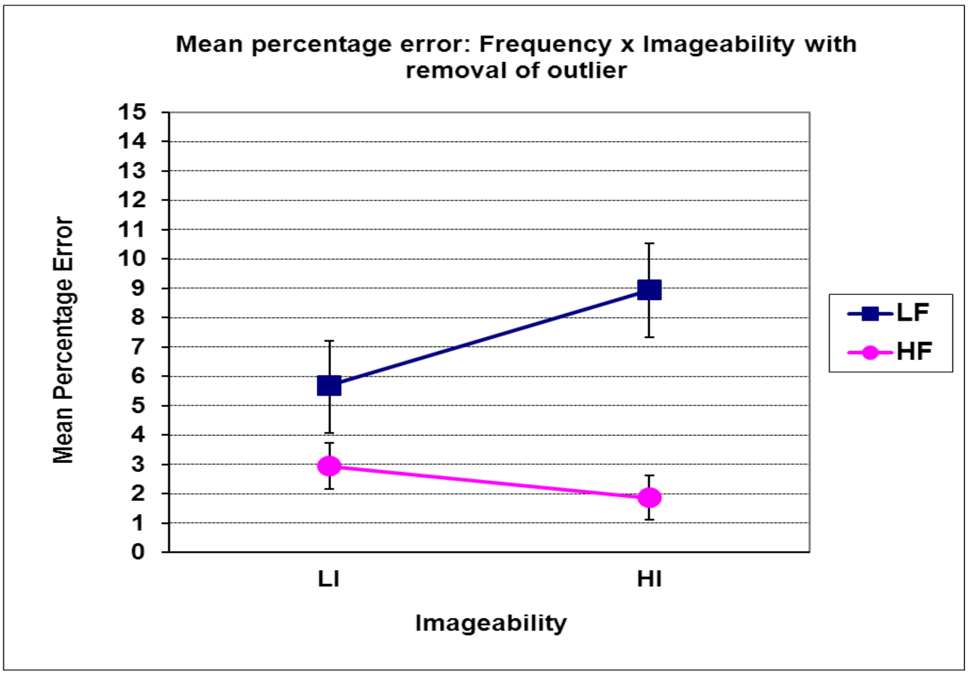
Condition	RT	%Error
LF-LI	568 (69)	5.69 (7.45)
LF-HI	567 (60)	8.93 (9.08)
HF-LI	503 (54)	2.93 (4.45)
HF-HI	498 (55)	1.87 (4.17)

Note: LF-LI = low frequency-low imageability; LF-HI = low frequency-high imageability; HF-LI = high frequency- low imageability; HF-HI = high frequency-high imageability

We conducted Bonferroni post-hoc t-tests to determine which groups differed. These four groups were LF-LI versus LF-HI words; HF-LI versus HF-HI words; LF-LI and HF-LI and finally LF-HI and HF-HI words. The first comparison showed that participants made significantly less errors on *LF-LI* words than on LF-HI words (5.69 vs. 8.93 %Err; F_2=5.39, $p<.05$). Second, the difference in percent error between HF-LI and

HF-HI words was not significant [$F < 1$]. Third, the data showed a trend for participants making more errors on LF-LI than they did on HF-LI (5.69 vs. 2.93 %Err; F_2=2.16, p=.15]. Finally, participants made significantly more errors on LF-HI words than on HF-HI (8.93 vs. 1.87 %Err; F_2=19.47, p<.001).

Figure 2.4. Percentage Error rates (with standard error bars) to low imageability and high imageability words by frequency (low and high) with removal of outlier item



Note: LF = low frequency; HF = high frequency; LI = low imageability; HI = high imageability

2.3.3 Results summary

The patterns of results in the present experiment are somewhat unclear regarding our research aims. The results are summarised in Table 2.4 and 2.5 below. In the RT results, the only significant result was a significant main effect of frequency. The %Err results showed a nonsignificant main effect of imageability but a significant main effect of frequency as well as a significant interaction of frequency x imageability. Follow-up analyses to the significant interaction revealed that LF-HI incurred significantly more errors than HF-HI words. For the equivalent comparison for LI words, LF words incurred more errors than HF words, being significant in subjects analysis but trend in items analysis. LF-LI words incurred more errors than LF-HI words: this was marginal in subjects analysis; in items analysis, however after removal of an outlier item, LF-LI incurred fewer errors than LF-HI item. For the equivalent comparison of LI-HI words which were HF, there were no significant differences in error rates in subjects and items analyses.

To summarise, the pattern of RT results showed a significant main effect of frequency, a nonsignificant main effect of imageability as well as a nonsignificant interaction. The pattern of %Error showed a main effect of frequency, a nonsignificant main effect of imageability and a significant interaction of frequency x imageability.

Table 2.4 Summary of ANOVA Results in Reaction Time and Percentage Error by Subjects (F_1) and by Items (F_2)

Main effect frequency						
Measure			***df***	***F***	***MSE***	***p***
RT	**HF < LF**	F_1	1,29	188.85	722	<.001
		F_2	1,24	154.42	825	<.001
% Err	**HF < LF**	F_1	1,29	41.77	27	<.001
		F_2	1,24	15.01	63	<.001
Main effect imageability						
RT		F_1	1,29	*<1*		*ns*
		F_2	1,24	*<1*		*ns*
% Err		F_1	1,29	*<1*		*ns*
		F_2	1,24	*<1*		*ns*
Interaction word frequency x imageability						
RT		F_1	1,29	*<1*		*ns*
		F_2	1,24	*<1*		*ns*
% Err		F_1	1,29	4.26	15	<.05
		F_2	1,24	*<1*		*ns*
		F_2 **(re-analysis)**	1,23	4.33	39	<.05

Note: RT = reaction time; %Err = percentage error

Table 2.5 Follow up Contrasts to Significant Interaction of Frequency x Imageability in Percentage Error by Subjects (F_1) and by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
% Err		LF-LI > LF-HI		HF-LI vs. HF-HI		LF-LI > HF-LI		LF-HI > HF-HI	
	F_1	3.45	=.073	1.13	>.15	21.56	<.001	57.17	<.001
		LF-LI < LF-HI		HF-LI vs. HF-HI		LF-LI > HF-LI		LF-HI > HF-HI	
	F_2 **(re-analysis)**	5.39	<.05	*<1*	*ns*	2.16	=.15	19.47	<.001

Note: %Err = percentage error; LF = low frequency; HF = high frequency; LI = low imageability; HI = high imageability

2.4 Discussion

In this Experiment, we made use of a lexical decision task with two standard dependent variables, mean RT data (the time taken to discriminate between a word and a nonword) and mean percentage errors (%Err; in discriminating between a word and a nonword) in order to investigate the relationship between frequency and imageability. Our study was designed as a first step to investigate the time course of imageability on early lexical access, as indexed by word frequency (e.g., Sereno & Rayner, 2003). Previous research investigating frequency and imageability has been limited and the small number of existing studies which have investigated this relationship have not controlled for AoA. This variable covaries with imageability (e.g., Barca et al., 2002; Bird et al., 2001; Gilhooly & Logie, 1980; Reilly et al., 2007) and a failure to control for this variable suggests that any observed results of main effects and interactions of imageability could be due to the effect of AoA rather than to imageability. Therefore in the present study, the low imageability and high imageability words were not correlated on imageability and AoA (see Figure 2.1.).

2.4.1 Reaction time

In the current study, the RT data showed a highly significant frequency effect where participants responded significantly faster to high frequency words than to low frequency ones. This frequency effect is well attested having been reported in all previous studies (e.g. De Groot, 1989; James, 1975; Kroll & Merves, 1986). A number of previous lexical decision studies have shown faster RTs to concrete words than to abstract words, that is the concreteness effect (Binder, 2005; Bleasdale, 1987; Chiarello et al., 1987; de Groot, 1989; Howell & Bryden, 1987; James, 1975; Kroll & Merves, 1986; Ransdell & Fischler, 1987; Rubin, 1980; Schwanenflugel & Shoben, 1983; Schwanenflugel et al., 1988; Whaley, 1978). However, our results were in line with those studies in which there were no significant differences in participant reaction time responses to abstract and concrete words (e.g., Feldman et al., 2006; Kroll & Merves, 1986: Experiment 3; Rubenstein et al, 1970; Richardson, 1976; Samson & Pillon, 2004; Schwanenflugel et al., 1988: Experiment 1; Schwanenflugel & Shoben, 1983: Experiment 2; Van Hell & De Groot, 1998: Experiment 3; 2008: Experiment 1). Also nonsignificant was the interaction between frequency and imageability. There have only been a small number of lexical decision studies which have looked at the interaction between these two variables and of these, some have reported a significant interaction between frequency and imageability (James, 1975: Experiments 1 & 2; Kroll & Merves, 1986; Experiments 1 & 2; De Groot, 1989: Experiment 4) while others like us have reported a lack of an interaction as well as a lack of the main effect of imageability (Rubenstein et al., 1970).

Rubenstein et al., (1970) in a lexical decision study showed a main effect of word frequency and no main effect of concreteness or an interaction between frequency and concreteness. Rubenstein et al.'s study was slightly larger than ours – they used 185 critical words presented to 39 participants and our study used 100 critical words to 30 participants. Rubenstein et al. did not obtain concreteness ratings from pre-existing norms, or alternatively carry out their own norming task; rather the authors themselves rated the words. In contrast, we used pre-existing imageability *and* concreteness norms which allowed us to obtain an average imageability score. This was in order to use as much of the pre-existing information as possible since previous researchers have made little attempt to make a distinction between concreteness and imageability. Therefore our study in not showing a main effect of imageability is surprising. In other words, regardless of frequency, there were no significant differences in reaction time to low imageable versus high imageable words. Numerically high imageable words did elicit faster responses by 3 ms than low imageable words but this difference was too small to show as significance. The nonsignificant interaction between frequency and imageability indicates that the conditions did not significantly differ from each other in participant reaction time responses.

It could be that the present study did not show the imageability main effect (in both RT and %Err results) and a frequency x imageability interaction for RT because the response time measurement was not sensitive enough to distinguish potential differences between low and high imageable words. It could be that that, for example, there is an interactive relationship between frequency and imageability in RT but that differences in response time to low versus high imageable words were not able to be detected by the response time measurement scale. A more sensitive measure is likely to be eye movement fixation durations as well as EEG recording which could detect potential differences between the two types of words, especially when they are all high frequency. That is, high frequency words typically lead to shorter response times anyway. It is possible that the response time measurement may not be sensitive enough to detect the differences that do potentially exist between abstract and concrete words which are all high frequency. In addition, displaying words in isolation has been criticised in terms of the extent results generalise to natural reading (e.g., Juhasz & Rayner, 2003). However, this explanation appears unlikely given there are some previous studies which have shown the main effect of concreteness and also the interaction (de Groot, 1989, Experiment 4; James, 1975, Experiments 1 and 2; Kroll & Merves, Experiment 2; Experiment 1, significant main effects only). Whilst there are other studies (Richardson, 1976; Rubenstein et al., 1970) which have not shown the main effect of imageability or the significant interaction, we had expected our results to resemble those of the first set of studies since our rationale was that frequency had been shown to have an interactive relationship with contextual predictability in lexical decision (e.g., West & Stanovich, 1982) such that it remains possible that frequency may also interact with word imageability. Therefore, one possibility that we did not show an imageability effect is because our low imageability and high imageability nouns were not sufficiently discriminated from each other. The mean rating (on a scale of 1 to 700; low imageability to high imageability respectively) for the low imageability nouns was 442 (SD=52.6) and the mean rating for the high imageability nouns was 567 (SD=44.3) (these values are collapsed across frequency). Across the four conditions, LF-LI; LF-HI; and HF-LI; HF-HI the mean ratings were: 478; 587; and 406, 546. This can be contrasted with a previous study (De Groot, 1989, Experiment 4) which did show the main effect of imageability; mean imageability ratings in the four conditions, LF-LI; LF-HI; and HF-LI; HF-HI were: 280; 630; and 270; 640. In another previous study which also showed the main effect of imageability (Kroll & Merves, Experiments 1 and 2, 1986) and the significant interaction (Kroll & Merves, Experiment 2, 1986) the mean rating for low imageability words was 270 and for high imageability words was 620. In contrast, the equivalent comparisons in the present study were 442 and 567 respectively. In addition, there was a small overlap in ratings between low imageability (minimum rating was 366 and maximum rating was 543) and high imageability words (minimum rating was 475 and maximum rating was 621). This highlights the difficulty in controlling for covariates which could potentially influence lexical decision performance. In attempting to control for a number of these other factors between the conditions (see Table 4.1: number of syllables, phonemes, morphemes, orthographic neighbours), we tried to ensure that these other variables were matched between the conditions of the experiment but perhaps imageability should have been better discriminated.

2.4.2 Percent error

The %Err data showed a highly significant main effect of frequency where participants made significantly more errors on low frequency words than on high frequency ones. Previous studies which have also shown this are Kroll and Merves (1986; Experiment 2) and James (1975; Experiments 1 & 2). The results also showed that the main effect of imageability was nonsignificant, that is, the difference in error between low imageability words (5.3%) and high imageability words (5.7%) was nonsignificant. This result resembles those of Rubenstein et al., (1970). They reported percentage of correct responses in each condition from which we can calculated percentage of error incurred in the conditions: low and high frequency abstract words (8% & 2% respectively); low and high frequency concrete words (5% and 1% respectively). There were no statistical tests performed on the correct responses data; the numbers (10% error with abstract words and 6% error with concrete words) while a larger difference than in the present results, are possibly suggestive of there being a nonsignificant main effect of imageability. In contrast, James (1975, Experiment 2) and Kroll and Merves (1986; Experiments 1 & 2) reported significantly more errors in discriminating abstract than concrete words. It should be noted however that in Kroll and Merves second experiment this difference was significant only in the items analysis.

In the present study, the frequency x imageability interaction was significant in both subjects and items analyses (with the removal of an outlier item). The most reliable difference was that low frequency-high imageable words incurred significantly more errors in discrimination than high frequency-high imageable words. Also in the expected direction was that low frequency-low imageable words incurred more errors than high frequency-low imageable words, significant in subjects but trend in items. Another contrast showed that there were no significant differences between low and high imageability with high frequency words. This was also shown in James' (1975; Experiment 1) study in which there was no significant differences between high frequency abstract and concrete words. In the final contrast, the low frequency-low imageable words incurred more errors than low frequency-high imageable words (expected direction), however this was only marginally significant in subjects analysis; in the equivalent comparison in the items analysis (with the removal of the outlier value) there was the opposite result in that low frequency-low imageable words incurred *less* errors than low frequency-high imageable words. In contrast, James (1975; Experiment 1) reported significantly more errors (only subjects analyses was conducted) in the low frequency abstract condition (13.5%) compared to errors in the low frequency concrete condition (4.2%). Thus in the present study, in the items analysis of low frequency words, it appears as if there was a processing advantage for low imageability words.

A recent lexical decision study by Kousta et al. (Experiment 1, 2011) investigated abstract words, in particular how they may be processed and represented. In their study, there were 40 critical abstract and 40 critical concrete words. Abstract and concrete words were actually low and high concrete words respectively and the study made a distinction between imageability and concreteness. Thus, critical words differed on concreteness only (abstract rating was 354 and concrete rating was 552), and critical words were matched on a number of other variables including word length, imageability, AoA, and log frequency. Pseudowords were created by selecting a further 40 abstract and 40 concrete words matched with the critical words in terms of concreteness; the pseudoword was created changing one letter in these words. Results showed that low concrete/abstract words were actually recognised faster (had a faster reaction time) than high concrete words. In the accuracy measure, results showed a numerical advantage for low concrete words compared to high concrete words (that is, participants made less error in discriminating words from pseudowords when they were low concrete compared to when they were high concrete); however this difference was marginally significant in the subjects analysis and nonsignificant in the items analysis. In addition, they carried out a regression analyses in lexical decision data from the ELP. These results showed that for accuracy rates, abstract words had an advantage over concrete words (imageability and context availability were partialled out).

Kousta et al. (2011) went on to reject both dual-coding and context availability models since their study had not demonstrated the concreteness effect, instead showing a reaction time and accuracy advantage (numerically only) for abstract words. Instead, the authors argued for an earlier suggestion which had previously been argued by Altarriba, Bauer, and Benvenuto (1999). That is, Altarriba et al. (Experiment 1) investigated abstract, concrete and emotion words (that is, words which denote emotional states). They argued that emotion words tend to be classified as abstract without any reasoning, and that it is possible that emotion words in the abstract category could increase or decrease the effect of concreteness in language processing, that is serving as a possible confound. Therefore, in their study, they collected ratings for abstract, concrete and emotion words (respective categories were decided a priori). Emotion words were classified as such if they had an affective meaning and have (un)pleasantness and arousal. Participants rated a total of 326 words on three scales: concreteness, imageability and context availability. The results were that concrete, abstract and emotion words received different ratings on the three scales; the suggestion being that concrete, abstract and emotion words have different amounts of concreteness, imageability and context availability. Therefore, Altarriba et al. posited that emotion words should not be included in the category of abstract words when investigating concreteness effects and if they are included, then it is possible that this may have increased the ratings of the abstract word on imageability and may have reduced the ratings of the abstract words on the concreteness and context availability scales. Based on this view, Kousta et al. suggested that the abstract words in their study appeared to have more affective associations than the concrete words; they do not state whether it is pleasant or unpleasant but from the materials provided, the abstract words category includes 20 words with either pleasant or unpleasant affect: words such as ‘horror’, ‘grief’, ‘demon’, ‘crime’, ‘angel’, ‘joy’, ‘love’ ‘paradise, whereas in the concrete words category, there were 4 words which are so and all negative: ‘cancer’; ‘asbestos’; ‘disease’ and ‘weapons’. However, Kousta et al. state that not many words referred directly to emotions but when those that did were excluded from analyses, the pattern of results stayed the same. Nonetheless, since the abstract words had far more affective associations than concrete words, it is possible that the observed results (where abstract words had an advantage over concrete words in RT, and numerically in accuracy) may have been due in part to the confounding of concreteness and affective association. In the present study, it appears as if the low frequency-low imageability critical words also had many more (negative) affective associations than the low frequency-high imageability words. For example, the first category had seven words with negative associations (rust, dent, liar, brawl, grief, poison, temper) while the low frequency-high imageable words had just one (rifle). With the high frequency words, the spread was even: high frequency-low imageable words had two negative association words (error, force) and high frequency-high imageable one (fight). Therefore, in the present study, the accuracy results in the items analysis (where low frequency-low imageable words incurred less error in discrimination than low frequency-high imageable words) could be due to there being more negative emotion words in the first category than in the latter: it is possible that there was a confound of imageability and affective association. In particular, the inclusion of the 7 emotion words in low frequency-low imageable category could have increased the overall rating of imageability of this group (which was 477.68) than the comparable case for the low frequency-high imageable words (which was 587.29). However, since the present results did not find the abstract word advantage for the RT data, more research is needed to fully understand the extent of the relationship. In particular regression analyses would be useful in manipulating this data.

2.4.3 General discussion

We previously argued that past studies which have shown the concreteness effect had not controlled for AoA and hence the observed effect could be attributed to AoA rather than to imageability or concreteness. A natural negative correlation exists between imageability and AoA in that high imageable words such as zebra are more likely to also have an early AoA score, and low imageable words like dogma tend to have a later AoA rating (e.g., Bird et al., 2001; Gilhooly & Logie, 1980). AoA ratings are scored on a 7-point scale where the following bands correspond to the age ranges 0-2 (100); 3-4 (200); 5-6 (300); 7-8 (400); 9-10 (500); 11-12 (600), 13 plus (700).

In the present study we attempted to control for AoA by using words which were not correlated on imageability and AoA. It is possible that is still an existing, albeit weak correlation, between imageability and AoA (see Figure 4.1). The AoA rating (see column 9, Table 2.1, also Appendix B) for low frequency-low imageable words was slightly higher than the AoA rating for low frequency-high imageable words (414.76 vs. 333.91 respectively). This could indicate that low frequency-low imageable words were learned later in life than the low frequency-high imageable words. Similarly, the AoA rating for high frequency-low imageable words was higher than the AoA rating for high frequency-high imageable (379.03 vs. 336.63 respectively) perhaps indicating that high frequency-low imageable words were learned later in life than the high frequency-high imageable words. These numerical differences are so small though that this seems unlikely. Also if it the case that is AoA and not imageability which drives imageability, then our results should have shown the imageability main effect. Even though the RT’s were suggestive of the pattern that would indicate the processing advantage with words learned earlier in life, they did not reach significance: high frequency-low imageability words had a longer reaction time (i.e., took longer to recognise) than the high frequency-high imageable words – this would be the expected direction of reaction time since words learned earlier in life (high frequency high imageable) would be processed more quickly than words which are learned later in life (high frequency low imageable) (e.g., Juhasz 2005). Similarly, low frequency words-low imageable words (learned later in life) had a longer reaction time than the low frequency-high imageable words (learned earlier in life).

This study therefore indicates that it is very difficult to separate these naturally highly related variables. A number of early factorial design studies which showed the concreteness effect (e.g., Boles, 1983; Day, 1977; Paivio & O’Neill, 1970; Rubenstein, Garfield, & Millikan, 1970; Winnick & Kressel, 1965) have potentially confounding variables. For example, word familiarity of a given word (which was not measured in most early studies) could account for at least some of the observed concreteness effect (e.g., Schwanenflugel et al., 1988). Recent work has suggested that imageability (as well as frequency) is correlated with AoA and thus that caution is needed when attributing the effect of each respective variable (Izura et al., 2011). The present study attempted to do this, with mixed results. In another study using the naming task, researchers showed effects of both frequency and AoA when imageability was controlled but when they controlled for AoA and frequency, no effect of imageability was found (Monaghan & Ellis, 2002). Thus many of the factors that influence the ease with which a word is recognised are correlated with each other.

Factorial designs may not be optimal in the task of only manipulating variables of interest whilst controlling the other correlated variables which affect word recognition time. It may be useful to utilise a correlation design where all of the factors are part of a larger scale regression analyses. In particular, factorial designs are useful for discrete categories but psycholinguistic variables are continuous and grouping such variables into discrete categories results in a statistically less powerful design.

The lexical decision task has been subject to criticism. When participants make a lexical decision (to a real word or a nonword), they must engage in the lexical access process by accessing the stored representation in the brain. Such a representation may include the word's orthography, phonology, or meaning. However RT is relatively slow and research from eye movements indicates that the average time take to access a word's representation is around 250 ms. RT can reflect conscious prediction strategies as well as the time taken to initiate and execute the motor response required in this task. Such decision processes and motoric responses are not normally part of word recognition. There is also the issue of the speed accuracy trade-off (the faster the participant responds, the more errors they will incur; Pachella, 1974). Therefore, care needs to be taken in specific instructions: if participants are encouraged to respond accurately, then there is a likelihood of them responding accurately but more slowly; similarly, if they are encouraged to respond as fast as possible then this often results in faster responses but with more errors. However, even though lexical decision performance depends on the combination of the ease with which lexical information is processed along with the influence of post-lexical decision processes, the lexical decision task is extremely useful in the field of word recognition. There is growing consensus that word recognition processes are best understood by considering converging evidence obtained from utilising various experimental methodologies (Sereno & Rayner, 2003).

While this study concerned the use of the lexical decision task to investigate the relationship between frequency and imageability, the relationship between these two variables has been investigated using other methods. Despite the goal of converging evidence across a variety of experimental paradigms, findings are very variable. For example, Richards (1976) reported two threshold identification experiments (in this method, words are degraded by brief tachistoscopic presentations) where there were main effects of concreteness and frequency and no interaction. In an earlier threshold identification study, Winnick & Kressel (1965) investigated the effects of concreteness and frequency on threshold identification performance. The variable they manipulated was how many 10 ms exposures participants needed before they were able to identify the target word correctly. Results showed that there were actually no significant differences in quantity of exposures required to identify abstract and concrete word; in both cases participants needed 9.65 exposures. However, there are converging results from a threshold identification study by Paivio & O'Neill (1970) in which in the significant interaction, low frequency concrete words took longer to respond to than low frequency abstract words.

The technique of monitoring participants' eye movements while they are reading text (which includes a target word which has been manipulated with the variable/s of interest) is recognised as an ecologically valid method of measuring word recognition processes. It is assumed that fixation time reflects the time taken to process the word (Inhoff & Radach, 1998; Rayner, 1998; Reichle et al., 1998). Only one eye movement study to date has investigated the time course of word imageability (Juhasz & Rayner, 2003). However, this study examined five intercorrelataed variables on word recognition (using a multiple regression analysis); there was no independent manipulation of word frequency. Future research could investigate the combined effect of frequency and word imageability whilst participants' eye movements are recorded. A multiple regression could serve to elucidate the relative contribution of the various lexical and semantic variables to RT and %Err.

2.4.4 Conclusions

The present study addressed the question of whether semantic factors affect early lexical processing, specifically lexical access. The lexical variable, word frequency, was used as an indication of lexical access. Word imageability is also a lexical level variable and indicates the meaning of a word. Specifically, we orthogonally manipulated frequency and imageability and measured reaction time and participants' accuracy on making word/nonword decisions in the lexical decision task. The theoretical implications from the present study i.e., the timing of imageability effects has implications for the architecture of the language processing system, in particular whether processing in the system is discrete or cascaded. In a modular architecture, semantics can only operate on the output of the lexical processor (e.g., Fodor, 1983; Forster, 1979). Therefore in such a model, there should be main effects of frequency and of imageability and no significant interaction. Such an additive result is suggestive of discrete processing stages in the language processing system. However, an interactive model posits that semantics can directly affect lexical access presumably via parallel processing stages (e.g., McClelland, 1987; Morton, 1969). Thus, evidence of such an interactive system would be evidenced by the main effects of frequency and of imageability, as well as the significant interaction, indicating that both variables are influencing the same stage of processing. This would suggest that semantic processing takes place in the system at the time of early lexical access.

The prominent view of visual word recognition, that is traditional bottom-up models of the language processing system, is that readers have to first match the visual stimulus to some internal representation before the meaning of the stimulus is made available. That is, recognition precedes meaning access. For example, in Morton's autonomous model, word recognition devices (logogens) have to receive a certain amount of activation through featural detectors before the word is identified. The meaning of the word is not made available until this threshold has been reached (see Becker, 1980; Forster, 1979; Norris, 1986, for similar views). This idea that recognition has to precede meaning access certainly seems reasonable, that is, how can a reader possibly access the referent of a word without first assessing the identity of the word? How could the system have access to the meaning without knowing what the stimulus is? However, experimental reports which show that semantic variables associated with lexical representations, such as the imageability of the word, can affect lexical access (presumably via cascaded top-down activation) challenge this assumption. In the present study, the pattern of reaction time results does not support an additive or an interactive view. However, the pattern of percent error is somewhat supportive of an interactive view; our study showed a significant main effect of frequency as well as a significant interaction, however, the main effect of imageability was nonsignificant.

Chapter 3

Word frequency and contextual predictability effects in an eye movement reading study: evidence for early interactive processing

Pre-test 1

Rating task

3.1 Method

3.1.1 Participants

The rating task was administered to a group of 20 participants (15 female, 5 male) all of whom were Native English speakers and attending University of Glasgow as undergraduate students (mean age 23.4 years). Care was taken to ensure that these participants did not later go on to participate in the eye tracking experiment or the Cloze task part of. The experiment was advertised as offering two course credits. Participants did not have any learning or reading disorders.

3.1.2 Apparatus

This word rating experiment was administered using the pencil and-paper method. The passages of text, with the target word in bold font, were typed and presented on double-sided A4 pages stapled together to form a booklet. An instruction sheet and consent form were typed and included also.

3.1.3 Materials

The passage of text was presented in its entirety; the sentence frame after the target word i.e., the post-target context was presented as well (as in Hand et al., 2009). An instruction sheet asked participants to circle on a 7-point scale how predictable they considered the word in bold (the target word) to be given the preceding context up to that point; 1 corresponded to 'low predictable' and 7 corresponded to 'highly predictable' (this appeared at the top of every page as a title bar). Participants were instructed to read each passage carefully and after doing so, to circle the first number that came to mind. Figure 3.1. below shows the presentation format of 12 example passages; two from each condition (these passages were later used in the eye tracking experiment). Conditions are labelled here only; participants were presented with a random order of passages with no reference to condition labels. In total, participants were presented with 180 passages of text and after scoring this task and the Cloze task (see pre-test 2 on p.108), 150 final experimental passages were selected.

Figure 3.1. Example presentation used in the Rating task

Low predictable **High predictable**

1 **2** **3** **4** **5** **6** **7**

[Condition 1 LF-LP]

Arthur was at home, preparing vegetables to accompany his dinner.

He steamed some **peas** and spooned them onto the side of his plate.

Overnight, vandals had ruined my prized cherry blossom tree.

They had ripped off the **bark** and scattered it across the lawn.

[Condition 2 LF-MP]

Colin couldn't resist the advances of the sexy new secretary.

He was overcome with **lust** and embraced her passionately.

The critically acclaimed restaurant was fully booked once again.

They had hired a talented **chef** who had transformed their menu.

[Condition 3 LF-HP]

Martin lost his temper and broke his wife's favourite ornament in two.

He would need to find some **glue** to repair it before she got home.

Jimmy ran down the hill, gripping the string in the strong winds.

He loved to play with his **kite** but rarely got the right conditions.

[Condition 4 HF-LP]

The zookeepers were busy preparing for their latest arrival.

They were getting a baby **bear** that had been born in America.

The sales team were pushing hard as the end of the month loomed.

They had been set a certain **goal** by the director of the company.

[Condition 5 HF-MP]

My favourite hobby is going to see musicals at the theatre.

I normally pay extra so that my **seat** is near to the stage.

I struggled to read the badly printed manual for my new computer.

It had little space between the lines of **text** and strained my eyes.

[Condition 6 HF-HP]

Gary had just started a new job helping tidy up the barber's shop.

He had to sweep up the **hair** from the floor at the end of each day.

When Alex arrived at his friend's house, he rang the bell.

He heard footsteps behind the **door** as his friend came to let him in.

3.1.4 Design

The design of the experiment reflected the later eye tracking experiment. That is passages of text used a within-subjects 2 x 3 design: the effect of frequency (LF, HF) on contextual predictability (LP, MP, HP) was of interest. The dependent variable was the rating given to the target: a 7-point scale was presented at the top of every page, the score of 1 was 'low predictable' and 7 was 'high predictable'. The final set of chosen 150 passages are displayed in Appendix C; Appendix D shows the mean rating score for each of the 150 target words and the mean rating scores for each respective condition are displayed in column 6 (highlighted in green here) in Table 3.1 (also shown is column 7 for pre-test 2).

Table 3.1 Summary Specifications of Target Words in the Six Conditions

Condition	Mean lett. (SD)	Mean freq. per mill. (SD)	Mean no. syll. (SD)	Mean img. score (SD)	Mean rating (SD) *	Mean Cloze (SD) **
LF-LP	5.88 (1.33)	6.98 (3.88)	2.08 (0.95)	562.47 (86.84)	4.28 (0.66)	0.01 (0.02)
LF-MP	5.88 (1.33)	7.26 (4.01)	1.64 (0.64)	561.20 (80.34)	5.58 (0.69)	0.52 (0.17)
LF-HP	5.88 (1.33)	6.67 (3.58)	1.88 (0.83)	575.17 (75.28)	6.21 (0.34)	0.96 (0.05)
HF-LP	5.88 (1.33)	179.57 (142.00)	1.80 (0.71)	501.10 (112.98)	4.47 (0.74)	0.01 (0.02)
HF-MP	5.88 (1.33)	178.86 (133.85)	1.64 (0.76)	513.84 (77.88)	5.81 (0.71)	0.56 (0.16)
HF-HP	5.88 (1.33)	179.86 (109.79)	1.76 (0.72)	542.20 (97.31)	6.26 (0.23)	0.97 (0.04)

* pre-test 1
** pre-test 2 (see p. 108)
Note: SD = standard deviation; LF = low frequency, HF = high frequency, LP = low predictability, MP = medium predictability; HP = high predictability; lett = number of letters; freq. per mill. = frequency per million; no. syll. = number of syllables; imageability score range is 1 (low imageability) to 7 (high imageability); rating range is 1 (highly unpredictable) to 7 (highly predictability); Cloze values are the mean probability of participants guessing the correct target presented in their given contexts

3.1.5 Procedure

Participants arrived at the School of Psychology to take part in the experiment. They were taken to a quiet booth and were given a printed copy of the task, complete with instructions and a consent form. Once they had read these, they were given the opportunity to ask questions for further clarification. Once they felt ready, they proceeded with the task. The experimenter left the booth whilst the participant completed the task. The experiment took approximately 50 minutes to complete. At the end, participants were debriefed and given their two course credits.

3.2 Results

The mean rating scores obtained across the six experimental conditions, LF-LP; LF-MP; LF-HP; HF-LP; HF-MP and HF-HP were as follows: 4.28 (.66); 5.58 (.69); 6.21 (.34); 4.47 (.74); 5.81 (.71) and 6.26 (.23). We were interested in the effects in the materials so the ANOVA was conducted by items ($F_{2)}$ analysis. Follow-up tests were Bonferroni multiple comparisons corrections.

The results showed that the main effect of frequency was significant [$F_2(1,24)=4.26$, $MSE=.223$, $p<.05$]. Thus, HF words were rated as (unexpectedly) significantly higher in predictability than LF ones (HF: 5.52 vs. LF: 5.36) despite the very small numerical differences. The results also showed a significant main effect of predictability [$F_2(2,48)=145.56$, $MSE=.316$, $p<.001$]. Bonferroni follow-ups to this main effect showed that all three contrasts significantly differed from each other (LP vs. MP and LP vs. HP and MP vs. HP). The results were as follow: LP versus MP (4.37 vs. 5.7: $F_2=138.95$, $p<.001$); LP versus HP (4.37 vs. 6.24: $F_2=274.79$, $p<.001$] and MP versus HP (5.7 vs. 6.24: $F_2=22.94$, $p<.001$). These contrasts show that HP targets were rated as significantly more predictable than MP and LP ones, and that MP targets were rated as significantly more predictable than LP ones.

As expected, the interaction of frequency x predictability was nonsignificant [$F_2<1$]. However, since the main effect of frequency had shown to be unexpectedly significant, it was decided to examine the frequency contrasts comparing low and high frequency at each of the three levels of predictability to see if results here were in the desired direction. When examined, these contrasts were indeed in the predicted direction as well as all three frequency contrasts at low, medium and high predictability being nonsignificant. Specifically, the LP frequency contrast compared LF-LP versus HF-LP (4.29 vs. 4.47; $F_2=1.13$, $p>.25$); the MP frequency contrast compared LF-MP versus HF-MP (5.58 vs. 5.81; $F_2=1.68$, $p>.20$); the HP frequency contrast compared LF-HP vs. HF-HP (6.21 vs. 6.26; $F_2<1$). These results showed that at each of the three levels of predictability, there were no significant differences in the rating score for low and high frequency targets. These results indicate that at low, medium and high predictable conditions, rating scores were matched across the two frequency conditions.

Since we examined frequency contrasts, we went on to examine the six predictability contrasts (comparing LP vs. MP; LP vs. HP and MP vs. HP for LF and HF targets). The predictability contrasts at low frequency and at high frequency were all significant. First the LF contrasts compared LP versus MP (4.29 vs. 5.58; $F_2=52.83$, $p<.001$); LP versus HP (4.29 vs. 6.21; $F_2=115.85$, $p<.001$); MP versus HP (5.58 vs. 6.21; $F_2=12.22$, $p<.01$). The same contrasts for HF targets were also significant. First, LP versus MP (4.47 vs. 5.81; $F_2=56.26$, $p<.001$); second LP versus HP (4.47 vs. 6.26; $F_2=100.12$, $p<.001$) and third MP versus HP (5.81 vs. 6.26; $F_2=6.28$, $p<.05$). These results show that across high frequency targets (and low frequency targets), high predictable targets had a significantly higher rating score than both medium and low predictable targets and medium predictable targets had a significantly higher rating score than low predictable targets. These findings indicate the validity of our predictability variable when targets were assigned to low, medium and high predictability conditions.

3.2.1 Results summary

The results are summarised in Tables 3.2, 3.3, 3.4 and 3.5 below. Participants were asked to rate targets according to how well they fit a passage of text (with ratings ranging from a scale of 1-7, where the higher the rating, the better the target fit with the context). When results were analysed across items, they showed a main effect of frequency, of predictability and a nonsignificant interaction. Since the main effect of frequency had shown to be significant, it was considered that the follow-ups to the nonsignificant interaction could tell us whether effects were occurring in the desired direction. When examined, these follow-ups showed the expected pattern of effects because the frequency contrasts were all nonsignificant. This indicates that the rating scores for low and high frequency targets were matched in the low predictability, medium predictability and high predictability conditions. In addition, we also examined the six predictability contrasts and all of numeric differences were in the right direction as well as being very highly significant. Thus, in the high frequency targets (and also low frequency ones), the higher predictable condition contained targets rated higher in predictability than those in the medium predictable as well as low predictable condition, as well as the medium predictable containing targets higher in predictability than those in the low predictable conditions. All of these numeric differences were significant, and most were highly so as indicated by the very high *F*-ratios. These serve to indicate that our predictability manipulations were very highly valid.

Table 3.2 Summary of ANOVA Results in Rating Task by Items (F_2)

Main effect frequency					
Measure		***df***	***F***	***MSE***	***p***
Rating score	F_2	1,24	4.26	.223	<.05
Main effect contextual predictability					
Rating score	F_2	2,48	145.56	.316	<.001
Interaction word frequency x contextual predictability					
Rating score	F_2	2,48	*<1*	.403	*ns*

Table 3.3 Follow up Contrasts to Significant Main Effect of Contextual Predictability by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***
Rating score	***F* 2**	LP < MP		LP < HP		MP < HP	
		138.95	<.001	274.49	<.001	22.94	<.001

Note: LP = low predictability; MP = medium predictability; HP = high predictability

Table 3.4 Follow-up Frequency Contrasts (LF vs. HF) by Items (F_2)

Measure		**LP words**		**MP words**		**HP words**	
Rating score	***F* 2**	***F***	***p***	***F***	***p***	***F***	***p***
		LF < HF		LF < HF		LF < HF	
		1.13	>.25	1.68	>.20	*<1*	*ns*

Note: LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

Table 3.5 Follow-up Contextual Predictability Contrasts by Items (F_2)

Measure		**LF words**						**HF words**					
Rating score	***F* 2**	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
		LP < MP		LP < HP		MP < HP		LP < MP		LP < HP		MP < HP	
		52.83	<.001	115.85	<.001	12.22	<.01	56.26	<.001	100.12	<.001	6.28	<.05

Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability

3.3 Discussion

The purpose of the word rating task, also known as the predictability task was to determine the rating score of each potential target word that fitted into a particular context. The aim was to conduct the rating task (along with the Cloze task, see pre-test 2 on p. 108) and decide the final set of materials to be used for the eye tracking experiment. The pencil-and-paper method was used to administer the task. Thus, the two line passage of text was presented in their entirety with the target in bold font. Participants were instructed to read the passage and to circle on a 7-point scale, with ‘1’ as ‘low predictable’ and ‘7’ as ‘high predictable, how predictable the target was once they had read the preceding context just prior to the target. Even though the experimenter arranged targets in one of six experimental conditions according to their intuition, participants were presented with a random order of passages.

Results showed significant main effects of frequency and of predictability and a nonsignificant interaction of the two variables. The significant main effect of frequency showed that high frequency targets were rated as more predictable than low frequency ones. The numeric difference in each frequency condition were so small (5.52 for high frequency targets and 5.36 for low frequency ones) that significance was not expected. However, it is worth thinking why this might have been the case. It could be that when words are higher in frequency, they are also more predictable because a high frequency score indicates that the reader has encountered the given word many times (i.e., more easily accessed from the lexicon) and that word is therefore likely to be more easily predictable in a given piece of text. In this particular case, since the same materials were used to administer a Cloze task to participants, and results in this latter experiment showed the desired nonsignificant main effect of frequency, it was decided that the final set of 150 targets in their given contexts was acceptable (see pre-test 2).

In line with expectations was a significant main effect of predictability. This result is desirable since it shows that the three predictability conditions of LP, MP and HP significantly differed in their ratings. Specifically these ratings respectively were 4.37, 5.7 and 6.24 – thus all were in the desired direction where the higher the predictability condition, the higher the rating score given to the targets. Given the importance of having properly defined predictability categories, this result indicates the high validity of our predictability variable.

Also in line with expectations was that the interaction of frequency x predictability was nonsignificant. This indicates equality of rating scores between respective conditions. Since the main effect of frequency had shown to be significant, it was decided that an examination of the frequency contrasts at each of the predictability conditions was warranted. An inspection of these follow-ups showed that at low, medium and high predictable conditions, there were nonsignificant differences in the rating score between low and high frequency targets. We would not expect there to be significant differences between low and high frequency targets at low, medium and high predictable conditions because the design of this experiment was to create three predictability conditions that were clearly delineated from each other. That is, low predictability ratings – for low and high frequency targets - were in a narrow range, as was medium and high predictability ratings. Thus, an eye-balling of the data shows that low predictable targets (when considering both low and high frequency targets) were rated between 3.17 and 5.94. Similarly low and high frequency medium predictable targets were rated between 3.72 and 6.56. Finally, low and high frequency high predictability targets were rated between 5.50 and 6.72. These results indicate that in each of the three predictability conditions, ratings were matched across low and high frequency targets.

Since we examined frequency contrasts, a further examination of predictability contrasts, for high frequency and for low frequency targets, was also implemented. Results here were all as expected where all six predictability contrasts were significant, and most were very strongly so, as indicated by very high *F* ratios. Specifically, these comparisons examined three predictability contrasts for high frequency targets as well as for low frequency ones. That is, for the high frequency targets, high predictable targets were rated as significantly more predictable than both medium and low predictable ones as well as medium predictable targets being rated as significantly more predictable than low predictable ones. Results were in the same direction for the low frequency targets with high predictable targets being rated as significantly more predictable than medium and low predictable ones and medium predictable ones as significantly more predictable than low predictable ones.

3.3.1 Conclusions

The purpose of the rating task was to present participants with a total of 180 passages of text (results are reported for the final chosen 150 passages). In this task, the target was presented in bold font participants were instructed to rate, on a 7-point scale, how predictable they thought the target was given the context prior to the target. We expected a nonsignificant main effect of frequency which would indicate that ratings were matched on low and high frequency targets. We also expected a significant main effect of predictability which would indicate that targets in the high predictable condition contained targets with a higher rating score (indicating that the target was predictable in the context), than those in medium and low predictable conditions – the latter two of which would also significantly differ from each other. Finally, the interaction of the two variables frequency and predictability was expected to be nonsignificant, which would indicate that we had controlled the rating scores between respective conditions.

Results overall were as expected: there was a significant main effect of predictability and a nonsignificant interaction. Despite small numeric differences in rating score for high and low frequency targets, the main effect of frequency was significant. However, an examination of the frequency contrasts (for the interaction of frequency and predictability) showed that effects were in the desired direction, where all three frequency contrasts comparing low and high frequency ratings scores at respectively low, medium and high predictability were nonsignificant. These results indicate that for low, medium and high predictability conditions respectively, ratings scores across the frequency conditions were matched.

Pre-test 2 presents the Cloze task which was also set up so that participants read the same 180 passages of text as those in the rating task. The overall aim was that the rating task could be used in conjunction with the Cloze task as a manipulation check of our variables.

Pre-test 2

Cloze task

3.4 Method

3.4.1 Participants

Twenty participants took part in the Cloze task (14 female, 6 male). All participants were Native English speakers and attending the University of Glasgow as undergraduate students (mean age 20.4 years). After participation, two course credits were given. Care was taken to specifically ensure that no two same participants took part in the rating task as well as the eye tracking experiment. Participants did not have any known learning disabilities.

3.4.2 Apparatus

The Cloze task was administered as a pencil-paper task. Passages were typed double-sided onto A4 pages and stapled together to form a booklet. The instructions were typed on the first page of the booklet. A consent form was also typed and attached to the booklet.

3.4.3 Materials

The passages were presented up to, but not including, either the target or the post-target context. The instruction sheet told participants to write down what they thought the next (single) word in the passage could be given the sentence frame up to that point. A blank underlined space was given in order to do this. In addition participants were instructed to ensure that they read the whole passage before filling in the ‘blank’ and also that this word, as a continuation of the sentence, would not be the last word in the passage. Figure 3.2. below shows the presentation of 12 example passages; two from each condition (these passages were later used). Conditions are labelled in this Figure only; participants were presented with a random order of passages with no reference to which condition the experimenter thought it should be in. Participants were presented with the same 180 passages of text as in the rating task and after scoring both this and the rating task, 150 final passages were selected.

Figure 3.2. Example presentation used in the Cloze task

[Condition 1 LF-LP]

Arthur was at home, preparing vegetables to accompany his dinner.

He steamed some _______________

Overnight, vandals had ruined my prized cherry blossom tree.

They had ripped off the _______________

[Condition 2 LF-MP]

Colin couldn't resist the advances of the sexy new secretary.

He was overcome with _______________

The critically acclaimed restaurant was fully booked once again.

They had hired a talented _______________

[Condition 3 LF-HP]

Martin lost his temper and broke his wife's favourite ornament in two.

He would need to find some _______________

Jimmy ran down the hill, gripping the string in the strong winds.

He loved to play with his _______________

[Condition 4 HF-LP]

The zookeepers were busy preparing for their latest arrival.

They were getting a baby _______________

The sales team were pushing hard as the end of the month loomed.

They had been set a certain _______________

[Condition 5 HF-MP]

My favourite hobby is going to see musicals at the theatre.

I normally pay extra so that my _______________

I struggled to read the badly printed manual for my new computer.

It had little space between the lines of _______________

[Condition 6 HF-HP]

Gary had just started a new job helping tidy up the barber's shop.

He had to sweep up the

When Alex arrived at his friend's house, he rang the bell.

He heard footsteps behind the _______________

3.4.4 Design

The design of the experiment was the same as in the later eye tracking experiment: passages of text were constructed using a within-subjects 2 x 3 design: frequency (LF, HF) on contextual predictability (LP, MP, HP). The dependent variable was the word written in the blank space. Responses were scored as "1" if the correct target word was written down and "0" for any other word given. The final set of chosen 150 passages is displayed in Appendix C; Appendix D shows Cloze values (that is, the probability of guessing the correct target word in the context) generated for each of the 150 target words. The mean Cloze value (i.e., the mean probability) across the six experimental conditions are displayed in column 7 in Table 3.1 on p.101.

3.4.5 Procedure

Participants were emailed from a pool of participants who had previously signed up to take part in language experiments in the School of Psychology. When they arrived at their pre-agreed day and time, they were taken to a quiet booth in a lab. They were given the instructions, consent form and the task itself. After reading the instructions and consent form, they were given the chance to ask questions if they needed to. When they were ready to start the task, the experimenter left the room. The experiment took approximately 50 minutes to complete. After successful completion, participants were debriefed and given their two course credits.

3.5 Results

Across the 8 conditions (LF-LP; LF-MP; LF-HP; HF-LP; HF-MP and HF-HP), the mean Cloze values respectively were, 0.01 (0.02); 0.52 (0.17); 0.96 (0.05); 0.01 (0.02); 0.56 (0.16) and 0.97 (0.04). Results were carried out for items analysis (F_2) in the ANOVA. Bonferroni multiple comparisons were used for the follow-up tests.

The results showed that the main effect of frequency was nonsignificant [$F_2<1$]. The Cloze probability for HF targets was 0.51 (SD=0.04) and for LF targets 0.50 (SD=0.04). This means that there were no significant differences in Cloze scores obtained for HF targets and for LF targets. The predictability main effect was highly significant [$F_2(2,48)=1639.1$, $MSE=.007$, $p<.001$]. Follow-ups to this significant main effect showed that all three predictability conditions significantly differed from each other. Specifically, targets in the HP condition had a significantly higher Cloze probability than targets in the MP condition (0.96 versus 0.54: $F_2=638.26$, $p<.001$) and also than those in the LP condition (0.96 versus 0.01: $F_2=3263.67$, $p<.001$). Also highly significant was the difference in Cloze probability obtained for MP and LP targets (0.54 versus 0.01: $F_2=1015.36$, $p<.001$). As expected, the interaction of frequency x predictability was nonsignificant [$F_2<1$].

3.5.1 Results summary

The results obtained in the Cloze task are summarised in Tables 3.6 and 3.7 below. Results were as expected: there was a nonsignificant main effect of frequency, a significant main effect of predictability and a nonsignificant interaction between frequency and predictability. Follow-ups to the main effect of predictability showed that all three contrasts were in the desired direction and all were highly significant: high predictable targets had a significantly higher Cloze probability than both medium and low predictable targets and medium predictable targets had a significantly higher Cloze probability than low predictable targets.

Table 3.6 Summary of ANOVA Results in Cloze Task by Items (F_2)

Main effect frequency					
Measure		***df***	***F***	***MSE***	***p***
Cloze probability	F_2	1,24	*<1*	.001	*ns*
Main effect contextual predictability					
Cloze probability	F_2	2,48	1639.1	.007	<.001
Interaction word frequency x contextual predictability					
Cloze probability	F_2	2,48	*<1*	.012	*ns*

Table 3.7 Follow up Contrasts to Significant Main Effect of Contextual Predictability by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***
Cloze probability	F_2	LP < MP		LP < HP		MP < HP	
		1015.36	<.001	3263.67	<.001	638.26	<.001

Note: LP = low predictability; MP = medium predictability; HP = high predictability

3.6 Discussion

The Cloze task was used to determine what the next word in the passage of text should be, given the context in a sentence frame. The pencil-and-paper method was used to administer the task to a group of participants who did not take part in the rating task or the later eye tracking experiment. The passages were presented up to, but not including, the target or the post-target context. Participants were instructed to write in one word on the underlined space which they thought best fitted in with the context of the passage. Participants were also told to make sure that they had read the whole passage carefully up to the underlined space. They were also made aware that their word would be a continuation of the passage rather than the last word in the passage. The order of passages of text was arranged randomly with no reference made to the experimenter’s intuition as to what condition each passage best fitted.

Results were expected where there was a nonsignificant main effect of frequency, a highly significant main effect of predictability and a nonsignificant interaction of frequency and predictability. The nonsignificant main effect of frequency showed that there were no significant differences in Cloze probability obtained for low and high frequency targets. This result indicates that the Cloze probability was matched across low and high frequency targets. Also as expected was a significant main effect of predictability and results showed this was a very large effect, as indicated by the size of the *F* ratio. Follow-ups to this significant main effect showed that all three predictability conditions were highly significant from each other. That is, the Cloze probability obtained for high predictable targets was significantly higher than the Cloze probability for medium predictable targets and both were significantly higher than Cloze probability for low predictable targets. These follow-ups serve as a manipulation check for the experimenter's intuition as to whether to classify passages of text as low, medium and high predictable contexts. Therefore, these highly significant contrasts indicate the high validity of the predictability variable.

As expected, the interaction of frequency by predictability was nonsignificant. This result indicates that numerical differences between comparable conditions were equally matched and such suggests that we adequately controlled for Cloze probability between respective conditions. Thus, there is strong evidence of the strong validity of our predictability variable, with levels of low, medium and high contexts manipulated with low and high frequency target words.

3.6.1 Conclusions

In the Cloze task, participants were given a total of 180 passages of text as in pre-test 1. For the Cloze task, they were asked to generate the next word in the passage. Results are reported for the final chosen 150 passages of text which were chosen to best fit one of the six experimental conditions. We expected a nonsignificant main effect of frequency which would indicate that Cloze probability was equally matched over low and frequency targets. Also expected was a significant main effect of predictability which would indicate that targets in the high predictable condition contained targets with a higher Cloze probability than those in medium and low predictable conditions – the latter two of which would also significantly differ from each other. Finally, a nonsignificant interaction between frequency and predictability serves to indicate that we had controlled the Cloze probability between respective conditions.

Results were all as expected: there was a nonsignificant main effect of frequency, a significant main effect of predictability and a nonsignificant interaction of the two variables. Thus results suggest that Cloze probability was matched for low and high frequency target. Also shown was that the three predictability conditions significantly differed from each other, with the numeric differences indicating that high predictable targets had a significantly higher Cloze than both medium and low predictable targets and that medium predictable targets had a significantly higher Cloze probability than low predictable targets. These results suggest that our manipulation of the predictability variable was as desired. Finally the nonsignificant interaction of frequency and predictability indicates that for low, medium and high predictability conditions respectively, Cloze probabilities across the frequency conditions were matched. Also, numeric differences between predictability conditions for high frequency and for low frequency targets were equally matched.

Once word ratings and Cloze probabilities had been calculated, passages of text with low and high frequency targets were validated as belonging to one of six conditions, either LF-LP, LF-MP, LF-HP, HF-LP, HF-MP or HF-HP (see Appendices C and D for full set of materials and target word characteristics).

Experiment 2

3.7 Introduction

How long readers look at a word is influenced by the ease or difficulty associated with accessing the meaning of the word. Two factors, word frequency and contextual predictability effects, have been reliably demonstrated across a variety of measures – lexical decision reaction times (RTs), eye fixation durations, and event-related potentials (ERPs). However, previous research has been inconsistent as to whether these factors, when examined simultaneously, are additive or interactive. Behavioural RT studies have typically demonstrated interactive effects (e.g., Stanovich & West, 1983), with a greater predictability difference for low frequency (LF) than for high frequency (HF) words. Sereno, Brewer, and O'Donnell (2003) obtained a similar pattern of effects in their ERP voltage amplitude data. In contrast, eye movement reading studies have shown additive effects of frequency and predictability on fixation time measures (e.g., Rayner, Ashby, Pollatsek, & Reichle, 2004) - except for word skipping which showed interactive results in Rayner et al. (2004). More recently, Hand, Miellet, O'Donnell, and Sereno (2010) also found additive fixation time effects. However, when launch distance to the target (used as a metric of parafoveal preview) was additionally considered as a factor, an interactive frequency-predictability effect emerged. Word skipping was also interactive. Whether frequency-predictability effects are additive or interactive has implications for models of word recognition (modular, Fodor, 1983 vs. interactive, Morton, 1969 models) as well as for models of eye movement control (serial (E-Z Reader; Rayner et al., 2004) vs. parallel (SWIFT; Engbert, Nuthmann, Richter, & Kliegl, 2005) allocation of attention).

From the eye movement literature, the general finding is that the more frequent the word in the language, the less time readers spend fixating that higher frequency word (HF) compared to words which are less frequent in the language (LF; low frequency) (Hand et al., 2010; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kliegl, Grabner, Rolfs & Engbert, 2004; Kliegl, Nuthmann, & Engbert, 2006; Kliegl, Olson, & Davidison, 1982; Rayner, 1977; Rayner et al., 2004; Rayner & Duffy, 1986; Rayner & Raney, 1996; Rayner, Sereno, & Raney, 1996; Schilling, Rayner & Chumbley, 1998; Sereno, O'Donnell & Rayner, 2006; Sereno, Pacht & Rayner, 1992; Sereno & Rayner, 2000; Slattery, Pollatsek & Rayner, 2007). A word's frequency refers to its frequency of occurrence in a language. This is obtained from large-scale corpus counts (e.g., the British National Corpus; BNC, http://natcorp.ox.ac.uk).

Similarly, the general consensus from eye movement studies is that readers spend less time fixating words which are highly predictable (HP) from their prior context (i.e., words which are constrained by the prior context) compared to words which are low predictable (LP) from the previous text (i.e., words which are not constrained by the prior context) (Balota, Pollatsek & Rayner, 1985; Carroll & Slowiaczek, 1986; Ehrlich & Rayner, 1981; Hand, Miellet, O'Donnell, & Sereno, 2009; Kliegl et al., 2004, 2006; Lavigne, Vitu, & d'Ydewalle, 2000; McDonald & Shillcock, 2003a, 2003b; Morris, 1994; O'Regan, 1979; Rayner et al., 2004; Rayner & Well, 1996; Schustack, Ehrlich, & Rayner, 1987; Zola, 1984; but cf. Hyona, 1993). In addition, HP words are more likely to be skipped than LP words (Altarriba et al., 1996; Balota et al., 1985; Brysbaert & Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Ehrlich & Rayner, 1981; Hand et al., 2009; Rayner et al., 2004; Rayner & Well, 1996).

Since predictability refers to the likelihood of a given word following a particular sentence context fragment, the effect of predictability is also referred to as contextual constraint or contextual predictability. When a prior context is highly constrained or predictable, very few words can complete the context fragment; when the prior context is low constrained or predictable, many words are able to complete that fragment. Therefore, the typical way to assess the predictability of a given word is to present participants with a modified version of the Cloze task (Taylor, 1953) in which participants are asked to write in what word they think best fits in with the sentence fragment they have read. This method allows the experimenter to calculate the likelihood of a certain word appearing in a particular sentence context. Experimenters can also administer a second task, the rating task, sometimes referred to as the predictability task. In this task, participants are presented with the entire sentence fragment with the target word highlighted. They are instructed to rate, typically on a 7-point scale, how well they think the highlighted target word fits in with the fragment they have read prior to the target. Generally, results from the rating task (which may or may not be carried out) are used to confirm results from the Cloze task since the latter is typically solely used to determine the contextual predictability of a target.

The locus of contextual predictability is unresolved. That is, does this variable, representing top-down semantic information, affect early, lexical processing such as lexical access and identification? The implication is that sentence context can speed up the process of lexical access, for example by limiting the word candidates that are chosen. Alternatively, does sentence context only affect later post-lexical stages of processing? The suggestion here is that sentence context is used to aid the semantic integration of an already-accessed lexical item into the meaning of the on-going discourse being built up by the reader.

One way in which this question has been researched is to examine whether contextual predictability interacts with word frequency (e.g., Sternberg, 1969b), specifically whether the combined effects of a word's frequency and its predictability from a prior context are additive or interactive. The reasoning is that the presence of the word frequency effect (e.g., in eye movement studies, longer fixations on LF words compared to HF words), indicates that lexical access has taken place – the visual stimulus has been recognised and all the stored information – orthographic and phonological, and controversially semantic - about the lexical item becomes available to the reader (Balota; 1990; Balota and Chumbley, 1984; Sereno & Rayner, 2003). There is plentiful support for this position; eye movement and electrophysiological measures have shown frequency effects 'early' in processing. For example, Sereno and Rayner (2000) showed that HF words were read faster than LF words, with this difference appearing on the first fixation on the word. In studies recording ERPS, many findings have reported the word frequency effect in 'early' ERP components (i.e., those which occur before 200 ms). For example, word frequency effects have been reliably demonstrated in the N1 component beginning at around 130 ms post-stimulus (Hauk & Pulvermüller, 2004; Scott, O'Donnell, Leuthold, and Sereno, 2010; Sereno, Brewer, & O'Donnell, 2003; Sereno, Rayner, & Posner, 1998). Other studies have demonstrated the frequency effect even earlier in the P1 component, beginning at around 80 ms (Hauk, Patterson, Woollams, Watling, Pulvermüller, 2006). Thus, since there is reliable evidence that the word frequency effect indexes lexical access, by examining word frequency with contextual predictability, the locus of this latter variable can be examined. In particular, the motivation for this experiment was to examine whether there is a frequency and contextual predictability interaction in early processing, for example in first fixation. Such an interaction in an early measure would indicate that both variables share the same processing stage and thus suggest that semantic information is made available during the early stages of word identification. On the other hand, additive effects of frequency and contextual predictability would suggest that the two variables occur at different stages of processing. Such a finding would suggest that orthography and phonology are sufficient for word recognition or lexical access to take place, with semantic information following lexical selection (Forster & Hector, 2002).

3.7.1 Behavioural investigations of frequency by predictability

Behavioural RT experiments have examined the simultaneous effects of frequency and contextual predictability and the tasks employed were lexical decision and naming (Stanovich & West, 1979, 1981, 1983; West & Stanovich, 1978, 1982). The finding from these early studies was an interactive pattern of results (however, initially a small number of RT studies emerged which had shown an additive pattern of results, see Fischler & Bloom, 1979; Forster, 1981; Schuberth & Eimas, 1977). However, subsequent RT studies contradicted these early additive findings, and thus of the suggestion of autonomous lexical processing. For example, in naming studies (Stanovich & West, 1981, 1983), participants read sentences and named the final word of the sentence aloud, as quickly as possible. These researchers examined pronunciation latencies on these end-of-sentence HF and LF words. The results showed that a word was named faster when it appeared in a congruent sentence context compared to when it appeared in either neutral or incongruent sentence contexts. That is, they reported main effects of frequency and predictability, and a significant interaction between the two variables, in which LF words were facilitated more by a predictable context than HF words. An interactive pattern of results was also shown in a lexical decision task (West & Stanovich, 1982) and in earlier naming tasks (Stanovich & West, 1979; West & Stanovich, 1978). Thus, there is strong support for an interactive view of lexical processing. However, there are aspects of this research that indicates a degree of caution regarding the extent to which such findings can be generalised to *normal* reading. That is, even though the naming task is considered a good measure of lexical access, this technique is limited in terms of investigating sentence context effects. The naming task makes use of a number of presentation methods (self-paced, rapid-serial visual presentation or RSVP). Such methods vary in amount of information available to the participant at any one point (e.g., word-by-word, phrase-by-phrase etc.) and thus disrupt the flow of reading as well as relying on a secondary task to assess reading behaviour. It is possible then that this confounds the dependent variable measure and does not provide a reliable picture of the relationship between sentence context and lexical access.

A second methodological flaw concerns both naming and lexical decision techniques. There is often a delay in the time between offset of the prior sentence context and the onset of the target word. During this delay, it is possible that participants have the time to consciously make predictions about the upcoming target word (e.g., Stanovich & West, 1983). Even if there is no time for conscious prediction strategies, the time between the presentation of the prior context and the presentation of the target word is unlike normal reading situations which are uninterrupted with any delays. Another potential problem is that the sentence contexts in some of these studies were relatively short and contained intralexical primes (e.g., Stanovich & West, 1983). Therefore, it is possible that the context effects found could actually have been modulated by associative priming from the other words in the sentence rather than top-down effects from higher order levels of representation. Yet another possible confound is where the target word was the last word in the sentence (e.g., Stanovich & West, 1981). Subsequent research has demonstrated that the last fixation on a line is approximately 5-7 letters from the end of the line (Rayner, 1998). Therefore if the target is the last word in the sentence, readers could end up not reading it. Finally, LP conditions tended to be somewhat anomalous (i.e., using target words in contexts which they would normally never occur – given what we know of the meaning of the word) (e.g., Stanovich & West, 1983).

3.7.2 Eye movement investigations of frequency by predictability

The technique of monitoring eye movements while participants are reading for comprehension is considered a method which approximates natural reading extremely well (e.g., Rayner, 1998, 2009; Rayner & Liversedge, 2004; Sereno, 1992; Sereno & Rayner, 2003). For example, it is of course necessary to make eye movements while reading and therefore this method takes advantage of what is a natural phenomenon. In addition, this method is unobtrusive such that measuring participants' eye movements while they read does not disturb their normal reading rate. As such there is no need for a secondary task to make inferences about comprehension. Instead, eye movement measures (fixation durations) tell us where and for how long readers look in the text and are therefore hugely informative. The assumption is that lexical processing is to a certain extent reflected in fixation durations, which therefore allow us to make inferences about underlying cognitive processes (e.g., Kliegl et al., 2006; Rayner 1998, 2009). Indeed this technique is described as an on-line method, indicating its ability to tap into comprehension processes at precisely the moment they are occurring (Sereno & Rayner, 2003).

Several measures of eye fixations on the target word are available from the eye movement record, including those that reflect the first pass processing time for the target word i.e., the first reading of the target word before any regressions back to the target word. These measures are the first fixation duration (FFD) on the target word (the duration of the very first fixation on the target before any regressions back to it) and the single fixation duration (SFD) on the target word (the duration of the fixation on a target word when only one fixation was made on the target word). In addition such first-pass measures can be assumed to reflect early lexical access processes, for example these early measures may indicate the initial access to representations of a word's orthography, phonology, or meaning (e.g., Rayner et al., 1989). Gaze duration (GD) is the sum of all fixations on the target word before an eye movement to another word and is likely to be a 'middle' processing measure reflecting processing activities which occur between initial lexical access to later processing activities. Finally, total time (TT) spent on the target word is the aggregate duration of all fixations and regressions on the target word, and since this measure includes second pass time (that is, re-reading) is assumed to be a 'later' processing measure indicating later processing activities. In addition, researchers can also compute the skipping rate, for example the percentage of times the target word was not fixated at all on the first reading of the word.

The predictability effect, which refers to the finding that readers spend less time fixating words which are HP than words which are LP is not disputed. In terms of the eye movement record, this effect has been shown in various eye movement measures, from the early to the later measures. For example, the predictability effect has been shown in FFD (Altarriba, Kroll, Sholl, & Rayner, 1996; Binder, Pollatsek, & Rayner, 1999; Hand et al., 2009; Rayner et al., 2004; Morris, 1994); in SFD (Hand et al., 2009; Lavigne et al., 2000); in GD (Balota et al., 1985; Binder et al., 1999; Hand et al., 2009; Lavigne et al., 2000; Rayner et al., 2004; Rayner & Well, 1996); in TT (Calvo & Meseguer, 2002; Ehrlich & Rayner, 1981; Hand et al., 2009; Rayner & Well, 1996). In terms of skipping rate, many studies have shown that HP words were skipped more than LP words (Altarriba et al., 1996; Balota et al., 1985; Brysbaert & Vitu, 1998; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Ehrlich & Rayner, 1981; Hand et al., 2009; Rayner et al., 2004; Rayner & Well, 1996).

One early eye movement reading study, Inhoff (1984), examined the combined effects of frequency and predictability. Predictability was defined by a rating task. Inhoff found, just like in the early RT studies, an interaction in gaze duration (the sum of all fixations on a word prior to moving to another word). However, Inhoff's results represent the combined data from normal reading and a condition in which a three-character foveal mask moved with the eyes and slowed down fixation times. In addition, the experimental passages used in this study were excerpts from *Alice in Wonderland*, and thus the word length of the target words chosen from these passages was not able to be formally controlled (low frequency words tend to be longer in length than high frequency words).

More recently, a small number of eye movement studies have appeared in the literature. However, in four of these studies, the frequency by predictability interaction was not the actual focus of the study (Altaribba, Kroll, Sholl & Rayner, 1996; Ashby, Rayner & Clifton, 2005; Hand, Sereno, & O'Donnell, 2012; Lavigne, Vitu, & d'Ydewalle, 2000; Rayner, Binder, Ashby & Pollatsek, 2001). All of these studies included a manipulation of frequency and predictability of targets words embedded in sentences and in all of these studies, predictability was defined by a Cloze task. The results from these studies were main effects of frequency and of predictability on fixation times but no interaction. The focus of Lavigne et al. and Rayner et al. was in examining the effects of predictability on the eyes' initial landing position in words and Altarriba et al.'s focus was Spanish-English bilinguals. Ashby et al. (2000) compared highly skilled readers with average skilled readers and even though they found differential effects of frequency and predictability between the two groups, there was no frequency by predictability interaction. The focus of Hand et al. (2012) was how word-initial letters influence lexical access during reading.

In contrast to the above studies, the frequency by predictability interaction was the primary focus of investigation in a German study (Kliegl, Grabner, Rolfs & Engbert, 2004); a French eye movement study (Miellet, Sparrow & Sereno, 2007) and two eye movement studies in English (Rayner, Ashby, Pollatsek & Reichle, 2004; Hand, Miellet, Sereno & O'Donnell, 2010). Again, the overall pattern of results in these three studies, like the studies above, showed an additive relationship between frequency and contextual predictability with two exceptions: Rayner et al. reported an interaction in PrF (i.e., probability of skipping a target word) and Hand et al. found both additive and interactive effects by additionally examining launch site.

Kliegl et al. (2004) examined the effects of word length, frequency and contextual predictability on several eye movement measures while participants read from the German Potsdam Sentence Corpus (which has 144 sentences; sentence length ranged from 5 to 11 words, with average word length 7.9 words). Kliegl et al. determined the predictability of target words via a Cloze task given to three participant groups (group 1, N=116, 17-19 years; group 2, N=76, 19-38 years; group 3: N = 80, 66-80 years). The participants generated every word in a sentence (the sentences were presented on a computer screen). Participants were asked to guess the first word of a sentence, typing in their guess. The computer then displayed the (correct) first word of the sentence and the participant then guessed the second word in the sentence and so on, until the end of the sentence was reached. Kliegl et al. used a method called *logit predictability* in which the Cloze values they generated were converted into log values. There were five groupings of predictability; the upper boundaries of each were 4%, 12%, 27%, 50% and 100%. Results from the eye movement data showed independent effects of word length, frequency and predictability on FFD, SFD and GD (i.e., measures which do not include regressions back to the target word). Their results showed that there was a nonsignificant tendency of predictability when the effects of word length and frequency were controlled. However, in contrast, predictability was significant in the eye movement measures which include regressions (i.e., total fixation time). In addition, Kliegl et al. also examined the effects of word length, frequency and contextual predictability on a subset of target words from the corpus. In this case, there were significant predictability effects in the SFD and GD measures. Kliegl et al. argued that *a priori* selection of target words gave a benefit to the reliability of predictability effects in SFD and GD measures (i.e., measures of first-pass reading).

Miellet et al. (2007) recorded participant's (N=15) eye movements whilst they read meaningful short stories (in French) of 134 words; 20 target words were selected from this passage and divided into four conditions (LF-LP, LF-HP and HF-LP, HF-HP) such that there were five items per condition (word length did not differ significantly). The predictability of target words was determined by administering a Cloze task to 20 participants. This task was administered on a word by word basis, that is, participants were asked to guess the next word in the passage (if this was not 'correct', they were told the correct word) and they then guessed the next word, and so on until the whole passage had been presented. For HP targets, the Cloze probability was 71% (where all probabilities were greater than 50%) and for LP targets, the Cloze probability was 4% (where all probabilities were less than 20%). In terms of their fixation time measures since results in the SFD and FFD measures were similar, they reported results on FFD and GD measures. Thus there was a main effect of frequency (in FFD and in GD; longer fixations on LF than on HF words) and a main effect of predictability (in FFD only; longer fixations on LP than on HP words). The interaction was not significant in any of the measures supporting an additive, late processing account. In addition, they calculated probability of skipping a target word but found no significant differences between conditions. They accounted for their results by extending the parameters of the E-Z Reader additive version 7 model (Rayner et al. 2004) of eye movement control (Reichle, Rayner, & Pollatsek, 2003).

Rayner et al. (2004) investigated the effects of frequency and predictability of target words embedded in single-line sentences. Participants read 32 single-line sentences containing target words varying in frequency and predictability. The target word was either

HF or LF and either predictable or unpredictable from its prior context (four conditions: LF-LP, LF-HP and HF-LP, HF-HP). That is, there were two possible target words for each of the 32 passages; participants read one version. Predictability was determined by the Cloze task as well as a separately conducted rating task. Participants who took part in either of these tasks did not participate in the eye tracking reading experiment. In the rating task, Rayner et al. presented 20 participants with the sentence frames up to and including either the predictable target word or the unpredictable target word (but not the sentence frame after the target word). Participants were asked to indicate on a scale of 1-7 how well they considered the (target) word to fit into the sentence where 1 meant that they word did not fit very well and 7 meant that the word fitted very well. The mean ratings for the four conditions, LF-LP, LF-HP and HF-LP, HF-HP, were 4.6, 6.3, 4.4, and 6.6 respectively. For the Cloze task, another group of 20 participants generated the next word in the experimental sentences. Participants received the sentence frame up to but not including either the target word or the post-target context. For HP targets, the Cloze probability was 78% and for LP targets, the Cloze probability was less than 1%. In terms of their results, in FFD, SFD and GD measures, results were additive. Specifically, there was a main effect of frequency (longer fixations on LF than on HF words) and a main effect of predictability (longer fixations on LP than on HP words) in all three measures. The interaction was not significant in any of these measures. The authors pointed out that while the interaction was not statistically significant, the fixation time data in all three measures did suggest an interaction. That is, in the unpredictable conditions, there were larger differences in word frequency than in the highly predictable conditions. The TT measure had the same result: main effects of frequency and predictability and a nonsignificant interaction. Rayner et al. also computed how often the target word was skipped. Here, the main effect of frequency was significant in the subjects analysis only; the main effect of predictability was not significant, but the interaction was significant. Follow-up tests showed that there was a significantly higher likelihood of participants skipping a target word when it was both high frequency and highly predictable (i.e., the HF-HP condition) compared to words in any of the other three conditions (which did not significantly differ from each other).

Rayner et al. countered a possible explanation for the different results in their fixation time measures (i.e., additive results) and probability of skipping the target word (i.e., an interaction between the two variables). Specifically, they reasoned that their observed results may be dependent unequally on word of differing lengths. Specifically, shorter words may contribute more to the skipping data whereas longer words to fixation duration data. Therefore, the observed effects may simply be due to shorter and longer word length. In contradiction to this possibility, Rayner et al. put forward that in their study, word length did not span a wide range, since words from 5 to 8 letters in length were used, and more importantly that of the 32 target words, 29 were either 5 or 6 letters in length. Rayner et al. stated that when 5 and 6 letter words were examined, these words did not appear to be supportive of the above word length argument. That is, when both the 5 and 6 letter words skipping rates were examined, for the predictable words there was a large effect of frequency whereas the unpredictable words showed either no effect, or small reversed effect of frequency. However, for the 5 and 6 letter gaze durations, both predictable and unpredictable words showed relatively large effects of frequency.

Rayner et al.'s study has some methodological limitations. First, in each of the four experimental conditions, a participant read only 8 sentences. Experimental results could be due to idiosyncratic aspects of the relatively small number of items used and may not be replicated in using a wider range of materials. Second, materials were single-line sentences with the target word placed in approximately the middle of the sentence. It could be argued that longer contexts prior to reaching the target word allows for context effects to develop more fully.

A recent study by Hand et al. (2010) was conducted to address the limitations of Rayner et al. (2004) mentioned above. In addition, a further focus of the study was to investigate the effect of parafoveal preview in post-hoc analyses. Specifically, Hand et al. examined the frequency and predictability of target words embedded in text while participants' eye movements were recorded. Materials were two-line sentences such that a longer context than that used in Rayner et al. allowed for the context to develop more fully. As in Rayner et al. there were two possible contexts for each target word; the target word was either HF or LP and either predictable or unpredictable from the prior context (the four conditions were LF-LP, LF-HP, and HF-LP, HF-HP). Number of materials was greater than in Rayner et al. also; participants read 88 sentences where there were 22 sentences per condition. Predictability was determined by two tasks: rating and Cloze. In the rating task, 10 participants did each of the two versions of the materials. They were presented with the whole passage of text (including the post-target context, unlike Rayner et al.) with the target word in bold, and asked to indicate on a scale of 1-7 how well they thought the word in bold font fitted in with the sentence context where 1 indicated that the word did not fit well and 7 indicated that it fitted very well. The mean ratings across the four conditions, LF-LP, LF-HP and HF-LP, HF-HP, were 3.69, 6.05, 4.07 and 6.20 respectively. For the Cloze task, a separate group of 20 participants (the full set could be administered because the target word was not given) were asked to write the next word in the sentence. Participants received the sentence frame up to but not including either the target word or the post-target context. For HP targets, the Cloze probability was 0.57 and for LP targets, the Cloze probability was less than 0.02 (in comparison to 0.78 and 0.01 in Rayner et al. 2004). The Cloze probabilities in the four conditions, LF-LP, LF-HP and HF-LP, HF-HP, were 0.02 (SD=0.06); 0.53 (SD=0.31); 0.02 (0.06); and 0.60 (SD=0.31). Hand et al. hypothesised that with the improvements to the materials there may be an interactive pattern of results with larger effects of predictability for LF than for HF words.

However, in Hand et al.'s (2010) study, the results in the FFD, SFD, and GD measures were all additive. That is, in all three measures, there was a main effect of frequency (longer fixations on LF than on HF words) and a main effect of predictability (longer fixations on LP than on HP words) in all three measures. The interaction was not significant in any of the three measures. The authors pointed out that while the interaction was not statistically significant, the fixation time data in all three measures did suggest an interaction. That is, there were larger differences in word frequency in the unpredictable conditions than in the highly predictable conditions. The TT measure also showed an additive result: main effects of frequency and predictability and the interaction was marginally significant in the subjects analysis and nonsignificant in the items analysis. In terms of word skipping, Hand et al. computed the probability of fixating the target word for each condition. Here, the main effect of frequency was significant in the subjects analysis only; the main effect of predictability was not significant, and the interaction was significant (though marginal in items analysis). Follow-up tests showed that participants skipped words in the HF-HP condition more than in any of the other three conditions (which did differ significantly in likelihood of skipping).

A second focus of Hand et al.'s (2010) study was to examine, post-hoc, the additional effects of parafoveal preview benefit which was indexed by launch distance. Their results were similar in FFD, SFD and GD measures. That is, in the three measures, all three main effects were significant in the expected direction. Thus the main effects of frequency were significant (which was longer fixations on LF than on HF words) as was predictability (which was longer fixations on LP than on HP words) as well as preview. Follow-ups to the main effect of preview showed that the closer the launch distance, the shorter the fixation time on the target word. This finding suggests that shorter launch distances gives rise to greater parafoveal previews and as such there are shorter fixation times on the target.

Regarding the interactions in Hand et al.'s (2010) study, 'frequency by predictability' was nonsignificant; 'frequency by preview' and 'predictability by preview' were both significant. Also significant was the three-way interaction, follow-ups to which showed an interactive pattern in the 'near' and 'middle' conditions and an additive pattern in the 'far' condition. Hand et al. suggested the initial additive result of the two-way interaction 'frequency by predictability' was the result of a combination of the interaction in 'frequency by predictability and launch distance' in the 'near' and 'middle' launch distances and an additive pattern of results in the 'far' condition. Their results suggested that at greater launch distances, the less the effect of parafoveal preview.

The results overall of the above frequency by predictability eye movement studies, therefore, provide overwhelming support that contextual predictability has a relatively delayed time course (the exception being the interaction in word skipping reported in both Hand et al., 2010 and in Rayner et al., 2004; however it should be noted that the main effect of predictability was nonsignificant in both studies; and the interactive finding when launch site was examined in Hand et al., 2010). However, one factor of interest regarding these prior eye movement studies is that of the level of predictability. To recap, predictability has been operationally defined as Cloze probability which is the proportion of participants who guess the target word when presented with the text up to but not including the target word. In the studies reviewed above, 'high' predictability targets have been those with Cloze probabilities ranging from the lowest of 0.53 to the highest at 0.94, and 'low' predictability targets with Cloze probabilities from 0.00 to 0.30. However, it could be that a sentence context with a Cloze probability of 0.53 is *moderately* predictable, rather than highly predictable. Similarly, a sentence context with a Cloze probability of 0.30 is moderately rather than low predictable.

Table 3.8 shows, for each study, the average Cloze for HP and LP sentence contexts, and if stated, the range and standard deviation. The range tells us the minimum and maximum Cloze probabilities which were in the complete data set, and the standard deviation gives an indication of the extent to which each Cloze probability number in the data set varies with respect to the mean. However as shown in Table 3.8, all of these prior studies have not reported the range of Cloze probabilities. Three of the studies have reported the standard deviation (Altaribba et al., 1996; Hand et al., 2010; 2012) which gives some indication of the range of Cloze probabilities. For example in Hand et al.'s study, the HP sentence contexts probably had a very large range as indicated by the relatively large standard deviations which possibly suggests that some of the Cloze probabilities at the lower and higher end of the range could have been low and moderate predictable respectively. However, without the range information, it is difficult to determine this definitively.

Two other studies (Kliegl et al., 2004; Miellet et al., 2007) have been somewhat more informative; even though the information about the range of Cloze probabilities was not provide. For example, Miellet et al. stated that all HP Cloze probabilities were greater than 0.50 and that all LP Cloze probabilities were less than 0.02. This suggests a possible range of HP Cloze probabilities of 0.50-1.00 and LP Cloze probabilities of 0-0.02. If this is the case, then the lower end of the high probability Cloze values were moderately predictable contexts. Similarly, Kliegl et al.'s study suggests that the LP Cloze probabilities ranged from 0.00-0.50 and HP from 0.50-1.00. The higher end of LP contexts were moderately predictable and similarly, the lower end of the HP contexts were moderately predictable.

Therefore in past investigations of frequency by predictability (see Table 3.8), with the exception of Lavigne et al., (2000), it is suggested that the additive relationship is between low and *moderately* predictable contexts and not highly predictable. Thus, this additive pattern of results which has been reported in all previous eye movement studies, may be the result of under-sampling at the high end of Cloze probability and an interactive pattern may emerge by the inclusion of high predictable contexts. Table 3.9 presents a modified version of where these 'HP' Cloze probabilities should be placed (i.e., in the MP; medium predictable condition). The present study was conducted to address these issues - by adequately defining Cloze probability conditions using a larger range of predictability values than has been the case so far, a more reliable picture of the relationship between frequency and contextual predictability may be obtained.

Earlier, Rayner and Well (1996) investigated the effect of contextual constraint (predictability) on target words in reading by using three levels of contextual predictability. However, frequency was not a variable in this study. Contextual predictability was defined by a Cloze task; LP (0.04; range 0.03-0.08), MP (0.41; range 0.13-0.68) and HP (0.86; range 0.73-1.00). Table 3.10 shows the Cloze values used in other past eye movement predictability only studies. Rayner and Well did not conduct their own Cloze task, instead they used Cloze probabilities for sentence contexts collected previously by Schwanenflugel (1986).

Schwanenflugel administered a Cloze task of 260 sentences to three separate groups of 40 participants (the 260 sentences were divided unequally among the three participant groups). Of the 260 sentences, 47 had appeared in an earlier norming study by Bloom and Fischler (1980). In Schwanenflugel's task, the target word was always the last word in the sentence. Participants were asked to list up to three (but not more) words that could complete the sentence frame (the sentence frame was presented with the last word of the sentence replaced with three blank lines). Schwanenflugel worked out the Cloze value for each sentence frame by using a method called *multiple production probability.* In this method, the probability of a completion is calculated across participants regardless of the order the target words are written down in; this was calculated by dividing the number of participants writing down the particular target word by 40 participants. Schwanenflugel's paper provided 97 of the 260 sentences that were administered.

Rayner and Well (1996) chose 36 sentences from Schwanenflugel's sentence list. However, since the targets were always sentence final, Rayner and Well added post-target context which meant that the target was now placed in approximately the middle of the sentence frame. The sentence frame was such that two alternative target words were inserted into the frame (participants read only one version) - a third of the sentences contained either a HP or a MP target word, another third contained either a MP or a LP target word and the final third of sentences contained either a HP or LP target word. Word length was not exactly matched: average target word length was 6.0, 5.9 and 6.2 letters for HP, MP and LP respectively.

Rayner and Well reported three fixation time measures: FFD, GD and TT and also reported skipping data. All three measures showed the same pattern of results. Specifically, fixation times on LP targets were longer than those on HP or MP targets which did not differ from each other. However, the skipping data showed that HP targets were skipped more often than MP or LP ones; there was no difference in skipping the target word between the MP and LP target words. In the FFD measure, predictability was significant by subjects only; this effect was nonsignificant in the items analysis. In the GD and TT measure, predictability was significant in both subjects and items analyses.

Unlike, past investigations, Rayner and Well did provide the associated ranges (generated from Schwanenflugel's Cloze task); Table 3.10 shows the Cloze values used in past eye movement predictability studies. Therefore, in Rayner and Well's study, even though the actual Cloze values were sufficient to make a distinction between high, medium and low predictability, when the range of values is examined, it can be seen that, in particular, the HP and MP conditions are close to overlapping: the highest MP Cloze value of 0.68 is very close to the lower range for HP target words 0.73: it could be argued that there was the HP and MP conditions were not differentiated sufficiently well. Indeed, their fixation time data show that that, on their three measures FFD, GD and TT, the HP and MP targets did not differ from each other (only the LP targets yielded longer fixation times than HP and MP targets).

In addition, like Rayner et al.'s (2004) and Miellet et al.'s (2007) studies, in Rayner and Well's study, participants read relatively few sentences in each condition. In Rayner et al.'s study, there were 8 sentences per condition; in Miellet et al.'s study, there were 5 sentences per condition, and in Rayner and Well's study, there were 12 sentences per condition. In using a relatively few number of experimental sentences, results can be dependent on possible idiosyncratic aspects of such small number of items. In addition, in Rayner and Well's study (and in Rayner et al., 2004), the sentences were single line sentences (with the target word appearing around the middle of the sentence frame); however it is unclear whether the sentence frame spanned over two lines of the display the participant viewed during the experiment. A longer pre-target text can provide a more salient context for the target word. Furthermore, in Rayner et al.'s study, some of the contexts used in the sentences were anecdotal in that the context pertaining to the target word relied on the reader knowing certain conventions. Example materials from Rayner and Well can be seen in Table 3.11.

Notes for Table 3.8 and 3.9 overleaf: EM = eye movement; SD = standard deviation; F = frequency; P = predictability; Exp. = experiment; info. = information; LP = low predictability; MP = medium predictability; HP = high predictability; HF = high frequency; LF = low frequency.

*There was no HP condition in Experiment 1.

**Collapsed over constraint.

***Inhoff (1984) used only a rating task.

****Kliegl et al. (2004) used a regression approach

***** Range is inferred

Table 3.8

Average Cloze Values and Associated Range and SD in Past EM Frequency by Predictability Studies

EM frequency by predictability studies finding additive effects		Average LP Cloze Value	Range (if stated)	Any other info.	Average MP Cloze value (if used)	Range (if stated)	Any other info.	Average HP Cloze Value	Range (if stated)	Any other info.
(F by P was not primary focus):										
Altaribba, Kroll, Sholl, & Rayner (1996)	*Exp. 1:*									
	HF condition	0.07		SD=0.12				0.70		SD=0.28
	LF condition	0.03		SD=0.10				0.70		SD=0.32
Ashby, Rayner, & Clifton (2005)	*Exp. 1*	0.05						*		
	Exp. 2	0.01						0.78		
Hand, O'Donnell, & Sereno (2012)**	*LF condition*	0.04						0.64		
	HF condition	0.04						0.62		
Lavigne, Vitu, & d'Ydewalle (2000)	*Exp. 1*	0.03						0.94		
	Exp. 2	0.07						0.86		
Rayner, Binder, Ashby, & Pollatsek (2001)	*Exp. 1*	0.01						0.64		
	Exp. 2	0.01						0.78		
(F by P was primary focus):										
Inhoff (1984)***										
Kliegl, Grabner, Rolfs, & Engbert (2004) ****										
Rayner, Ashby, Pollatsek, & Reichle (2004)	*LF condition*	0.01						0.78		
	HF condition	0.04						0.71		
Miellet, Sparrow, & Sereno (2007) *****			0.00-0.02	<0.02					0.50-1.00	>0.50
Hand, Miellet, Sereno, & O'Donnell (2010)	*LF condition*	0.02		SD=0.06				0.60		SD=0.31
	HF condition	0.02		SD=0.06				0.53		SD=0.31
Present study	*LF condition*	0.01	0.00-0.05	SD=0.02	0.52	0.20-0.75	SD=0.17	0.96	0.85-1.00	SD=0.05
	HF condition	0.01	0.00-0.05	SD=0.02	0.56	0.20-0.75	SD=0.16	0.97	0.90-1.00	SD=0.04

Table 3.9

Modified Version of Average Cloze Values and Associated Range and SD in Past EM Frequency by Predictability Studies

EM frequency by predictability studies finding additive effects		Average LP Cloze Value	Range (if stated)	Any other info.	Average MP Cloze value (if used)	Range (if stated)	Any other info.	Average HP Cloze Value	Range (if stated)	Any other info.
(F by P was not primary focus):										
Altaribba, Kroll, Sholl, & Rayner (1996)	*Exp. 1:*									
	HF condition	0.07		SD=0.12	0.70		SD=0.28	~~0.70~~		~~SD=0.28~~
	LF condition	0.03		SD=0.10	0.70		SD=0.32	~~0.70~~		~~SD=0.32~~
Ashby, Rayner, & Clifton (2005)	*Exp. 1*	0.05						*		
	Exp. 2	0.01						0.78		
Hand, O'Donnell, & Sereno (2012)**	*LF condition*	0.04			0.64		SD=0.20	~~0.64~~		
	HF condition	0.04			0.62		SD=0.29	~~0.62~~		
Lavigne, Vitu, & d'Ydewalle (2000)	*Exp. 1*	0.03						0.94		
	Exp. 2	0.07						0.86		
Rayner, Binder, Ashby, & Pollatsek (2001)	*Exp. 1*	0.01			0.64			~~0.64~~		
	Exp. 2	0.01			0.78			~~0.78~~		
(F by P was primary focus):										
Inhoff (1984)***										
Kliegl, Grabner, Rolfs, & Engbert (2004) ****										
Rayner, Ashby, Pollatsek, & Reichle (2004)	*LF condition*	0.01			0.78			~~0.78~~		
	HF condition	0.04			0.71			~~0.71~~		
Miellet, Sparrow, & Sereno (2007) *****			0.00-0.02	<0.02		0.50-1.00	>0.50		~~0.50-1.00~~	~~>0.50~~
Hand, Miellet, Sereno, & O'Donnell (2010)	*LF condition*	0.02		SD=0.06	0.60		SD=0.31	~~0.60~~		~~SD=0.31~~
	HF condition	0.02		SD=0.06	0.53		SD=0.31	~~0.53~~		~~SD=0.31~~
Present study	*LF condition*	0.01	0.00-0.05	SD=0.02	0.52	0.20-0.75	SD=0.17	0.96	0.85-1.00	SD=0.05
	HF condition	0.01	0.00-0.05	SD=0.02	0.56	0.20-0.75	SD=0.16	0.97	0.90-1.00	SD=0.04

Table 3.10 Average Cloze Values and Associated Range in Past EM Predictability Studies

EM predictability studies		Average LP Cloze Value	Range (if stated)	Any other info.	Average MP Cloze value (if used)	Range (if stated)	Any other info.	Average HP Cloze Value	Range (if stated)	Any other info.
Ehrlich & Rayner (1981)	*Exp.1*	0.15						0.93		
	Exp.2	0.00						0.60		
Zola (1984)		0.08		SD=0.10				0.83		SD=0.13
Balota, Pollatsek, & Rayner (1985)		0.01						0.64		
Hyönä (1993)*		0.32						0.65		
Rayner & Well (1996)		0.04	0.03-0.08		0.41	0.13-0.68		0.86	0.73-1.00	

Note: EM = eye movement; LP = low predictability; HP = high predictability; MP = medium predictability; SD = standard deviation; info. = information

* Hyona (1993) did not find the predictability effect. However, in the Cloze task, the first letter of the target word was given

Table 3.11 Example Materials used in Rayner & Well's (1996) study

<u>HP or MP sentences</u>

He campaigned so he would win the *election|primary* in his home state.

While away, James sent a *letter|package* to keep in touch.

<u>MP or LP sentences</u>

The girl crept slowly toward the *door|edge* without anyone hearing her.

The businessman brought his equipment to play *golf|ball* with his son.

<u>HP or LP sentences</u>

The friends were not talking because they had a *fight|scheme* last semester.

He scraped the cold food from his *plate|spoon* before washing it.

Note: One version of the sentence was presented: HP, MP or LP.
HP = high predictability, MP = medium predictability, LP = low predictability

3.7.3 Present study

Whether frequency and predictability interact remains a debated issue. While studies using techniques like lexical decision and naming (Stanovich & West, 1981, 1983) suggested that these two variables may interact, studies using eye movements have overall reported additive results of these two variables in fixation duration measures (Altarriba et al., 1996; Ashby et al., 2005; Hand et al., 2010; Lavigne et al., 2000; Miellet et al., 2007; Rayner et al., 2004; Rayner et al., 2001). These studies showed an additive pattern of results (main effects of frequency and contextual predictability and no significant interaction between the two variables in any of the fixation duration measures). There are two exceptions. An interaction has been found in the skipping data (Hand et al., 2010; Rayner et al., 2004). In addition, Hand et al. reported a three-way interaction when launch site was additionally examined with frequency and predictability; specifically they showed that when there was a near distance to the target word (within 3 character), there was a larger predictability effect for LF words; when there was a medium distance (4-6 character), the predictability effect was numerically larger for HF words; and there was a far distance, there was no interaction. The question of whether frequency and predictability have an additive or an interactive relationship remains unresolved and needs further investigation.

One factor of interest in all of the previous eye movement studies is that of the level of predictability. Predictability has been operationally defined as Cloze probability (the proportion of subjects who guess the target word when presented with the text up to but not including the target). In addition, the rating task is sometimes used to confirm the results from the Cloze task. In all the past eye movement studies, 'high' and 'low' predicable labels have used Cloze probabilities of around 0.50-0.80 and 0.00-0.30 respectively (Table 3.8 shows for each study the Cloze probabilities used in LP and HP conditions). However, it is possible that the reported additive pattern of results in fixation time measures may be the result of under-sampling the extreme ends of the full range of Cloze probability values, and that an interactive pattern may emerge given a larger differentiation in Cloze probabilities between predictability conditions. Although Rayner and Well (1996) used three levels of predictability, they did not manipulate word frequency. In particular, in the present study the Cloze probability values for the 'high' and 'low' predictability conditions were more extreme than has been used in previous studies: HP = 0.96 (0.85-1.00); MP = 0.54 (0.20-0.75); and LP = 0.01 (0.00-0.05). In the current experiment two norming tasks were administered to obtain the predictability of target words. In addition, target words were embedded in the second of two-sentence text passages (e.g., Hand et al., 2010). Thus, there was a considerable context preceding the target word compared to the single-line passages used by Rayner and Well (1996) and in Rayner et al. (2004). Therefore, the aim of the present study was to extend Rayner and Well by using three levels of predictability but additionally include word frequency as a variable. The present study used a 2 (frequency: LF, HF) x 3 (predictability: LP, MP, HP) design with 25 items per condition. We investigated whether orthogonally manipulating the frequency and predictability of target words presented in short passages of text produced additive or interactive effects on several eye movement measures: FFD, SFD, GD, TT as well as the likelihood of skipping the target word (as indicated by percentage of times the target word was not fixated in the first pass). It was predicted that in the 'early' measures (likely to reflect early lexical processes), FFD and SFD, frequency and predictability would have an interactive relationship where predictability would more beneficial for lower frequency words than high frequency words. We expected larger predictability effects (where lower predictable words are fixated for longer than higher predictable words) for LF words than for HF words. That is, LF words will be fixated less when these words appear in more predictable conditions. In other words, context would boost LF words more than HF words. In terms of skipping rate (as indicated by percentage of words which were skipped in each condition), we predicted an interactive relationship with more skipping for HF-HP words. Since the middle measure GD is likely to indicate processing which takes place between initial lexical access to later processing activities, and TT later processing activities, in these measures we expected an additive pattern of results (suggesting that frequency and predictability are occurring at different stages of processing), that is main effects of frequency and predictability and a nonsignificant interaction.

3.8 Method

3.8.1 Participants

Forty (28 female, 12 male) native English speaking University of Glasgow undergraduate and postgraduate students (mean age 23 years) participated voluntarily. Each participant received either the normal two course credits or payment of £6. All participants had normal or corrected-to-normal vision and had not been diagnosed with any reading disorder. The predictability of target words was determined via two norming tasks – the Cloze task and the rating task each of which was undertaken by two separate groups of 20 participants (see pre-tests 1 and 2). In addition, participants who took part in either of these norming tasks did not partake in the main eye tracking experiment.

3.8.2 Apparatus

Participants' eye movements were recorded with an EyeLink® 2K eye tracker manufactured by SR Research Ltd. (Mississauga, Ontario, Canada). The sampling rate was 1000 Hz using corneal reflection and pupil tracking and the spatial resolution was 0.01˚. The gaze tracking range is 32˚ horizontally and 25˚ vertically. The passages of text were presented in 14-point Bitstream Vera Sans Mono font (black characters on a white background) on a Dell P1130 19" flat screen CRT (1024 x 768 resolution; 100 Hz). This is a non-proportional font (all characters have the same width). A monocular desktop mount incorporated the camera and sat just below the CRT display. At a viewing distance of approximately 72cm, approximately 4 characters equalled 1˚ of visual angle. Viewing was binocular but eye movements were recorded from the right eye. A forehead/chin rest were used to minimise head movements though the eyetracker allows movements of up to 25 mm horizontally and 25 mm vertically.

3.8.3 Materials

The complete set of 150 passages arranged according to their six conditions are presented in Appendix C. Appendix D shows the individual specifications for the 150 targets and Table 3.1 (shown in pre-test 1; rating task) presents a summary of these specifications in the form of means and SD's. Pre-tests 1 and 2 presented earlier contain the details of how the rating and Cloze tasks respectively were administered.

Word frequencies were determined from the British National Corpus (BNC), a database of 90 million written words (http://natcorp.ox.ac.uk), frequency values for LF were occurrences of 17 million or below (range: 0.62-16.22 occurrences per million) and HF were occurrences of 43 million or above (range: 43-611.76 occurrences per million). Across the six conditions, word length was matched (average word length 5.88 characters; range: 4-8 characters; SD = 1.33). As well as this, the number of syllables and word imageability across the 6 conditions did not differ between conditions (number of syllables range: 1-4 syllables; word imageability range: 502-575). Imageability refers to the ease with which a word gives rise to a sensory mental image; this is different from "concreteness" which is defined as the ability to see, hear and touch something (Paivio, Yuille, & Madigan, 1968). Imageability scores for each target word were obtained from five recognised norms – Cortese and Fugett's (2004) imageability norms; the Bristol Norms (Stadthagen-Gonzalez & Davis, 2006; http://language.psy.bris.ac.uk/); Gilhooly and Logie's (1980) imageability norms; Morrison, Chappell, & Ellis's (1997) imageability norms and finally the MRC Psycholinguistic Database (Wilson, 1988; http://www.psych.rl.ac.uk/bristol_norms.html). These norms reflect participants' ratings of words on a scale of 1 (low imageability) to 7 (high imageability).

The level of contextual predictability was obtained via two tasks: the Cloze probability task and the rating task. Participants only took part in one of these tasks and in addition, these participants did not take part in the main eye tracking experiment.

3.8.4 Design

The experiment used a within-subjects 2 x 3 design: the effects of two types of frequency (low frequency, LF; high frequency, HF) on three types of contextual predictability (low, medium and high predictability; LP, MP, HP respectively) were investigated. Dependent variables were standard eye movement measures (FFD, SFD, GD, TT) as well as the percentage of times the target was skipped (%Skip). LF and HF target words appeared in approximately the middle of the second line of short two-line passages. The context in the first line was such that it either made the target word not all predictable (LP); moderately predictable (MP) or very predictable (HP). Therefore, there were six conditions in which sentences were constructed: LF-LP, LF-MP, LF-HP and HF-LP, HF-MP, HF-HP. In each of these conditions, there were an equal number of passages of text; 25 passages of text and hence 25 target words. Each participant read the total number of 150 passages. The first line was such that it contained the majority of the context pertaining to the target word. Care was taken to avoid the use of associative or intralexical primes in the pre-target context (i.e., the second line of the sentence). In addition, since the target word was always placed in approximately the middle of the second line, this meant that the target word was never the first word in the sentence or the final word in the sentence. Finally, the passages were presented in a different random order to each participant.

3.8.5 Procedure

When a participant arrived for the main eye tracking experiment, the forehead/chin rest were adjusted so that they were at the correct height for the participant. The participant was given written and verbal instructions about the reading task. They were instructed to read the passages normally for meaning as they would read a story and told that yes-no questions would appear on a random basis following some of the passages to make sure they were paying attention. The experiment first involved calibration which took about 5 minutes. This is where participants were asked to follow a series of 9 dots around the screen; these calibration points extended over the full horizontal and vertical range over which the passages were presented.

The experiment involved the participant reading the five practice trials followed by the 150 experimental trials. A total of 48 questions appeared randomly, half with yes answers and the other half with no answers. Participants used a joypad to indicate that their answer was ‘no (left-click) or ‘yes’ (right-click). Recalibration was performed whenever necessary, for example, when there was difficulty in the sentence starting when the participant was looking at the black box. This indicated that the calibration was off.

Each passage began when participants fixated the upper left-most calibration point which corresponded to the first character of text – this was a black square box. When they had read the passage, participants looked away to the bottom-right of the screen and pressed a key to clear the screen. The black box re-appeared immediately or, if the passage had had a question, then the participant would answer it first before the black box re-appeared. Yes-No comprehension questions were asked on 40% of the trials. There was no difficulty in answering the questions (average 98% correct). Participants were given a break approximately half way through. The experiment took approximately 50 minutes to complete.

3.9 Results

The target region included the space before the target word and the target word itself. Data were eliminated from analyses for the following reasons: (a) a blink or track loss occurred on the target word; (b) the fixation on the target was either the first or the last fixation on the line; or (c) the duration of any individual fixation shorter than 100 ms or longer than 750 ms. In the study, the percentages of data for single fixation, immediate refixation, and skipping of the target word were 66, 4 and 25% respectively.

The eye movement dependent measures which were analysed on the target word were: (a) first fixation duration (FFD; the duration of the very first fixation on the target word, provided that the word was not skipped, regardless of whether it is the only fixation or the first of many fixations on that word); (b) single fixation duration (SFD; the duration of the first fixation on the target word if it received only one fixation on the first pass); (c) gaze duration (GD; the sum of all fixations on a target word on the first pass prior to making an eye movement to another word); and (d) total time spend fixating the target word (TT; the sum of all fixations on the target word including any re-reading of the target word i.e., regressions back to it). The first two measures can be said to represent first-pass or early measures of processing and therefore may indicate the initial access to representations of a word's orthography, phonology, or meaning. The third measure, GD, is likely to indicate processes that occur between initial lexical access and later processing activities. The fourth measure is likely to reflect those later post-lexical processing activities. Also computed was the likelihood of skipping the target word as indicated by the percentage of skips (%Skip). Two-way within-subjects repeated measures analyses of variance (ANOVAs) were carried out on each measure both by subjects (F_1) and by items (F_2). Where conducted, follow-up tests made use of Bonferroni multiple comparisons corrections.

3.9.1 Frequency by predictability analyses

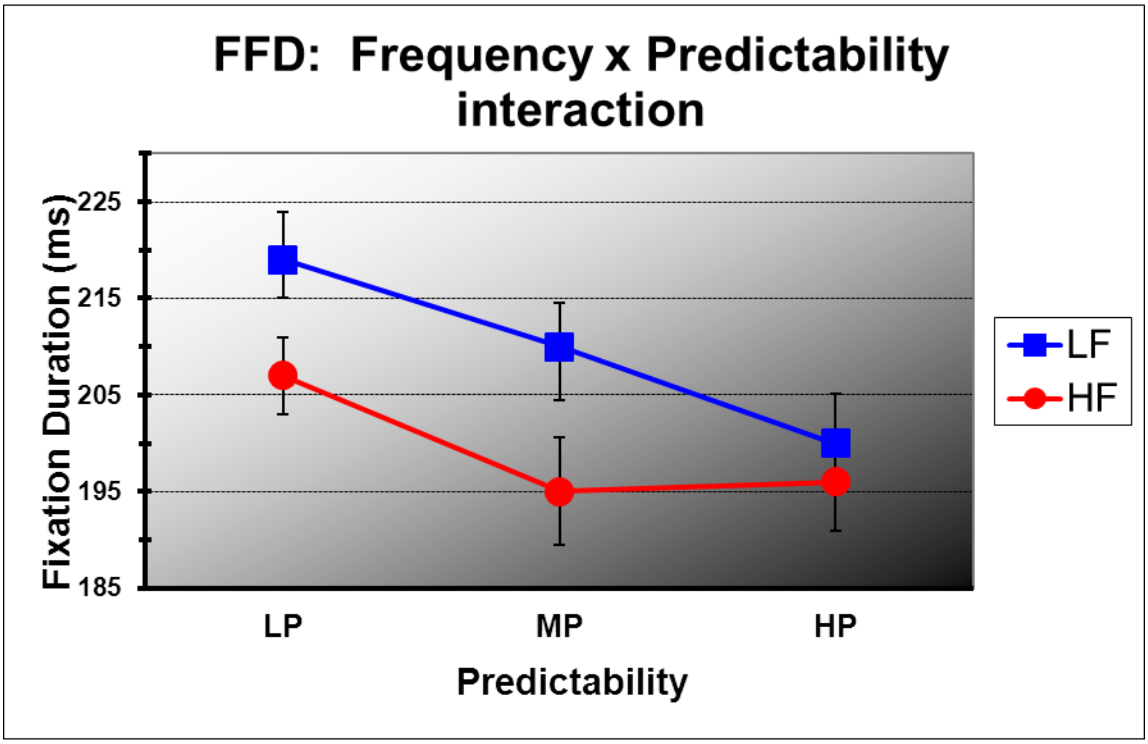
In the sections below, the results for each measure are presented describing main effects (frequency, predictability), follow-up contrasts for predictability (LP vs. MP vs. HP) when it was significant, the interaction (frequency x predictability), as well as follow-up contrasts of the interaction when it was significant. The follow-up tests were Bonferroni multiple comparisons corrections. Mean fixation times across the six experimental conditions are presented in Table 3.12. Means with standard error bars are displayed in Figure 3.3. for first fixation duration and Figure 3.4. for the single fixation duration.

Table 3.12 Mean Fixation Time Results in Milliseconds (Standard Deviation) and Percentage of Times the Target Word was Skipped in the Conditions

Measure		**LF**				**HF**	
	LP	**MP**	**HP**		**LP**	**MP**	**HP**
FFD	219 (28)	210 (27)	200 (26)		207 (26)	195 (26)	196 (23)
SFD	221 (28)	210 (27)	201 (27)		208 (27)	194 (25)	196 (23)
GD	234 (32)	219 (31)	213 (36)		217 (29)	201 (27)	200 (25)
TT	252 (40)	240 (41)	227 (40)		237 (34)	219 (31)	209 (31)
%Skip	21 (12)	24 (15)	25 (14)		25 (13)	28 (14)	29 (16)
Sent.1	28 (4)	29 (5)	29 (5)		29 (5)	29 (5)	29 (5)

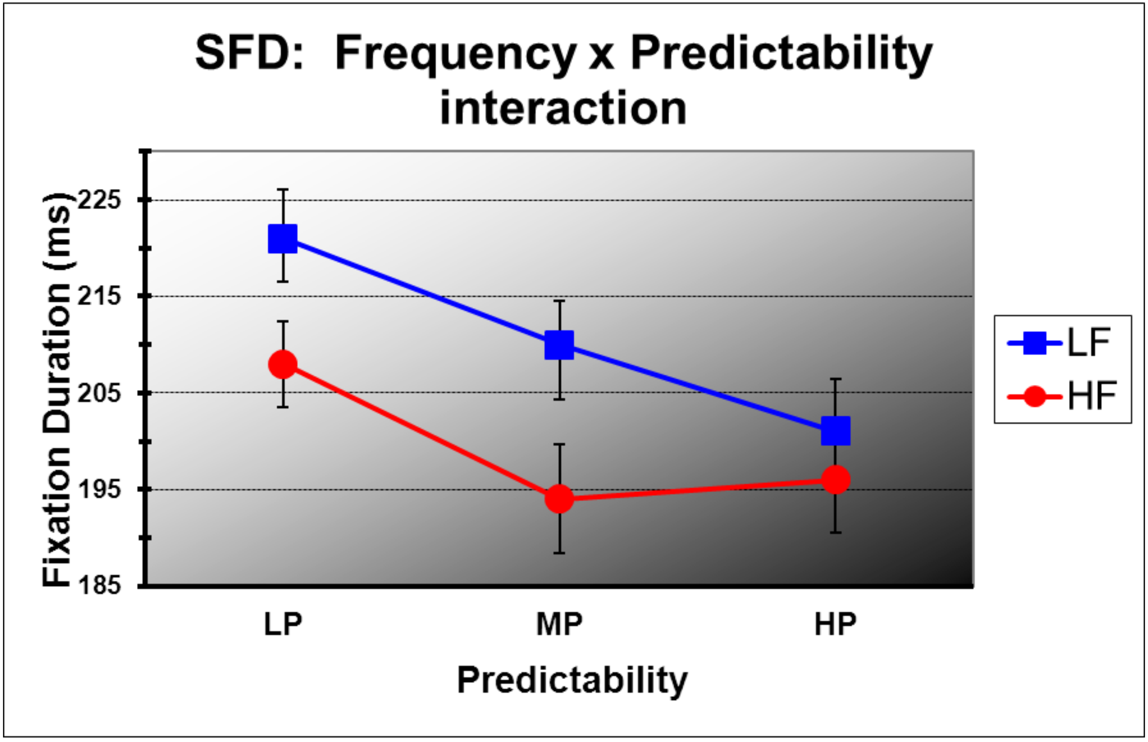
Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; TT = total time including regressions; %Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Figure 3.3. Average first fixation duration (FFD) as a function of frequency and contextual predictability



Notes: ms = milliseconds; SFD = single fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

Figure 3.4. Average single fixation duration (SFD) as a function of frequency and contextual predictability



Notes: ms = milliseconds; SFD = single fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

3.9.1.1 First and single fixation durations

Since the FFD and SFD data showed the same pattern of results (see Table 3.12) these two measures are discussed together. Basically, there were main effects of frequency and predictability and a significant interaction. The follow-up contrasts to the significant interaction also showed the same pattern of results in both measures. Both measures also showed the same numerically unexpected (but statistically nonsignificant) result in one of these follow-ups, which is that HF-MP words were fixated for less time than HF-HP words. The main effect of frequency was significant in the FFD and SFD measures [FFD: $F_1(1,39)=38.34$, $MSE=184$, $p<.001$, $F_2(1,24)=20.18$, $MSE=191$, $p<.001$, SFD: $F_1(1,39)=39.89$ $MSE=193$, $p<.001$, $F_2(1,24)=21.15$, $MSE=215$, $p<.001$] as was the main effect of predictability [FFD: $F_1(2,78)=22.21$, $MSE=212$, $p<.001$, $F_2(2,48)=13.14$, $MSE=198$, $p<.001$, SFD, $F_1(1,39)=24.69$, $MSE=216$, $p<.001$, $F_2(1,24)=16.67$, $MSE=178$, $p<.001$]. HF target words were fixated for less time than LF targets (FFD: 199 ms vs. 210 ms, SFD: 199 vs. 211 ms respectively). For the significant main effect of predictability, follow-up contrasts showed that of the three groups, two groups (LP vs. MP and LP vs. HP) had significant differences in FFDs and SFDs on target words; fixation time on LP targets was significantly longer than on MP targets [FFD; LP: 213 ms vs. MP: 203 ms; F_1=21.59, p<.001, and F_2=14.45, p<.001, SFD; LP: 215 ms vs. MP: 202 ms; F_1=26.36, p<.001, and F_2=17.62, p<.001]; fixation time on LP targets was significantly longer than on HP targets [FFD: LP: 213 ms vs. HP: 198 ms; F_1=41.74, p<.001, and F_2=23.80, p<.001, SFD: LP: 215 ms vs. HP: 199 ms; F_1=45.19 p<.001, and F_2=30.61, p<.001]. In the final comparison, the difference in FFDs between MP and HP targets (203 ms vs. 198 ms) was marginally significant in the subjects analysis and nonsignificant in the items analysis [F_1=3.25, p=.074, and F_2=1.16, p>.25]. In the SFD measure, the difference between MP and HP targets (202 vs. 199 ms) was trend in subjects analysis and nonsignificant in items analysis [F_1=2.52, p=.116, and F_2=1.78, p>.15].

In the FFD measure (see Figure 3.3.), the frequency x predictability interaction was significant in the subjects analysis and marginally significant in the items analysis and in the SFD measure (see Figure 3.4.) the interaction was significant in both subjects and item analysis [FFD: F_1(2,78)=3.95, *MSE*=158, p<.05, and F_2(2,48)=2.93, *MSE*=202, p=.063, SFD: F_1(2,78)=3.96, *MSE*=65, p<.05, and F_2(2,48)=3.57, *MSE*=217, p<.05]. Follow-up contrasts examined the effect of frequency (LF vs. HF) at each level of predictability (LP, MP, & HP); the effect of predictability for LF words and the effect of predictability for HF words.

First, frequency contrasts (LF vs. HF targets) were significant in two predictability conditions, LP and MP. Specifically, HF targets were fixated for significantly less time than LF targets in the LP condition [FFD: 207 ms vs. 219 ms respectively; F_1=20.82, p<.001, F_2=10.83, p<.01, SFD: 208 vs. 221 ms respectively; F_1=21.72, p<.001, F_2=11.65, p<.01] and also in the MP condition [FFD: 195 ms vs. 210 ms respectively; F_1=29.15, p<.001, F_2=13.76, p<.001]. For LF and HF targets appearing in the HP condition, the FFDs (196 vs. 200 ms respectively) was trend for significance in the subjects analysis and nonsignificant in the items analysis [F_1=2.62, p=.110, F_2<1]. In the SFD measure, the HF and LF times (196 vs. 201 respectively) were marginally significant in subjects analysis and nonsignificant in items analysis [F_1=2.87, p=.094, F_2<1].

Second, we compared, for LF words, three groupings of predictability (LP vs. MP; LP vs. MP; and MP vs. HP). All differences between the respective predictability conditions were significant in FFD (LP>MP: 219 vs. 210 ms; F_1=11.53, p<.01, F_2=6.01, p<.05, LP>HP: 219 vs. 200 ms; F_1=45.89, p<.001, F_2=22.88, p<.001, and MP>HP: 210 vs. 200 ms; F_1=11.42, p<.01, F_2=5.43, p<.05) as well as in SFD (LP>MP: 221 vs. 210 ms respectively; F_1=14.04, p<.001, F_2=5.66, p<.05, LP>HP: 221 vs. 201 ms respectively; F_1=47.93, p<.001, F_2=25.01, p<.001, and MP>HP: 210 vs. 201 ms respectively; F_1=10.09, p<.01, F_2=6.88, p<.05).

The same predictability comparisons were carried out for HF words. Two groupings from the three were significant (LP>MP: 207 vs. 195 ms respectively; F_1=17.91, p<.001, F_2=8.24, p<.01, LP>HP: 207 vs. 196 ms respectively; F_1=14.66, p<.001, F_2=4.20, p<.05 – in FFD) and similarly in SFD (LP>MP: 208 vs. 194 ms respectively; F_1=20.84, p<.001, F_2=8.94, p<.01; LP>HP: 208 vs. 196 ms respectively; F_1=15.67 p<.001, F_2=4.31, p<.05). The difference in FFD between MP (195 ms) and HP (196 ms) conditions shows that MP targets incurred a *faster* reading time than HP targets – this small difference being nonsignificant [all *F*s<1]. Similarly, in the SFD measure, MP targets (194 ms) were also read faster than HP ones (196 ms) – this difference was nonsignificant too [all *F*s<1].

3.9.1.2 Gaze duration

The GD data showed a significant main effect of frequency [F_1(1,39)=38.62, *MSE*=384, p<.001, F_2(1,24)=18.63, *MSE*=449, p<.001] and a significant main effect of predictability [F_1(2,78)=27.29, *MSE*=291, p<.001, F_2(2,48)=13.62, *MSE*=329, p<.001]. The frequency x predictability interaction was nonsignificant [all *Fs*<1]. In terms of the significant frequency main effect, there were faster GDs on HF words than on LF words (206 vs. 222 ms respectively). For the significant main effect of predictability, follow-up contrasts showed that there were slower GDs on LP targets (226 ms) than on both MP targets (210 ms) [F_1=31.76, p<.001, F_2=16.90, p<.001] and HP targets (207 ms) [F_1=48.36, p<.001, F_2=23.44, p<.001]. The difference in GDs between MP (210 ms) and HP (207 ms) targets was nonsignificant [F_1=1.74, p>.15, F_2<1].

3.9.1.3 Total Time

The TT data showed the same pattern as the GD data. That is, there were main effects of frequency and predictability and a nonsignificant interaction. Participants TT was significantly less on HF (222 ms) than on LF (240 ms) targets [F_1(1,39)=28.34, *MSE*=658, p<.001, F_2(1,24)=10.18, *MSE*=900, p<.01]. The main effect of predictability was also significant [F_1(2,78)=21.11, *MSE*=698, p<.001, F_2(2,48)=13.69, *MSE*=626, p<.001]. Follow-up contrasts showed that all three group comparisons were significant: TT on LP targets (245ms) was greater than TT on MP targets (230 ms) and both were greater than TT on HP targets (218 ms) and MP was significantly greater than HP [LP-MP:

F_1=13.48, p<.001, F_2=10.08, p<.01; LP-HP: F_1=41.96, p<.001, F_2=26.92, p<.001; MP-HP: F_1=7.88, p<.01, F_2=4.06, p<.05). The interaction of word frequency and contextual predictability was nonsignificant [all *Fs*<1].

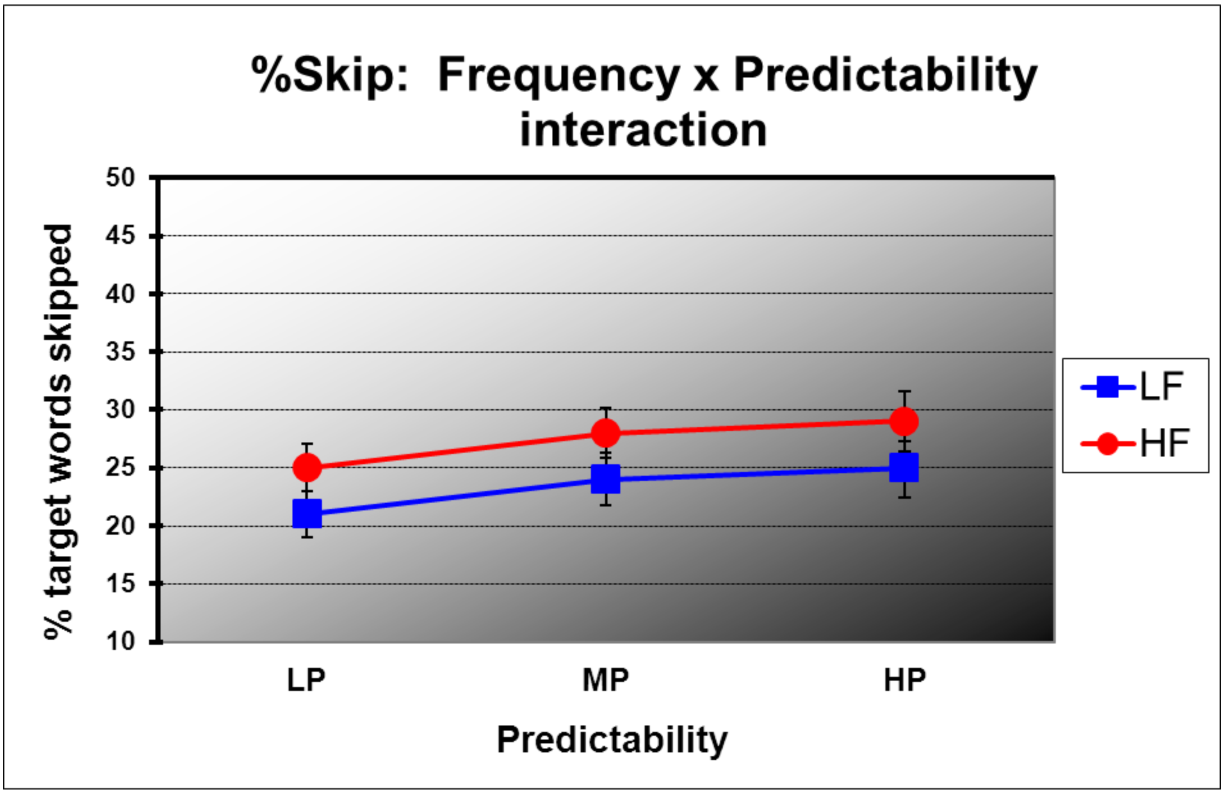
3.9.1.4 Percentage of times the target was skipped

The %Skip data showed a significant main effect of frequency [F_1(1,39)=10.79, *MSE*=93, p<.01, F_2(1,24)=10.26, *MSE*=61, p<.01] and a significant main effect of predictability [F_1(2,78)=5.19, *MSE*=48, p<.01, F_2(2,48)=4.36, *MSE*=58, p<.05]. However, the frequency x predictability interaction was nonsignificant [all *Fs*<1]; see Figure 3.5. In terms of the significant frequency main effect, participants skipped HF targets more than LF targets (27% and 23% respectively). For the significant main effect of predictability, Bonferroni follow-up contrasts showed that patterns of skipping were in the expected direction: participants skipped MP targets more than LP target (26% vs. 23%); this difference was significant in the subjects analysis only, being marginally significant in items analysis [F_1=4.49, p<.05, F_2=3.77, p=.058]. HP targets (27%) were skipped more than LP targets (23%); significant in both subjects and items analyses F_1=9.99, p<.01, F_2=8.39, p<.01]. The third and final comparison showed that there was no difference between skipping rates of HP targets (27%) and MP targets (26%) [F_1=1.08, p>.30, F_2<1].

3.9.1.5 Sentence 1 first pass time across the conditions

Even though line 1 of the materials was such that for all the sentences, the line spanned approximately one line of text, it is possible that because line 1 was not exactly matched in number of characters some part of the obtained results could be dependent on possible idiosyncratic aspects of varying character length. Therefore, one further analysis was carried out in which the whole of sentence 1 was computed as one region and the first pass reading time (milliseconds, ms, per character) was calculated across the six conditions (shown in row 8 in Table 2.6). These analyses showed that both main effects of word frequency and contextual predictability were nonsignificant [all *Fs*<1]. Thus, there was a nonsignificant difference in first pass time when the later target was HF (29 ms) or LF (28.6 ms). Also, the interaction of word frequency by contextual predictability was nonsignificant [all *Fs*<1]. Thus, there were no significant differences in fixation time on line 1 across the six conditions.

Figure 3.5. Percentage of times the target word was skipped as a function of word frequency and contextual predictability



Note: %Skip= percentage of times the target word was skipped; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

3.9.1.6 Results summary

The results are summarised in Table 3.13, 3.14, 3.15, and 3.16. Our hypotheses regarding the FFD and SFD were overall supported; hypotheses regarding skipping the target word were not supported; and GD and TT hypotheses were supported. In the two early processing measures (FFD and SFD), a similar interactive pattern of results was observed, namely that there was a main effect of frequency and of predictability and a significant interaction; the difference in the measures was that in the FFD measure, the interaction was marginal by items. For both FFD and SFD follow-ups to the main effect of predictability showed that of the three contrasts, two (LP-MP and LP-HP) were significant. However, the difference in fixation time in both FFD and SFD measures for MP-HP was marginally significant in the subjects analysis and nonsignificant in the items analysis. In both FFD and SFD measures, the follow-ups to the significant interaction showed that there was a significant frequency effects (that is, longer fixations on LF than on HF words) for LP and MP conditions; in the HP condition, the difference between fixations on LF and HF words was marginally significant by subjects and nonsignificant by items. For the predictability effects, for LF targets, all predictability conditions significantly differed from each other (LP vs. MP; on MP vs. HP; and MP vs. HP). Analogous comparisons for HF words showed that only the first two comparisons were significant. In addition in both FFD and SFD measures, there was an unexpected numeric difference for HF words where MP words were read faster than HP words. However, because we did not predict results in this direction, they warrant further explanation.

In the %Skip measure, additive results were shown: there were main effects of frequency (with significantly more skipping of HF than LF targets) and of predictability (participants significantly skipped more MP than LP targets as well as more HP than LP ones. The interaction was not statistically significant. In the GD measure, there was an additive pattern of results: a significant main effect of frequency (with significantly shorter GDs on HF than on LF targets), a significant main effect of predictability (with significantly longer GDs on LP targets than on both MP targets and HP targets). The interaction was not statistically significant. In the total time measure, the additive pattern of results was maintained. There was a significant main effect of frequency (with participants fixating for less time on HF than on LF targets); a significant main effect of predictability (participants fixated longer on LP targets than on MP targets, and both were fixated longer than HP targets).

Table 3.13 Summary of ANOVA Results in all Fixation Duration Measures and Skipping by Subjects (F_1) and by Items (F_2)

Main effect frequency						
Measure			***df***	***F***	***MSE***	***p***
FFD	**LF > HF**	F_1	1,39	38.34	184	<.001
		F_2	1.24	20.18	191	<.001
SFD	**LF > HF**	F_1	1,39	39.89	193	<.001
		F_2	1,24	21.15	215	<.001
GD	**LF > HF**	F_1	1.39	38.62	384	<.001
		F_2	1,24	18.63	449	<.001
TT	**LF > HF**	F_1	1,39	28.34	658	<.001
		F_2	1,24	10.18	900	<.01
% Skip	**LF < HF**	F_1	1,39	10.79	93	<.01
		F_2	1,24	10.26	61	<.01
Sent. 1	**LF vs. HF**	F_1	1,39	*<1*		*ns*
		F_2	1,24	*<1*		*ns*
Main effect contextual predictability						
FFD		F_1	2,78	22.21	212	<.001
		F_2	2,48	13.14	198	<.001
SFD		F_1	2,78	24.69	216	<.001
		F_2	2,48	16.67	178	<.001
GD		F_1	2,78	27.29	291	<.001
		F_2	2,48	13.62	329	<.001
TT		F_1	2,78	21.11	698	<.001
		F_2	2,48	13.69	626	<.001
% Skip		F_1	2,78	5.19	48	<.01
		F_2	2,48	4.36	58	<.05
Sent. 1		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
Interaction word frequency x contextual predictability						
FFD		F_1	2,78	3.95	158	<.05
		F_2	2,48	2.93	202	=.063
SFD		F_1	2,78	3.96	65	<.05
		F_2	2,48	3.57	217	<.05
GD		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
TT		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
% Skip		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
Sent. 1		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*

Note: LF = low frequency; HF = high frequency; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Table 3.14 Follow up Contrasts to Significant Main Effect Contextual Predictability by Subjects (F_1) and by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP > MP		LP > HP		MP vs. HP	
	F_1	21.59	<.001	41.74	<.001	3.25	=.074
	F_2	14.45	<.001	23.8	<.001	1.16	>.25
SFD		LP > MP		LP > HP		MP vs. HP	
	F_1	26.36	<.001	45.19	<.001	2.52	=.116
	F_2	17.62	<.001	30.61	<.001	1.78	>.15
GD		LP > MP		LP > HP		MP vs. HP	
	F_1	31.76	<.001	48.36	<.001	1.74	>.15
	F_2	16.90	<.001	23.44	<.001	*<1*	*ns*
TT		LP > MP		LP > HP		MP > HP	
	F_1	13.48	<.001	41.96	<.001	7.88	<.01
	F_2	10.08	<.01	26.92	<.001	4.06	<.05
% Skip		LP < MP		LP < HP		MP vs. HP	
	F_1	4.49	<.05	9.99	<.01	1.08	>.30
	F_2	3.77	=.058	8.39	<.01	*<1*	*ns*

Note: FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; LP = low predictability; MP = medium predictability; HP = high predictability

Table 3.15 Frequency Contrasts (LF vs. HF): Follow ups to when the Word Frequency x Contextual Predictability Interaction was Significant by Subjects (F_1) and by Items (F_2)

		LP words		**MP words**		**HP words**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***
FFD		LF > HF		LF > HF		LF vs. HF	
	F_1	20.82	<.001	29.15	<.001	2.62	=.110
	F_2	10.83	<.01	13.76	<.001	*<1*	*ns*
SFD		LF > HF		LF > HF		LF vs. HF	
	F_1	21.72	<.001	29.15	<.001	2.87	=.094
	F_2	11.65	<.01	13.76	<.001	*<1*	*ns*

Note: FFD = first fixation duration; SFD = single fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

Table 3.16 Contextual Predictability Contrasts: Follow-ups when the Word Frequency x Contextual Predictability Interaction was Significant

		LF words						HF words					
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP > MP		LP > HP		MP > HP		LP > MP		LP > HP		MP vs. HP	
	F_1	11.53	<.01	45.89	<.001	11.42	<.01	17.91	<.001	14.66	<.001	<*1*	*ns*
	F_2	6.01	<.05	22.88	<.001	5.43	<.05	8.24	<.01	4.20	<.05	<*1*	*ns*
SFD		LP > MP		LP > HP		MP > HP		LP > MP		LP > HP		MP vs. HP	
	F_1	14.04	<.001	47.93	<.001	10.09	<.01	20.84	<.001	15.67	<.001	<*1*	*ns*
	F_2	5.66	<.05	25.01	<.001	6.88	<.05	8.94	<.01	4.31	<.05	<*1*	*ns*

Note: FFD = first fixation duration; SFD = single fixation duration; LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability

3.10 Discussion

The current experiment investigated the word frequency x predictability interaction in several target word fixation times as well as the likelihood of skipping the target word during reading passages of text whilst participants' eye movement were recorded. Target words were either low frequency (LF) or high frequency (HF). These words were embedded in the second sentence of two-sentence passages whose predictability (as determined from prior Cloze probability and confirmed via rating task norms) were low, medium, or high predictability (LP, MP, HP). Previous findings from RT studies (using techniques like naming and lexical decision) have demonstrated interactive effects of frequency and predictability (i.e., main effects of frequency and predictability along with an interaction), however, later eye movement studies have, on the contrary, demonstrated additive effects of these two variables (i.e., main effects of frequency and predictability but no interaction). It was suggested that there were several possible methodological limitations with both the RT and eye movement studies and the current experiment attempted to address these limitations, for example, using more experimental items per condition in sentence contexts which were more accurately defined in terms of their contextual predictability. In addition, low predictable conditions avoided anomalous contexts and materials consisted of two-line sentence passages to allow sentence context to develop more fully. In regards to the previous additive eye movement studies, it was suggested that the reported additive results were between low and medium predictability conditions, and that the high end of predictability (as indicated by Cloze probabilities near 1.00) had not been sampled. Thus, we expected interactive results between frequency and predictability when considering very high predictability.

3.10.1 Main effect and interaction effects in the first and single fixation duration measures

Results in the two measures, first and single fixation duration measures, yielded similar main effect and interaction results (see Table 3.13 for a summary). In both measures, there was a significant main effect of frequency, a significant main effect of predictability as well as a significant interaction. In contrast, both Rayner et al. (2004) and Hand et al. (2010) showed the two significant main effects but no interaction. The main effect of frequency in all three studies were shorter fixations on high frequency than on low frequency targets. However, mean fixation times in first and single fixations were considerably faster in the current study than in both previous studies. Specifically, for low and high frequency targets, the mean first fixation duration was 210 ms and 199 ms respectively. In Rayner et al. and in Hand et al., the comparable times were 276 ms and 260 ms and 284 ms and 260 ms respectively. Similarly, the mean single fixation duration in the current study for low and high frequency targets was 210 ms and 199 ms respectively. The comparable times in Rayner et al. and in Hand et al. were 278 ms and 263 ms and 290 ms and 264 ms respectively. However, the effect sizes across studies were comparable to Rayner et al.'s study: in the present study (FFD, 16 ms; SFD, 17 ms) and in Rayner et al. (FFD, 15 ms; SFD, 15 ms). In Hand et al.'s study, the effect sizes were slightly larger than in both of these: Hand et al. (FFD, 24 ms; SFD, 26 ms). In all three studies, the mean frequency per million for low and high frequency targets was about the same as in the present study, 7 and 179 compared to 5 and 150 in Hand et al., 6 and 144 in Rayner et al. That is, there was a similar frequency range for low and high frequency targets across all three studies.

In terms of the significant main effect of predictability in both early first and single fixation measures, results showed significantly longer fixations on low predictable targets (FFD; 213ms, SFD; 215 ms) than on medium predictable ones (FFD; 203 ms, SFD; 202 ms). This result was also reported in Rayner et al. and in Hand et al. In addition, our results also showed longer first and single fixations on low predictable than on high predictable targets (HP, FFD; 198 ms; SFD; 199 ms). However, in both first and single fixation durations, there were no statistically significant differences between the medium and high predictable targets.

As with the first and single fixations in the main effect of frequency, predictability condition means in the present study were also much faster than in both Rayner et al. (2004) and Hand et al. (2010) studies. Specifically, in the present study, the low predictable mean was 213 ms in first fixation and 215 ms in single fixation; compared to

277 ms and 282 ms in Hand et al.; 272 ms and 278 ms in Rayner et al., respectively. Similarly, in the present study, the medium predictable condition mean was 203 ms in first fixation and 202 ms in single fixation. This can be compared to the slower first and single fixations respectively of 268 ms and 272 ms in Hand et al.; 262 and 264 ms in Rayner et al. The new results from the present study was the reading times in a very high predictable condition, here we showed reading times of 198 ms and 199 ms (FFD & SFD respectively).

In terms of the effect sizes in the present study, for a comparison with Hand and Rayner et al, we calculated the size of the predictability effect excluding our HP condition to allow for an equivalent comparison. The effect sizes in our study (FFD, 11 ms; SFD, 13 ms) were comparable to those in both Hand et al. (FFD, 9ms; SFD, 10 ms) and Rayner et al. (FFD, 11 ms; SFD, 14 ms) studies.

Present results differed from Hand et al. (2010) Rayner et al. (2004) in that in both first and single fixation duration, the current study showed a significant interaction of frequency x predictability (though in the first fixation, the interaction was marginal in items). Results showed the same patterns of significance in the first and single fixation measures. Frequency contrasts (LF vs. HF) for low, medium and predictable conditions showed that there was a significant frequency contrast (in the desired direction where HF targets were read faster than LF ones) for low and medium predictable targets. The frequency contrast was nonsignificant for high predictable targets. Predictability contrasts (comparing reading times at two levels of predictability) for low frequency targets showed that all contrasts were significant: high predictable targets were read faster than medium and low predictable ones, as well as medium predictable targets being read than low predictable ones. Predictability contrasts for high frequency targets showed that two out of three contrasts were significant: low predictable targets had significantly slower first (and single) fixation durations than both medium and high predictable targets. The unusual result was between the medium and high predictability conditions, first the contrast was nonsignificant, but also there was a reversed predictability effect where medium predictable targets had a shorter time than high predictable ones.

Overall the results in the first and single fixation durations showed significant main effects of frequency and of predictability as well as a significant interaction, which was frequency effects (that is shorter fixations on HF than on LF targets) for low and medium conditions only; and secondly predictability effects in low frequency targets (that is, longer fixations on LP vs. MP; on MP vs. HP; and MP vs. HP respectively). Analogous comparisons for high frequency targets showed that only the first two comparisons were significant. There was no effect of frequency in the high predictable condition, and also no differences between medium-high predictability conditions when reading high frequency targets.

If we consider our results at just low and medium (or past studies, high) predictable conditions (see Figure 3.3. and 3.4. for FFD and SFD interaction plots respectively), we can see that the lines look fairly parallel – suggesting that an interaction is not present but rather looked like additive results between frequency x predictability. Our results then – where our medium predictable is the past eye movement studies high predictable (given the Cloze probabilities used) – could be said to replicate the additive result between frequency x predictability.

Two results need explaining. First, the lack of a frequency effect for high predictable targets and second, the nonsignificant results for the medium versus high predictable condition for high frequency targets. A possible explanation is to do with the quality of the display used in the current study. Materials were presented to participants using a clear, good quality font i.e., black letters on a white background on a flat screen CRT in a well lit room. It could be that this is a better quality text and as such is easier to read than font presentations using dot-matrix. This is supported by the observation that in two previous eye movement studies, Rayner et al.'s (2004) and Hand et al.'s (2010) single fixation times were much longer than in the present study (see Table 3.17).

Table 3.17 shows that mean single fixation in the present study was 205 ms – considerable faster the 271 ms in Rayner et al.'s study and 277 ms in Hand et al.'s study. In both of these previous studies, there was a more 'unnatural' reading situation than in the current study; a Dual-Purkinje eye tracker was used and materials were presented in a dot-matrix font (cyan letters on a black background) in a dimly lit room. This can make the text more pixelated or blurred and as such is more difficult to read hence the longer reading times in both these studies.

Table 3.17 Single Fixation Durations across Experimental Conditions in Current Study and in Hand et al. (2010) and in Rayner et al. (2004)

	LF			HF			
	LP	MP	HP	LP	MP	HP	Mean SFD
Current study	221	210	201	208	194	196	205
Rayner, Ashby, Pollatsek, & Reichle (2004)	287	269	na	268	258	na	271
Hand, Miellet, Sereno, & O'Donnell (2010)	294	285	na	269	259	na	277

Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability; SFD = single fixation duration

The fast fixation times apparent in the present study therefore could explain why there were no frequency differences in the high predictable condition, and also no differences between medium-high predictability for high frequency targets. In particular, it could be that the reading times were so fast that there are floor effects in the data. That is, because readers were already reading so fast in the conditions which are prone to fast reading times anyway (i.e., high frequency targets, high predictable contexts) any effects of frequency and/or predictability may not show because constraints posed by the visual system means that it is simply not possible to fixate any faster, hence the ‘floor’. This explanation could also account for the unexpected reversed predictability effect in high frequency targets where the high predictable targets had a slightly slower fixation time than medium predictability ones. Fixation times when words are both high frequency and high predictable should lead to the least time spent on the word but excellent quality displays will have speeded up reading time and due to visual system constraints, there is a lower limit posed by the system and as such, it is not possible to read any faster. Therefore, it is suggested that a floor effect could potentially be masking frequency and/or predictability effects, and it is possible that in the absence of such a floor effect (that is, if there were no natural restriction on fixation times getting faster), the least time would indeed be spent on high frequency-high predictable target. In such a case, results would be additive i.e., no interaction would be present. In a more recent study on a corpus of eye movement data by Kennedy and colleagues, they found no evidence that frequency and predictability interact in early fixation durations or in gaze duration (Kennedy, Pynte, Murray & Paul, 2012).

3.10.2 Skipping data

In terms of skipping rate across the conditions of the experiment, we calculated the percentage of times the target word was skipped per condition. There was a significant main effect of frequency, predictability and no interaction (see Figure 3.5. in the Results section). In contrast, Rayner et al.'s (2004) and Hand et al.'s (2010) studies found only the main effect of frequency (in the latter study, frequency was significant in subjects analysis only) as well as a significant interaction. In the present study, follow-ups to the significant main effect of predictability were in the expected direction: medium predictable targets were skipped significantly more than low predictable ones targets (though marginally significant in items analysis). Results also showed that high predictable targets were skipped significantly more than low predictable ones. In the expected direction was that high predictable targets had a higher percent of skipping than medium predictable targets; however this difference was nonsignificant.

The skipping results in both Rayner et al. (2004) and Hand et al. (2010) studies showed a significant interaction in which words in HF-MP (in their studies, labelled HF-HP) were skipped more than targets in any of the other three condition (LF-LP; LF-MP; and HF-LP, which did not differ from each other). Table 3.18 show the percentage of times the target word was skipped; in the current study (row 2); in Rayner et al. (2004) (row 3); in Hand et al. (2010) (row 4) and in Hand et al. (2012) (row 5).

Table 3.18 Percentage of Times the Target Word was Skipped in Current Study; Rayner et al. (2004); Hand et al. (2010) and in Hand et al. (2012)

	LF				HF				Mean percent skipping
	LP	**MP**	**HP**		**LP**	**MP**	**HP**		
Current study	21%	24%	25%		25%	28%	29%		25%
Rayner, Ashby, Pollatsek, & Reichle (2004)	19%	14%	na		15%	23%	na		18%
Hand, Miellet, Sereno, & O'Donnell (2010)	18%	17%	na		19%	23%	na		19%
Hand, O'Donnell, & Sereno (2012)	20%	30%	na		17%	34%	na		25%

Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability

In the present study, then, in the skipping data we obtained an additive result between frequency and predictability. There is a relationship between the percentage of times a word is skipped and its length (Rayner & McConkie, 1976). That is, word skipping occurs more often with short words than longer words (Blanchard, Pollatsek, & Rayner, 1986; Brysbeart & Vitu; 1998; Rayner, 1979) and longer words tend to be fixated (Rayner & McConkie, 1976). Short words typically refer to 4 letters or less and long words to 8 letters or more. One suggestion is that when readers skip a word, they do so because they have already identified it on the prior fixation (Rayner & Duffy, 1988). Thus, it is possible that the two measures (fixation time measures and skipping) depend unequally on words of varying length. That is, words of shorter length in the study may have contributed more to the skipping data and longer words to the fixation duration data and therefore the pattern of data observed might actually reflect such a pattern for shorter and longer words. However, this explanation seems implausible since word length in the study spanned a relatively narrow range from 4 to 8 letter words. In addition, the majority of our words were actually 5 and 6 characters in length.

As Table 3.18 shows, skipping rates in the present study were considerably higher than in Rayner et al. (2004) and in Hand et al. (2010) studies. The reason for increased skipping in the present study than in these two past studies could be because we had more shorter words in our study than did Rayner et al. and Hand et al. in their studies. That is, we used 4 letter words in our study whereas in the other two studies, the shortest word was 5 letters in length. However, this explanation is not fully plausible when it is considered that in the present study, of the 25 words used per condition, only five targets were 4 letters in length. As in Rayner et al. (2004) and Hand et al. (2010) studies, the majority of our words were also 5 and 6 letters in length: in Rayner et al., of the 32 targets, 29 were 5 or 6 letters; in Hand et al., of the 44 targets, 33 were 5 or 6 letters; and in the present study, of the 25 targets 11 were 5 or 6 letters. Thus another more likely explanation of the high rates of skipping in the present study could be to do with the font and parafoveal word information.

As discussed earlier in the present study, the good quality presentation display is likely to have led to faster fixation durations in the present study than in previous studies which have used point-plotting displays. Furthermore, it is possible that a clear font makes the accessibility of parafoveal information better and this can possibly explain the increased skipping in the present study. That is, when words are skipped, this suggests that readers are able to extract information not only from the currently fixated word but also from the next (parafoveal) word as well. The suggestion is that a clearer font leads to a more advantageous situation in which the reader can completely identify a word in the parafoveal region and in such a case, that word would subsequently be skipped. A recent study by Hand, O'Donnell and Sereno (2012) showed similar rates of skipping to the current study (see row 5 in Table 3.18). That is, in Hand et al. (2012), the percentage of times the target word was skipped was 20%; 30%; 17%; 34% in LF-LP; LF-MP; HF-LP; HF-MP respectively (see Table 3.18, column 5). Interestingly, Hand et al.'s (2012) materials were also presented using the identical display as that used in the present study. However, the interaction in Hand et al. (2012) was significant, in which there was more skipping in moderately predictable conditions (labelled high predictable in the study) for high and low frequency targets. Thus, this finding was somewhat in contrast to previous research which had shown increased skipping in moderately predictable high frequency conditions only (Hand et al., 2010; Rayner et al., 2004).

3.10.3 Gaze duration and total time

In terms of the gaze duration data, there was a main effect of frequency and of predictability and a nonsignificant interaction. This was also the result reported in Rayner et al. (2004) and Hand et al. (2010) studies. Across all three studies, the significant frequency effect was shorter gaze durations on high frequency than on low frequency targets. Further support for the view that clearer display qualities gives rise to faster reading times can be seen from the comparison of gaze duration times across the three studies. Specifically, in the current study, gaze durations were faster than in both the previous studies: the mean gaze duration for high and low frequency targets was 199 ms and 211 ms respectively. These times can be compared with the much slower 282 ms and 299 ms in Rayner et al.'s (2004) study and 280 ms and 312 ms in Hand et al.'s (2010) study, respectively. The frequency effect sizes were somewhat comparable across the three studies: in the present study the frequency effect size in gaze duration was 24 ms, compared to 17 ms in Rayner et al. (2004) and 33 ms in Hand et al. (2010).

In gaze duration, the main effect of predictability in Rayner et al. (2004) and in Hand et al. (2010) studies was that medium predictable targets were fixated significantly shorter time than low predictable targets; present results also showed this. In addition, the new finding from the present study was that there were significantly shorter gaze durations on high predictable than on low predictable targets. As in the first and single fixations, the medium-high predictability contrast remained nonsignificant. In addition, predictability condition means in the present study were faster than those in Rayner et al. (2004) and

Hand et al. (2010) studies. In the present study, the average gaze duration in low, medium and high predictable conditions was 226, 210 and 207 ms respectively. In Rayner et al. and Hand et al. studies, the average gaze duration in the low predictable condition was 300 and 302 ms respectively; and in the medium predictable condition, the average was 281 and 300 ms respectively. In terms of predictability effect sizes, again this was calculated excluding the high predictable condition to allow for an equivalent comparison with the present study. The effect size in gaze duration (16 ms) was comparable to that in Rayner et al. (19 ms) and in Hand et al. (13 ms) studies. The interaction of frequency x predictability was nonsignificant, replicating Hand et al. (2010) and Rayner et al. (2004) studies who also reported no frequency by predictability interaction in this measure.

For the final measure, total time, present results replicated those in Rayner et al. (2004) and Hand et al. (2010) studies. We showed a significant main effect of frequency with shorter fixation time on high than on low frequency targets; a significant main effect of predictability and a nonsignificant interaction of frequency x predictability. Follow-ups to the significant main effect of predictability followed the same patterns as in Hand et al. (2010) and Rayner et al. (2004) studies: there were significantly shorter fixations on targets in the medium predictable condition than in low predictable one. The results from the present study also showed that there was significantly shorter time on high versus low predictable conditions as well as high versus medium predictable conditions. The reason that the medium-high predictability contrast was significant here but not in first, single and gaze duration is because total time includes regressions back to the target, hence it not being a measure of initial processing of a word.

3.10.4 Conclusions

Our experiment examined the single and combined effects of word frequency and contextual predictability on words during normal reading while readers' eye movements were measured. Past reaction time studies have generally found interactive effects of these two variables. However, the methods used in these behavioural studies make them generally limited when it comes to normal reading. Past eye movement studies have overall found additive effects of these variables; however we argued that if Cloze values are properly assigned to their correct categories, then it is possible that there might be an interactive relationship between these two variables. To this end, we examined three levels of contextual predictability with word frequency.

Results were that in first and single fixation duration measures, there was an interaction of frequency x predictability, and additive results of frequency x predictability in gaze duration and total time. If first and single fixations are taken reflect the process of lexical access then this interaction lends support to an interactive architecture of the language processing system. However, the present experiment had very fast reading times, as seen when comparing our single fixation durations with past eye movement frequency-predictability studies (Hand et al., 2010; Rayner et al., 2004). In comparison with these two earlier eye movement studies, the present study also had considerably higher rates of skipping.

It is possible that that there are floor effects in the data, which are masking an actual additive relationship between frequency and predictability. It appears that there may be a floor effect in some of the conditions (suggested by relatively less variance). For example, for the single fixation duration, the standard deviations in the HF-MP and HF-HP conditions indicate less variance in the data compared to the other conditions. If it is the case that there are floor effects operating, then results in first and single fixations would be in favour of a modular view, which would state that the interaction in first and single fixation durations exists because floor effects are masking an actual additive relationship between frequency and predictability. This view would go on say that there should be an interaction of frequency and predictability in later measures, that is, gaze duration and total time. However, in both these measures, we observed additive results of frequency and predictability.

The frequency-predictability relationship needs further investigation and in the next experiment (Experiment 3, Chapter 4), we continue to explore the frequency-predictability relationship by using a novel way to slow down reading times, that is, using the boundary technique (Rayner, 1975) in the materials.

Chapter 4

Word frequency and contextual predictability effects in an eye movement-contingent boundary change paradigm

4.1 Introduction

Experiment 2 (Chapter 3) examined the word frequency by contextual predictability relationship and it was shown that these two variables interacted in early fixation time measures, single and first fixation durations. Since these measures are likely to capture the initial processing of a word such as lexical access processes, results from this first experiment suggested that frequency and predictability may exert an early influence at the lexical access stage of word processing. Experiment 2 was the first study to date in the literature to sample very highly predictable target words from the high end of Cloze predictability values. Thus, while a number of other eye tracking studies have shown that frequency and predictability are additive in fixation time measures (e.g., Altaribba et al., 1996; Ashby et al., 2005; Hand et al., 2012; Lavigne et al., 2000; Rayner et al., 2001; Rayner et al., 2004) suggesting that frequency and predictability do not influence the same stage of word encoding, by using very high predictability in Experiment 2, it was shown that the two variables can interact in early fixation durations. Therefore, these results were in agreement with early behavioural studies which used technique like lexical decision and naming (Stanovich & West, 1981; 1983). In addition, results from Experiment 2 added to a small body of research which have shown an interactive relationship between frequency and predictability in skipping data (Hand et al., 2010; Rayner et al., 2004) and when launch site is taken into consideration (Hand et al., 2010).

The aim of the current experiment (Experiment 3) was to examine the frequency-predictability interaction (obtained in first and single fixation duration measures in Experiment 2) when readers were prevented from parafoveally processing the target. This is because it is possible that some of the observed interaction in the first and single fixation duration measures in Experiment 2 could have been due to the parafoveal processing aspect of reading high frequency and high contextual predictability targets (e.g., Inhoff & Rayner, 1986; Balota et al., 1985). In normal reading, we process not only the word that is fixated (i.e., in foveal vision, word n) but also some aspects of the adjacent word (i.e., the parafoveal word, n+1). When readers are fixating word n, the ability to extract information from word n+1 has been shown to be influenced by the frequency and predictability of word n+1. That is, parafoveal preview benefit is greater for high frequency than for low frequency words (Inhoff & Rayner, 1986) and also, the extraction of information is more efficient when aided by sentential context (Balota et al., 1985). These findings suggest that word frequency and contextual predictability affect lexical access as opposed to having their effect on later post-lexical stages such as semantic integration. In Experiment 2, therefore, it is possible that lexical access was facilitated when the parafoveal word was of higher frequency (e.g., Inhoff & Rayner, 1986) and also of higher contextual predictability (e.g., Balota et al., 1985). If this is indeed the case, then this suggests further support for an interactive view of lexical access (e.g., McClelland & O'Regan, 1981; McClelland & Rumelhart, 1981). The cross Experiment 2 and 3 comparison was further data analyses in which a three-way design (word frequency x contextual predictability x parafoveal preview) was computed with the data collected in the respective Experiments.

In Experiment 3, it was considered important to examine the frequency-predictability interactive relationship when targets were viewed (processed) without the possible parafoveal processing advantage from high frequency words and high contextual predictability (given the preceding context) in order to evaluate the foveal component of the frequency-predictability interaction. More specifically by inhibiting participants from parafoveally processing the target, they would be prevented from acquiring the processing advantage from high frequency words versus low frequency words (e.g., Inhoff & Rayner, 1986) and the processing advantage from higher contextual predictability words versus lower contextual predictability words (e.g., Balota et al., 1985). To inhibit processing the parafoveal word, a boundary technique was used so that participants initially viewed a letter string other than the target on the pre-target fixation. Only when the eyes crossed the pre-specified boundary location (placed after the last letter of the pre-target word) did the letter string change to the target itself. In this way, it could be ensured that readers only started processing the target when it was directly fixated in foveal vision. This would mean that targets were not pre-processed on the prior fixation.

By using the boundary technique we were also able to address another issue arising from Experiment 2. That is, in this experiment, there were very fast fixation duration's compared to equivalent past eye tracking studies (e.g., Hand et al., 2010; Rayner et al., 2004). It appeared that there may have been a floor effect in the data and in the absence of this, an additive result between word frequency and contextual predictability may have emerged. One reason that Experiment 2 showed very fast fixation durations (in particular for high frequency and for high predictability targets) was because of the very clear font used in our study compared to more pixelated fonts in the previous studies (Hand et al., 2010; Rayner et al., 2004). In support of this suggestion, results in Experiment 2 had a greater percentage of skipping than in past frequency-predictability studies by Hand et al. (2010) and by Rayner et al. (2004). If it is the case that lexical access starts from parafoveally viewing the target word then one possible reason that words are skipped is if they are identified on the fixation prior to a given target. Therefore, it is possible that the clear font in our study made parafoveal information about the target word (on the pre-target fixation) more easily obtained, and as such the target word was skipped because more information about it was acquired on the pre-target fixation. This seems particularly the case with the high frequency and high contextual predictability targets since these were skipped the most than any other words in Experiment 2.

A way to prevent the floor effect (in particular with higher frequency and high predictability words) is to slow down participants' rate of reading. This could be achieved by distorting the text in some way with the aim of forcing participants to fixate longer on the text. However, disrupting the text in such a way is atypical of normal reading. Therefore, it was reasoned that a more ecologically valid way of slowing down reading times was to inhibit participants from parafoveally viewing the target word by using the boundary technique (Rayner, 1975). We know from past studies that when parafoveal information is withheld, reading rate is slowed by as much as 40% (Rayner et al., 1982).

The latter part of this Chapter presents the cross Experiments 2 and 3 comparison. Here, additional data analyses were conducted with data collected in the two prior experiments, Experiments 2 and 3. The aim of these further analyses was to examine whether word frequency and contextual predictability of the parafoveal word affected parafoveal preview benefit. The usual way of computing parafoveal preview benefit is by comparing a condition of a valid preview of a target word with a condition of either no preview or an invalid preview of a given target (Rayner, 1998). Thus, for these further analyses, data from Experiment 2 provided a condition of valid preview because here readers had obtained the normal preview of the target. Experiment 3 provided data for the condition of invalid preview because in this Experiment, readers had read all the experimental passages with an invalid parafoveal preview of the target.

4.1.1 Parafoveal preview benefit

Parafoveal preview benefit is the term used to refer to the fixation time advantage on a target word (word n) when the parafoveal information of that target word (word n+1) provided useful information regarding the upcoming target word compared to when there was no useful information provided about that upcoming target word. The first condition is typically labelled as the 'valid parafoveal preview of the target word' and the latter as the 'invalid parafoveal preview of the target word'. Given the size of the perceptual span, it is not surprising that in normal reading readers are able to not only process information about the currently fixated word n but also some information about the upcoming parafoveal word n+1 (e.g., Rayner, 1998; Sereno & Rayner, 2000). However, exactly what information is processed about the parafoveal word is debated. In fact, several debates are evident in the literature (e.g., Liversedge & Findlay, 2000; Starr & Rayner, 2001) such as exactly what kinds of information are extracted about the parafoveal word (see Section 3.1.2. below) and what characteristics of the parafoveal word affect how much it is extracted on fixation n. It is the latter which is of interest in the present study.

4.1.2 The kinds of information extracted from the parafoveal word

There is overall agreement that readers can obtain information about the parafoveal word in terms of its abstract letter codes as well as its sub-lexical and phonological codes (Henderson et al., 1995; Inhoff, 1989; McConkie & Zola, 1979; Pollatsek et al., 1992; Rayner et al., 1980, 1982). When it comes to the question of whether or not readers can obtain semantic information such as word meaning about the parafoveal word, many studies have shown that this does not occur, during either first fixation durations or gaze durations (in alphabetic writing systems; Altarriba et al., 2001; Hyönä & Häikiö, 2005; Rayner et al., 1986; Rayner et al., 1980 but cf. Hohenstein et al., 2010). The latter finding is important to the present discussion because the theoretical implication is that a late-processing account of word recognition is supported with orthographic and phonological processing taking place first i.e., being made available to the reader first with semantic information being available to the reader at a later stage (e.g., Forster, 1976; 1989). However, in contrast to these positions, there are two early studies which indicate that lexical access is influenced by certain characteristics of parafoveal information (Inhoff & Rayner, 1986; Balota et al., 1985). These two studies indicate support for interactive views of lexical access (e.g., McClelland & O'Regan, 1981; McClelland & Rumelhart, 1981). For example, it could be that lexical access is an interactive-activation type process in which contextual and parafoveal information are used to drive lexical access (McClelland & O'Regan, 1981). Inhoff and Rayner's (1986) and Balota et al.'s (1985) studies are described in the two subsequent sections below.

4.1.3 Parafoveal lexical processing: word frequency effects

Inhoff and Rayner (1986) examined the joint effects of word frequency and parafoveal information in a moving window study. They were interested in whether lexical properties of the parafoveal word, in particular word frequency of the parafoveal word, influenced parafoveal preview benefit. Low and high frequency target words were designed to fit into an identical single line sentence frame (participants read one version). Each pair of low and high frequency targets was matched on word length and contextual predictability. Parafoveal preview was manipulated by using the moving window technique so that sentences were read in one of three viewing conditions where information to the left was available at all times: one-word window (only the fixated word was seen); two-word window (the currently fixated word as well the word immediately next to it were seen) and full-line (the whole line of text was displayed). 24 participants read a total of 60 experimental passages.

Inhoff and Rayner (1986) predicted that if word frequency and parafoveal preview interact, that is, if there is a larger parafoveal preview benefit for high frequency words than for low frequency words, then this suggests that readers obtain more effective parafoveal previews from the high frequency than from the low frequency words. Also, when parafoveal previews are denied (i.e., in the one-word window condition) then there should be relatively small benefits of word frequency; when parafoveal previews are available (i.e., two-word and full-line viewing conditions) then relatively large benefits should accrue. The interaction was predicted in the first fixation duration measure. Overall the results they obtained were as hypothesised.

The full-line window can be considered as offering a valid preview of the target since the whole line of text was viewed in its entirety. The two-word viewing condition can also be considered as offering a valid preview of the target since the reader saw the fixated word as well as the next to it, thus obtaining a parafoveal preview of the target. Finally, the one-word window can be considered as offering an invalid parafoveal preview of the target since only the currently fixated word was seen; there was no preview of the parafoveal word provided to the reader.

First fixation duration and gaze durations across the conditions were analysed. Only subjects analysis were carried out. Thus, in both the first and gaze duration measures, there was a significant main effect of frequency where there were significantly shorter fixations on high versus low frequency words; there was a significant main effect of parafoveal preview (i.e., window size) where fixations in the one-word presentation were longer than those in both two-word and full-line conditions. In terms of the interaction between word frequency and parafoveal preview, in the gaze duration measure the interaction was nonsignificant but significant in the first fixation duration (see Table 4.1). Basically, a parafoveal preview of high frequency words lead to larger preview benefits than a parafoveal preview of low frequency words; follow-ups showed that this was in all three window size conditions.

Table 4.1 First Fixation Durations in Valid (full-line and two-word window) and Invalid Parafoveal Preview (one-word window) Conditions with High and Low Frequency Targets in Inhoff and Rayner's (1986) study

	Valid preview (full-line)	**Valid preview (two-word)**	**Invalid preview (one-word)**	**Size of parafoveal preview benefit***	
Target					
HF	237 ms	233 ms	277 ms	Full-line vs. one-word: 40 ms	Two-word vs. one word: 44 ms
LF	255 ms	261 ms	277 ms	Full-line vs. one-word: 22 ms	Two-word vs. one word: 16 ms

Note: * Calculated by subtracting the fixation time in a valid preview of target with the invalid preview of target; HF=high frequency; LF=low frequency

As can be seen in Table 4.1, when the parafoveal target was high frequency, participants' first fixations were 237 ms with a valid preview of the target (full-line) versus 277 ms with an invalid preview of the target. Hence the size of the parafoveal preview benefit was 40 ms. In contrast, when the parafoveal target was low frequency, first fixations were 255 ms with a valid preview of the target (full-line) and 277 ms with an invalid preview of the target (one-word) meaning that the parafoveal preview benefit was 22 ms when readers viewed low frequency words. Thus, when readers parafoveally viewed high frequency words, the preview benefit was 40 ms but when they parafoveally viewed low frequency words, the preview benefit was 22 ms. Similar preview benefit effects were seen when considering the first fixation time in the two-word window vs. the one-word window (see Table 4.1). These results are important because they suggest that parafoveal processing is influenced by the lexical characteristics of the parafoveal word, with the specific suggestion that high frequency parafoveal words are processed more effectively than low frequency words as evidenced by faster fixation durations when parafoveally viewing a high frequency than a low frequency word. This result indicates that more information is extracted from a parafoveal word which is of high frequency than a parafoveal word which is of low frequency.

On a theoretical level, the interaction of word frequency and parafoveal preview may have arisen via interactive mechanisms. For example, in the interactive threshold model (Balota and Rayner, 1990), it is envisaged that high frequency words have lower thresholds than low frequency words (e.g., Morton, 1969) which means that the interactive threshold for high frequency preview words is more easily exceeded than that for low frequency words. Therefore, lexical access takes more time (evidenced by longer fixation durations) the more ineffective the parafoveal information i.e., from lower frequency words.

Results also showed that when parafoveal previews were not available (i.e., in the one-word window), then there was no effect of frequency in the first fixation measure (but there was in the gaze duration measure). This result is at odds with the view that first fixations reflect lexical access. That is, if first fixation duration reflects lexical access processes, there should have been a frequency effect (i.e., first fixations on high frequency words should have been shorter than those on low frequency words, irrespective of their being a preview of the word). When gaze duration was taken into account (that is, the sum of all fixations on the target prior to making an eye movement to another word) then there was a frequency effect in the one-word window condition. In Inhoff and Rayner's (1986) study, the gaze duration measure included refixations to the target. Typically, the gaze duration measure includes the first and single fixation durations; gaze duration is increased when readers make more than one fixation on a target. The absence of the frequency effect in the 'invalid preview' condition could be taken to indicate that some fixations in reading are not determined by the frequency of the word. However, given the wealth of studies showing that replication of the word frequency effect (e.g., Rayner, 1998), a more likely explanation is that the eyes may have moved to a different location before the meaning of the word was accessed. Moreover, Inhoff and Rayner included refixations to the target in the gaze duration measure. Refixations of a word are likely to reflect processes that occur relatively late in word processing.

4.1.4 Parafoveal semantic processing: contextual predictability effects

Balota et al. (1985) investigated the joint effect of contextual predictability and parafoveal preview in a boundary study. They used a 2 (Predictability: LP, HP) x 5 (Parafoveal Preview: 5 conditions) design. Target words were embedded in a single line sentence frame (spanning over two lines). Each sentence frame was constructed so as to accommodate either a predictable target word or an unpredictable target word (given the prior sentence context). Contextual predictability of target words was determined via both a rating and a Cloze task.

One example sentence in Balota et al.'s (1985) study was: 'Since the wedding was today, the baker rushed the wedding cake to the reception'. The target word 'cake' was high predictable. The same sentence frame was used to accommodate a low predictable target ('pies'). For each high and low predictable target word, participants saw one of the following previews (for the target 'cake') when they were fixated on the word 'wedding': a visually similar nonword (VS) (first two or three letters were the same and all other letters replaced with visually similar letters, e.g., 'cahc'); a visually dissimilar nonword (VD) (visually dissimilar words were selected from those created in the previous condition; e.g., 'picz'); identical (Ident) word (preview was same as target; i.e., 'cake'); semantically related words (SR) where the low and high predictable targets for that sentence frame were reversed (e.g., the low predictable 'pies' was used for the equivalent high predictable 'cake') and finally in the anomalous (AN) preview condition, an word unrelated anomalous word in that context was used (e.g., 'bomb'). This lead to the formation of ten experimental conditions and a total of 30 participants read 10 sentences in each condition thus a total of 100 experimental sentences were read.

Balota et al. (1985) suggested that an interactive model such as McClelland and O'Regan's (1981) model would predict an interaction between contextual predictability and parafoveal preview such that there would be a larger parafoveal preview effect for high contextual predictability targets than for low contextual predictability targets. This interaction is because contextual constraint on its own may not be enough to push a lexical representation over its threshold but within an interactive model when combined with parafoveal preview information, may be sufficient to exceed a threshold. Thus, Balota et al. (1985) predicted that there would be an interaction between contextual predictability and parafoveal preview where in the example sentence, participants would be able to use the parafoveal information in the two conditions 'identical' and 'VS' (visually similar nonword) because 'cake' and 'cahc' respectively bear a consistency with the eventual target than they would be able to use the equivalent parafoveal information consistent with the low predictable target: for the low predictable target 'pies', the identical word and VS words were 'pies and 'pirc' respectively.

Balota et al. (1985) presented results for when the target word was skipped as well those cases when the target word was fixated. In both cases, the data showed that contextual predictability affected the use of parafoveal information.

In terms of the skipping results, the data showed that when the word in the parafovea was high predictable (i.e., 'cake') or its visually similar nonword (i.e., 'cahc'), the probability of skipping the target area was 0.11 and 0.12 respectively. This can be compared to when the word in the parafovea was low predictable (i.e., 'pies') or its visually similar nonword (i.e., 'picz') in which case, the probability of skipping the target area was much lower at 0.03 and 0.02 respectively. When the target word was anomalous, then the skipping rate was very low (probability of 0.005). These results show that the high predictable target word (as well as the corresponding visually similar nonword) was skipped significantly more than the low predictable word (as well as the corresponding visually similar nonword). This indicates that contextual predictability affects the use of parafoveal information. In particular, the suggestion is that some of the parafoveal word was identified and this was based on how predictable the prior context was. Since both high predictable words and the corresponding visually similar nonwords were skipped to a similar degree, the implication is that the decision to skip the target word was based on only identifying some part of the parafoveal word, most likely the beginning two or three letters (that is in the high predictable word and the corresponding visually similar nonword had beginning letter overlap). This suggests that in the case of the corresponding nonword, a strong context assists the reader in filling in information which is not totally accessible in the parafovea.

When the target word was not skipped, Balota et al. (1985) examined the gaze duration on these fixated targets. In the ANOVA, only subjects analysis were reported. The results showed that there was a significant main effect of contextual predictability, a significant main effect of parafoveal preview information which was that when the parafoveal preview was visually related to the target, participants spent less time on the subsequent target as well as a significant interaction between predictability and parafoveal preview information. The interaction suggested an identical or visually similar parafoveal preview produced more facilitation when the target was highly predictable from the sentence context.

Balota et al. (1985) offered several ways to assess the significant interaction. In one analysis, they averaged across the conditions where there was an identical or visually similar preview of the target (two conditions: identical and visually similar nonword) and then across the parafoveal preview conditions which were visually dissimilar to the target (the other three conditions (semantically related and visually dissimilar; visually dissimilar nonword; semantically anomalous and dissimilar). For the present discussion, the first grouping can be considered as offering a valid preview of the eventual target; and the latter grouping as an invalid preview. Table 4.2 below shows the averaged data according to these distinctions.

Table 4.2 Gaze duration in Valid and Invalid Preview Conditions with High and Low Predictable Targets in Balota, Rayner & Pollatsek’s (1985) study

	Valid preview*	**Invalid preview****	**Size of parafoveal preview benefit**
Target			
HP	240 ms	284 ms	44 ms
LP	264 ms	285 ms	21 ms

Note: * Valid preview consisted of the mean of the following two parafoveal preview conditions: identical and visually similar.
** Invalid preview consisted of the mean of the following three parafoveal preview conditions: semantically related and visually dissimilar; visually dissimilar nonword and semantically anomalous and dissimilar.
HP=high predictable; LP=low predictable

As can be seen from this Table 4.2, the size of the parafoveal preview effect differed depending on whether the target was read in a high predictable or a low predictable context. When the target was read in the high predictable context, the parafoveal preview benefit was 44ms whereas when the target was read in the low predictable context, the parafoveal preview benefit was lower at 21 ms. That is, the parafoveal preview benefit was significantly greater for high predicable words compared to low predictable words. This significant difference reflected most of the significant interaction. This result suggests that visual information from the parafovea aided later foveal processing but more so when the target was high predictable given the sentence context versus low predictable given the sentence context.

Another significant follow-up contrast compared the difference in gaze duration between high and low predictable words and when they were read in the two ‘valid’ preview conditions, identical and visually similar nonword. Results are presented in Table 4.3.

Table 4.3 Comparison of Gaze Duration across Identical and Visually Similar Preview Conditions in Balota, Rayner & Pollatsek's (1985) study

	Valid preview conditions		
	Identical*	**Visually Similar****	**Size of parafoveal preview benefit**
Target			
HP	232 ms	248 ms	16 ms
LP	264 ms	263 ms	1 ms

Note: HP=high predictable; LP=low predictable.
*Identical preview condition was where participants viewed an identical word to the target.
** Visually similar preview condition was where participants viewed a visually similar nonword to the target; the first 2 or 3 letters were identical with the target and the remaining letters were visually similar to the target

As can be seen from Table 4.3, when the target was read in the high predictable context, there was a 16 ms difference between the two conditions. However, when the target was read in a low predictable context, the identical preview was 1 ms slower than the visually similar preview (264 ms vs. 263 ms respectively). This latter contrast was nonsignificant. These findings indicates that readers processed more than the first two or three letters of the parafoveal preview when the target was high predictable. Moreover, this indicates that partial word information acquired from the parafovea is helpful in identifying the word on the subsequent fixation.

Overall the results from Balota et al.'s (1985) suggest that more preview information was extracted when a target word was high predictable compared to when it was low predictable. The indication is that when the highly predictable target word was read with a prior identical parafoveal preview ('cake') there was much greater activation, such that performance was affected, than the activation produced when the highly predictable target word was read with a prior visually similar parafoveal preview ('cahc'). Thus the suggestion is that when words are highly constrained from their context, readers are able to use more parafoveal preview information about that upcoming word than from words which are less constrained in their context. This is an important finding because it suggests extraction of parafoveal information is enhanced when it is guided by higher-level sentence context. Balota et al. (1985) argue that this finding is the strongest evidence that contextual predictability affects lexical access as opposed to having its effect on later post-lexical stages such as semantic integration.

The extent to which Balota et al.'s (1985) experiment provide support for an interactive view of the language processing system (e.g., McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1985) needs further investigation. This is because the interactive result between contextual predictability and parafoveal preview was reported in the gaze duration measure which reflects not only early lexical processing but also aspects of processing which take place after initial lexical access processes. Furthermore, Balota et al. do go on to analyse their data in the first fixation duration measure. However the only significant result in this measure was a significant main effect of parafoveal preview information where the valid preview conditions (identical and visually similar) were 14.7 ms shorter in first fixations than the invalid preview conditions (the three visually dissimilar previews conditions) (see Table 4.4). However, one possible reason that their results from the gaze duration did not reflect in the first fixation duration measure could be do with the level of contextual predictability used in the sentence materials. For their high predictable targets, the Cloze value generated from the norming Cloze task was 64%. As we have suggested in Experiment 2, this constitutes a more moderate level of contextual predictability. It is possible that a strong sentence context (as would be obtained via a very high level of contextual predictability such as in Experiment 2) that are established adequately from the prior sentence context before processing the target word would have effects on lexical access time.

Table 4.4 First Fixation Duration's in Valid and Invalid Preview Conditions with High and Low Predictable Targets in Balota, Rayner & Pollatsek's (1985) study

	Valid preview*	**Invalid preview****	**Size of parafoveal preview benefit**
Target			
HP	220 ms	236 ms	16 ms
LP	224 ms	237 ms	13 ms

Note: * Valid preview consisted of the mean of the following two parafoveal preview conditions: identical and visually similar.
** Invalid preview consisted of the mean of the following three parafoveal preview conditions: semantically related and visually dissimilar; visually dissimilar nonword and semantically anomalous and dissimilar.
HP=high predictable; LP=low predictable

4.1.5 Inhibiting parafoveal processing of a given target in moving window and boundary studies

Studies which make use of either the boundary or moving window techniques work with the underlying assumption that the normal parafoveal preview of the target is denied in the condition of 'invalid parafoveal preview of the target'. This is because instead of viewing the target, participants view some stimuli other than the target. It is common to use a string of x's, or a blank space, or visually dissimilar letters (where word shape with the target is not maintained). Thus, such stimuli are assumed to inhibit processing of the target and therefore taken to fit the criteria of the 'invalid' parafoveal preview of a particular target.

Whilst it is the case that processing an invalid parafoveal preview of the target word means that readers are prevented from parafoveally processing the *target*, they are in fact processing the invalid stimuli. This is because on a fixation, readers' process the currently fixated word as well as some aspects of the parafoveal word (Rayner, 1998). Thus past studies that have used a string of x's or visually dissimilar letters (without maintaining word shape) cannot be said to be a condition where *no parafoveal preview* was provided to the reader. Rather, these are situations of providing incorrect parafoveal information to the reader (Hand et al., 2012). Of course, the very nature of the boundary paradigm and moving window technique means that something other than the target word has to occupy the space of that target which means that this other stimuli has to be processed by the reader. However, it could be that in a natural reading situation, a row of x's or a blank area or a letter string with different word shape (compared to the target) in the parafovea is likely to draw attention because of these stimuli having an unnatural appearance in the text which means for example that the reader will want to fixate it. This is more apparent in moving window studies since with every fixation, a new window is presented with invalid information presented outside the window. In contrast, in boundary studies, only one word is altered and hence one invalid preview stimulus is seen per experimental passage. Therefore, since the parafoveal preview is processed in both valid and invalid preview conditions, care needs to be taken in terms of the information supplied to the reader in their parafovea.

One logical suggestion is that the more 'word-like' parafoveal letters strings are, the better they are since the 'more correct' information is provided to the reader (Sereno & Rayner, 2000). For example, Sereno and Rayner (2000) previously suggested that when a letter string is orthographically legal (i.e., has avoided weird letter sequences) it can be said to be word-like because it follows the rules of orthography, that is adhering to letter combinations which exist.

In Inhoff and Rayner's (1986) experiment, since the moving window was used to manipulate parafoveal preview, information was presented outside the pre-determined window size (two-word and one-word windows). This information was a row of x's to half of the participants and 'confusable' letters to the other half. In terms of the latter, the only information given in Inhoff and Rayner's (1986) paper is that the letter 'b' was replaced with the letter 'd' and the letter 'n' with the letter 'm'. No other examples are given, so it appears as if ascenders, descenders and small letters were replaced respectively with the same type of letters. However because of the limited infomration given, it is difficult to ascertain whether the resulting letter string would have been orthographically legal.

In Balota et al.'s (1985) experiment, the boundary technique was used to manipulate parafoveal preview. Targets were viewed in one of five parafoveal preview conditions (see Section 3.1.4 above) with two conditions (identical and visually similar nonword) offering the 'valid' preview of a given target. Thus, the remaining three preview conditions used a preview other than the target itself. Even though none of these conditions used a string of x's and targets and parafoveal previews of them were of the same length, there are still issues with the stimuli used in these preview conditions.

The main issue in Balota et al.'s (1985) experiment is to do with the parafoveal preview condition of visually similar nonwords. There are two points to note here. First, even though overall word shape between preview letter string and target was maintained, the first two or three letters between target and preview letter string were identical. This means that it is likely that rather than being prevented from processing the later target, readers had already begun processing some aspect of the target due to the same initial letters appearing between preview and subsequent target. Second, by replacing the remaining letters of the word with visually similar letters (from the example given, most likely ascenders with ascenders, descenders with descenders etc.), the resulting preview letter string would have been orthographically illegal. For example, for the target 'cake' the visually similar nonword letter string was 'cahcs' and for the target 'pies', it was 'picz'. In both cases 'cahcs' and 'picz' contain letter combinations which do not exist in the English language and are thus unpronounceable.

In all other preview conditions (visually dissimilar nonwords; semantically related words; anomalous preview words) word shape between target and parafoveal preview word was not maintained. Since readers can view some aspects of the initial parafoveal word, when they came to view the actual target word which was of a different overall shape to what they had processed on the previous fixation, it is possible that they would have been aware of this change in word shape and hence could have adopted strategies that are not typically used in normal reading.

Sereno and Rayner (2000) discussed the issue of the information presented in the invalid preview conditions. In particular, they offered further analyses of their results based on whether readers saw orthographically legal letter strings as the parafoveal preview of the target or, alternatively, orthographically illegal letter strings.

Sereno and Rayner (2000) manipulated word frequency and parafoveal preview (as well as a third variable, word regularity). 'Word regularity' refers to spelling-to-sound regularity refers to 'regular' words versus 'exception' words. For example, regular words such as 'time' are easily pronounced since their spelling pattern is predictable. This is because all words which end in 'ime' rhyme with each other e.g., 'dime', 'crime', 'lime' and so on. However in case of exception words such as 'good', it is difficult to predict pronunciation based on the spelling of this word. This is because word neighbours which share the words ending 'ood' i.e., 'food' or 'mood' do not rhyme. Alternatively other exception words like 'doubt' or 'weird' do not have word neighbours (other words with the same ending) so the spelling to sound mapping is specific to a given word (Sereno & Rayner, 2000). The consensus from studies using behavioural techniques like lexical decision and naming tasks is that regularity does not affect high frequency words but does affect low frequency ones. That is, there are longer responses to low frequency exception words versus low frequency regular words. Therefore this result suggests that phonological information could be important in word recognition where exception words are more difficult to process than regular words. Words which are high frequency do not show the regularity effect possibly because word meaning is accessed from the words orthography without recourse to phonology; however low frequency words need phonology to access word meaning. Sereno and Rayner (2000) were interested in investigating these effects using the eye tracking technique since this method is ecologically valid compared to behavioural methods in terms of processes which occur in normal reading.

For the parafoveal preview variable in Sereno and Rayner's (2000) experiment, targets were viewed in one of two parafoveal preview conditions, 'valid' or 'invalid'. In the valid preview condition, readers viewed a preview of the target itself. In this invalid preview condition, a letter string of equal length to the target itself was presented. These letter string were generated using pre-existing information (Mayzner & Tresselt, 1965) which gives the probability of a given letter being in a given letter position in words of four, five, and six letter length. This probability information was used to create the letter strings used in Sereno and Rayner's (2000) experiment. Letter overlap between target and its preview was avoided. However, some of the letter strings were orthographically illegal. Some of these letter strings were: 'thsg'; 'fhmt'; 'lnts'; 'ntneh'; 'rlasd'.

Sereno and Rayner (2000) presented main effect and interaction results in first fixation duration and gaze duration measures. In the first fixation duration, there was a significant main effect of frequency and preview (the main effect of regularity was nonsignificant). Both main effects were in the expected direction: readers first fixations were significantly less on high frequency than low frequency targets as well as significantly less when targets were read in a valid preview versus an invalid preview of a given target. However the interaction of these two variables (as well as the three-way interaction) were all nonsignificant. The results were similar in the gaze duration measure: there were significant main effects of frequency and preview as well as a nonsignificant interaction of frequency by preview. Therefore, Sereno and Rayner's (2000) results showing significant main effects of frequency and preview but a nonsignificant interaction of the two variables in first fixation duration lends support to an additive view of the language processing system.

Since the two-way interaction frequency by regularity was marginally significant in subjects analysis (and nonsignificant in items), Sereno and Rayner (2000) offered a separate analysis of frequency by regularity ANOVAs in each preview condition (this was in both measures and the following results are discussed for the first fixation measure). In the valid preview condition, the main effect of frequency and regularity were both significant in the expected direction. The interaction of frequency and regularity was marginally significant by subjects and nonsignificant by items. Follow-up contrasts showed that there were significant differences in first fixations between regular and exception words for low frequency words only: low frequency regular words incurred significantly less first fixations than low frequency exception words. Finally, in the invalid preview condition, the only significant effect was that of the main effect of frequency.

Therefore, Sereno and Rayner's (2000) results showed that contrary to their predictions, in the invalid preview condition there was no frequency by regularity interaction (but there was in the valid preview condition). This interaction in the valid preview condition was that low frequency regular words had a processing advantage over low frequency exception words. This was not expected since the invalid preview condition can be viewed as somewhat equal to the presentation method in lexical decision and naming tasks where words are viewed without a parafoveal preview.

One explanation of this result offered by Sereno and Rayner was that in the invalid preview condition, it was not really the case that early phonological codes were activated before directly fixating the target and that this may have had a neutralising effect on the following foveal phonological processing. As a test of this explanation, Sereno and Rayner divided the nonwords used in the invalid preview condition into pronounceable and unpronounceable words (35 and 61 respectively). They analysed gaze durations and the results in the ANOVA showed that for the pronounceable previews, there was a significant main effect of frequency but a nonsignificant frequency by regularity interaction. In contrast, for the unpronounceable previews, both the main effect of frequency and the interaction of frequency by regularity were significant. Sereno and Rayner suggested that the low frequency regularity effect with unpronounceable previews arises because there is a processing cost with low frequency exception words. Specifically, when previews are unpronounceable, they are become odd in terms of their orthography and phonology. Past research has shown that when participants view such odd words, they are likely to show low frequency regularity effects (Waters & Seidenberg, 1985). This suggests that in Sereno and Rayner's study, the odd i.e., the unpronounceable previews lead to participants being overly aware of phonology such that there was a detriment in foveal processing of low frequency exception words. Thus, when parafoveal previews in the invalid condition were pronounceable (i.e., of a 'normal' orthography and phonology) the early influence of phonological codes does not produce a conflict leading to processing costs.

There is some further evidence from earlier work by Sereno (Sereno & Rayner, 1992; Sereno, 1995) that a preview of random letters which are orthographically illegal may be disruptive. That is, when a letter string is visually distinct, attention shifts on the parafoveal word (which later turns to the foveal word) and this may invoke different processes compared to normal reading.

Sereno and Rayner (1992) developed a new paradigm called the 'fast priming paradigm' to investigate semantic priming effects during reading. The semantic priming effect refers to the faster response to a given word when the participant views a semantically related word (i.e., the prime) prior to it. The paradigm used to investigate this has largely been lexical priming tasks such as the lexical decision or naming tasks in which single words are presented. However, the technique of eye movements has several advantages over the lexical priming methods. That is, compared to lexical decision and naming techniques, eye movements have a shorter latency. There is often a delay between the time between offset of the prior sentence context and the onset of the target word, and during this delay, it is possible that participants have the time to consciously make predictions about the upcoming target word. In addition, in contrast to studies using cross-modal priming paradigms (e.g., the prime is presented heard and the target is seen visually), in the eye movement method, materials are presented in the same (visual) mode which means that there is less scope for error when calculating the processing time of prime and target. Finally, lexical priming tasks require a secondary response, typically a button-press which means that this could be disruptive to the reading process which usually proceeds uninterrupted by such overt tasks.

The fast priming paradigm makes use of the eye tracking technique and specifically uses the boundary technique while participants are engaged in the participant experimental task. Basically, when the eyes are to the left of the invisible pre-specified boundary, a random letter string appears in the target location in order to prevent parafoveal processing of the prime. When the eyes cross the boundary, a prime word is presented briefly (which is timed from the onset of the fixation, not when the boundary is crossed) and the replaced with the target word (this word stays on the screen until the participant finishes the particular text).

In the original study using the fast priming paradigm, Sereno and Rayner (1992) conducted two experiments. The aim of the first experiment was to see if priming effects typically observed in lexical priming studies could be obtained in an eye movement paradigm. The experimenters prepared single line sentence contexts with a target noun place in approximately the middle of the sentence frame. When the eyes were to the left of the invisible boundary (the last letter but one of the word before the target), a row of random letters (e.g., 'gzsd') occupied the space of the target. During the saccade that crossed the invisible boundary, the random letter string was replaced (for a particular length of time) with one of three possible prime words: the prime was either semantically related to the target (e.g., 'love') or semantically unrelated to the target (e.g., 'rule') or finally an identical condition where the actual target was shown (e.g., 'hate'). Prime durations were 60, 45, and 30 ms. The target ('hate) then replaced the prime and stayed there until the participant had finished reading the sentence.

Sereno and Rayner (1992) analysed the gaze duration on the target word. They showed that when the prime was presented for 30 ms, gaze durations were significantly faster for targets preceded by the semantically related primes compared to targets preceded by the unrelated primes (there was a 28 ms difference in gaze duration between the two prime conditions). Of importance in terms of the present discussion were the gaze duration results on the target when preceded by the identical versus semantically related primes and the identical versus semantically unrelated primes. Basically, the results were that for all three prime durations, gaze durations were significantly different for the target preceded by the identical versus semantically related prime and also significantly different for the target preceded by the identical versus the semantically unrelated prime. This result suggests that when the prime was nonidentical to the target (not 'identical'), this had a disruptive effect. There was one exception to this: when the prime was presented for 30 ms, there was a nonsignificant difference in gaze durations between the identical and the semantically related conditions. In this case, it seems that the advantage gained by priming superseded the disruptive effects of the presence of a non-identical prime.

In Experiment 2 by Sereno and Rayner (1992), the aim was investigate the 30 ms prime presentation by using prime duration just above and below 30 ms: the prime durations were either 21, 30, or 39 ms. Each target noun has three corresponding primes: as in Experiment 2, semantically related and semantically unrelated. However instead of an identical prime (i.e., the word itself), instead they used a 'random letter string' condition (e.g., 'frxe'). Again, they showed a priming effect at the 30 ms duration level where there was a 31 ms facilitation for targets preceded by related primes.

The next two sections reviews other frequency and/or contextual predictability studies subsequent to those carried out by Inhoff and Rayner (1986) and Balota et al. (1985) in order to research the stimuli researchers have used to inhibit parafoveal processing of a given target.

4.1.5.1 Word frequency by parafoveal preview studies

Five studies, including Inhoff and Rayner's (1986) have included a manipulation of word frequency and parafoveal preview, implemented either with a boundary or a moving window method (see Table 4.5). Where this relationship has not been the sole focus of the paper, it is typically because researchers have been interested in 'foveal processing load'. This refers to the finding that the amount of parafoveal preview benefit obtained depends on the difficulty of the foveal word (e.g., Kennison & Clifton, 1995; Schroyens, Vitu, Brysbaert, & Y'dewalle, 2000). In addition, the focus of Sereno and Rayner's (2000) study was the additional effect of word regularity. Reingold, Reichle, Glaholt, and Sheridan (2012) were interested in presenting a novel method to analyse results in contrast to ANOVA. They did however also analyse mean fixation durations.

Of all these studies (Table 4.5) only Reingold et al. (2012) analysed results in the single fixation duration measure. In newer studies (most likely reflecting the clearer displays of modern machine), on the majority of trials, typically only a single fixation takes place. That is, single fixations are cases where the initial fixation on the target word is not followed by an additional fixation (this suggests that single fixation duration indicates complete lexical processing). Therefore, in newer studies at least, single fixations represent the majority of the data. In some earlier studies, researchers have calculated the 'probability of refixation' which is the likelihood of making more than one fixation on the target word.

Kennison and Clifton's (1995) study looked at foveal processing load but also examined the effects of word frequency and parafoveal preview. This was by using the boundary technique. For the condition of parafoveal preview, participants viewed targets in one of two conditions: 'full-preview' or 'no-preview'. The full-preview condition offered an identical preview to the target whereas in the 'no-preview' condition, the preview offered was a visually dissimilar and unpronounceable nonword but of equal length to the target itself. These letter strings were formulated by replacing the target word with a random consonant letter string. The only example given in Kennison and Clifton's (1995) study is the letter string: 'jnslzrw'. Thus that the letter strings were orthographically illegal most probably due to the imposed limitation of not using any vowels to create the letter string.

Table 4.5 Studies which have included a manipulation of word frequency and parafoveal preview

Study	Technique used to manipualte parafoveal preview of the target	How the invalid parafoveal preview of the target was implemented
Inhoff & Rayner (1986)	Moving window	A row of x's for half the participants and for the other half, presumably ascenders, descenders, extenders and nonextenders were replaced with letters of the same type to create confusable letters; the only information given is that the letters 'b' and 'n' were replaced with the letters 'd' and 'm' respectivley. Resulting letter strings were likely to be orthographically illegal.
Sereno & Rayner (2000)	Boundary	61 of the 96 invalid preview letter strings were orthographically illegal letter strings. All previews are provided in an Appendix: some examples are: thsg; fhmt; lnts; ntneh; rlasd.
Kennison & Clifton (1995)	Boundary	A visually dissimilar, unpronounceable non-word letter string (only consonants were used) but identical in length to the target word. Only one example given of letter string : 'jnslzrw'.
Schroyens, Vitu, Brysbaert, & Y'dewalle (1999)	Boundary	Masked preview condition was used where every pixel which formed a letter of the target word scrmabled up.
Reingold, Reichle, Glaholt, & Sheridan (2012)	Boundary	Pronounceable nonwords were used. These were of identical word length to their corresponding target and letter overlap was avoided. The overall word shape betweem target and the nonword letter string was not maintained. One example is given: for the targets 'table' and 'banjo' the preview was 'purty'.

Kennison and Clifton (1995) analysed results in first fixation and gaze duration measures (both subjects and items analyses were conducted). In the first fixation duration measure, there was a significant main effect of frequency and of preview. Thus, participants first fixations were significantly less on high frequency than on low frequency targets as well as when targets were read in the valid preview versus invalid preview of the target. The interaction of frequency by preview was nonsignificant. However, results were interactive in the gaze duration measure: there were main effects of frequency and preview as well as a significant interaction between frequency and preview. Follow-ups to the significant interaction showed that there was more of a parafoveal preview benefit for high than for low frequency words. This interaction was significant in the subjects analysis but only marginally significant in items analysis. Specifically, this interaction showed that word frequency affected preview with greater preview benefit for HF than LF target words

(HF, 41 ms; LF, 8 ms). This latter result indicates that the frequency of the target word affects preview benefit where more benefit (as displayed by shorter fixation durations) is obtained from high frequency words compared to low frequency words.

Kennison and Clifton's (1995) result in the gaze duration measure therefore indicates that the relationship between frequency and preview is interactive. However because this interactive result was obtained in the middle measure gaze duration and not in the earlier first fixation data, it cannot be concluded that the relationship between frequency and preview jointly affect the same early lexical processing stage. Therefore, Kennsion and Clifton's (1995) experiment differs from Inhoff and Rayner's (1986) experiment in that the latter authors reported the interactive results in the first fixation duration measure, however Kennison and Clifton showed the interaction in the gaze duration. However, Inhoff and Rayner (1986) found that frequency and window size were additive in the gaze duration measure. It could be that this difference in results is due to the differences in the probability of refixation in the two studies. If the probability of refixation increases, this means that there is less likelihood of the first fixation measure reflecting total first pass reading time. That is, Inhoff and Rayner's (1986) study had a lower refixation probability (10.7%) than Kennison and Clifton's study (20%). This means that it is possible that the first fixation duration in Inhoff and Rayner's study was more accurate in reflecting the first pass reading time than that in the latter study. Finally, the reason that Inhoff and Rayner's study showed the interaction in the first fixation duration measure but Kennison and Clifton's study only showed it in the gaze duration measure could reflect idiosyncratic aspects of Kennison and Clifton's materials.

Schroyens, Vitu, Brysbaert, & d'Ydewalle's (1999) study was conducted with words in the Dutch language and focused on foveal processing load; however they also manipulated word frequency and parafoveal preview. Thus, targets were read in one of two preview conditions: 'visible' or 'masked'. In the visible condition, participants read the target with a preview of the target itself. In the masked preview condition, targets were masked by using a computer to randomly scrambling the pixels of every letter in the word. Thus, the resulting letter strings are likely to have not resembled actual words since letter string with 'mixed-up' pixels would have been likely to have been nonword like.

As is the norm, Schroyens et al. (1999) calculated the parafoveal preview benefit by comparing fixation durations for the target word when it was viewed in the visible condition versus the masked condition. They presented results across several measures:

single, first, and gaze durations on the target word in subjects and items analyses. However they focused largely on the gaze duration measure.

In the gaze duration measure, the researchers did not report results of the main effect of frequency. However the main effect of preview as well as the frequency by preview interaction were both significant. The interaction was significant in subjects analysis only. Follow-up contrasts to the significant interaction showed that the parafoveal preview benefit was larger for high frequency words than for low frequency words (16 ms vs. 8 ms respectively). In the single fixation duration, the only reported result relevant to the present study is that there was a significant main effect of parafoveal preview. In the first fixation duration, the only relevant reported result was that of a main effect of parafoveal preview.

Overall, results from Schroyens et al.'s (1999) study showed a reliable main effect of preview in first, single and gaze duration measures. The main effect of word frequency was only tested in gaze duration where it was shown to be significant. The interaction of word frequency and preview was only tested in the gaze duration measure where it was significant in the subjects analysis. As stated earlier, the main focus of this work was on issues to do with foveal processing load rather than the effects of word frequency and parafoveal preview hence the somewhat sparse results relating to these manipulations (also see Henderson and Ferreira, 1995).

In another study by Reingold, Reichle, Glaholt and Sheridan (2012), the effects of word frequency and parafoveal preview were the sole focus of the investigation. In Reingold's (2012) study, participants read target words in one of two preview conditions, implemented via the boundary technique. In the valid preview condition, the sentence (a single line) was presented as normal so that participants viewed a preview of the upcoming word. In the invalid preview condition, readers were presented with a pronounceable nonword which was of the same length to its corresponding target and same position letter overlap between a target and its corresponding preview were avoided. However, overall word shape was not maintained between preview letter string and target. For example, a single line sentence frame was constructed to accommodate both a high and low frequency target; 'table' and 'banjo' respectively (participants read one version). The preview of both words was 'purty'. There are no other example sentences and previews given in the study.

The focus of Reingold et al.'s (2012) study was in presenting a novel method to analyse their results (a 'survival analysis' technique). However, they also analysed results in terms of mean fixation durations in light of the fact that most experimental studies have taken this approach and thus allowing for comparisons to be made with equivalent previous research. Basically, they presented results for the main effect of frequency and the main effect of preview as well as the interaction of these variables. They analysed results in first, single and gaze durations as well as probability of skipping the target in subjects and items analyses. Results in both the fixation durations and the skipping probability showed that word frequency affected the use of parafoveal information.

In the first, single and gaze durations, results were similar: there were significant main effects of frequency and preview as well as a significant interaction between frequency and preview. All effects were in the expected direction. The significant interaction was that there was a larger parafoveal preview benefit from high frequency parafoveal targets than low frequency ones; the size of the parafoveal preview benefit in first, single and gaze durations for high and low frequency targets respectively was 37 ms vs. 26 ms; 51 ms vs. 37 ms; and 62 ms vs. 51 ms. Therefore, these results provide support for the view that more information is acquired from a parafoveal word that is high frequency than from one that is low frequency. This study was also the first to reliably show all main effects (frequency and preview) as well as a significant interaction between the two in both first and single fixation durations.

The skipping data showed that, on the probability of skipping the target word, there was a significant main effect of frequency (where high frequency targets were skipped more than low frequency ones); a significant main effect of preview (so targets in valid preview condition were skipped more than those in invalid preview condition); and a significant interaction between frequency and preview. This interaction was that while there were no significant differences in probability of skipping low frequency targets read in valid and invalid preview condition (both incurred a probability of 0.06), for the high frequency targets read in a valid preview, probability of skipping was 0.10 compared to those read in an invalid preview where probability of skipping was 0.07. Thus the size of the parafoveal preview benefit was 0.03 for high frequency targets and 0.00 for low frequency ones.

Reingold et al.'s (2012) study supports interactive models of lexical access. That is, they showed interactive fixation duration results in the first and single fixation durations as well in the skipping measure. These results oppose modular positions of lexical access and are more consistent with interactive views (e.g., McClelland & Rumelhart, 1981).

4.1.5.2 Contextual predictability by parafoveal preview studies

The only study to have examined contextual predictability and preview was that by Balota et al. (1985) and is detailed in Section 4.1.4. Table 4.6 below shows the stimuli used in the invalid parafoveal preview conditions.

Table 4.6 Past studies which have included a manipulation of contextual predictability and parafoveal preview

Study	**Technique used to manipualte parafoveal preview of the target**	**How the invalid parafoveal preview of the target was implemented**	
Balota, Rayner, & Pollatsek (1985)	Boundary	*Three 'invalid parafoveal preview conditions:*	*For the target 'cake':*
		visually dissimilar (VD) nonword	picz'
		semantically related (SR) word	pies'
		anomalous (AN) word	bomb'

4.1.5.3 Word frequency by contextual predictability by parafoveal preview studies

Of particular interest was in the study which has manipulated both word frequency and predictability of target words along with parafoveal preview in the one study (Hand, O'Donnell, & Sereno, 2010). Hand et al.'s (2010) study focused on the effect of parafoveal information to the frequency and predictability relationship and they used launch distance as a metric of the extent of parafoveal processing.

Hand et al. (2010) directly examined the effects of orthogonally varying word frequency and contextual predictability in an eye movement study. They were also interested in the effect of parafoveal preview to the frequency by predictability interaction. For the variable of parafoveal preview, rather than use for example the boundary technique where the target word would be read with a valid and invalid parafoveal preview, they analysed their results in post-hoc analyses using launch distance to index the extent of parafoveal processing (see Table 4.7). Their reasoning behind this was based on the fact that visual acuity drops off because of retinal eccentricity. Thus, it is reasonable to assume that the amount of parafoveal preview obtained by readers is affected in part by the pre-target launch distance. This assumption is that the more the distance from the target word, the less of a preview obtained from the target by readers.

Table 4.7 Word frequency and contextual predictability and parafoveal preview manipulations in past studies

Study	Technique used to manipualte parafoveal preview of the target	How the invalid parafoveal preview of the target was implemented
Hand, Sereno, & O'Donnell (2010)	Parafoveal processing was examined post-hoc and was indexed by launch distance.	n/a

Note: n/a = not applicable

Launch distance was measured as the distance between the beginning of the target word (that is, the space before the target) to the location of the previous fixation. There were three levels of launch distance (near: 1-3 characters; middle: 4-6 characters; and far: 7-9 characters). Hand et al. (2010) suggested that the closer the prior fixation to the target word, the more preview of the target word the reader would have prior to fixating it. Thus, a lesser distance, ‘near’ (a distance of 1-3 characters from the target word) gives more of a preview than ‘middle’ (a distance of 4-6 characters from the target word) and both give more preview than ‘far’ (a distance of 7-9 characters from the target word). In general, the ‘near’ distance can be considered as offering the reader a valid preview of the parafoveal target and the ‘far’ distance as yielding an invalid parafoveal preview of the upcoming word.

Hand et al. (2010) hypothesised that their results would show a launch distance effect where there would be longer target fixations the greater the launch distance. Their results in this three-way design were similar in first, single and gaze duration measures. Specifically, in all three measures, the main effects of frequency (shorter fixation on high than low frequency targets), predictability (shorter fixations on high than low predictable targets) and preview were significant. Follow-up tests to the significant main effect of preview showed that the closer the launch distance, the shorter the fixation time on the target word; all contrast were significant: near vs. middle; near vs. far; and middle v. far. This result suggests that shorter launch distances gives way to greater parafoveal previews and thus there are shorter fixation times on the target. In terms of the interactions, ‘frequency by predictability’ was nonsignificant; however the other two-way interactions ‘frequency by preview’ and ‘predictability by preview’ were significant, as well as the three-way interaction. For the significant three-way interaction, Hand et al. (2010) conducted separate frequency by predictability ANOVAs at each level of preview: near, middle and far, conditions. The results here showed an interactive pattern in the ‘near’ and ‘middle’ conditions and an additive pattern in the ‘far’ condition. Specifically, in the ‘near’ condition, there were main effects of frequency and predictability and a significant interaction in which there was a larger predictability effect for low than high frequency targets. In the ‘middle’ condition, only the main effect of frequency was significant; the interaction was significant but there was an opposite pattern to that of the significant interaction in the ‘near’ condition: there was a larger predictability effect for high than low frequency targets. In the ‘far’ condition, the only significant effect was that of the main effect of frequency.

Hand et al. argued that their initial additive results of frequency and predictability were the result of a combination of the results in the three-way interaction: the interaction in frequency by predictability and launch distance in the ‘near’ and ‘middle’ launch distances and an additive pattern of results in the ‘far’ condition. They suggested that the overall pattern of results showed that the greater the launch distance, the less the effect of parafoveal preview. Therefore, overall they showed additive effects of frequency and predictability, but argued that these factors can exert interactive effects – but this interaction is dependent on parafoveal preview.

4.1.6 Present study

There is an issue concerning the information presented to participants in the invalid parafoveal preview of the target condition. As can be seen in Tables 4.5 and 4.6 none of the studies have used orthographically legal letter strings of equal length to the corresponding target whilst avoiding same letter position overlap but maintaining overall word shape with the target. These criteria are important to fulfil in order to make the assumption that parafoveal processing of a given target was inhibited on the pretarget fixation. Only one study has fulfilled most of these criteria (Reingold et al. 2012); the only omission being that that preview and target did not match in their overall word shape. Also contextual predictability was not manipulated in their study.

Only one study has examined the joint effects of word frequency and contextual predictability along with parafoveal preview (Hand et al., 2010). However, the effects of parafoveal processing were examined in post-hoc analyses using launch distance to the target as a metric of the extent of parafoveal processing. Thus, fixation durations in valid and invalid parafoveal preview conditions offers scope for a further study.

Therefore the aim of Experiment 3 was to block the parafoveal access of the parafoveal target word using the boundary technique by using better previews than has been the case in all prior studies. This was in order to explore the possibility that some of the observed interaction in the first and single fixation duration measures in Experiment 2 could have been due to the parafoveal processing aspect of reading high frequency (Inhoff & Rayner, 1986) and high contextual predictability targets (Balota et al., 1985). The boundary technique prevents readers from parafoveally obtaining information about the eventual target. For example, this technique can be used to block parafoveal access of high frequency words and of high predictable words. This is because a letter string other than the target word itself occupies the space of the target and only changes to the target when the readers eyes are on the pre-specified location. Therefore the foveal aspects of the frequency-predictability interaction observed in early fixation duration measures in Experiment 2 could be examined. A secondary aim was that by using the boundary technique, the rate of reading would be slowed (e.g., Rayner et al., 1982) and thus the possible floor effect suggested in Experiment 2's data could be further investigated.

The parafoveal preview of the target word used in the invalid preview condition were designed in an attempt to provide no useful information to the reader about the actual target word while not misleading (i.e., garden-pathing readers) them either. It was important to avoid letter overlap between each letter of the parafoveal letter string and the target word so that it could be said that readers were inhibited from acquiring lexical information from the parafoveal word. Thus, the present study was the first to use parafoveal previews which deliberately tried to be less disruptive as possible than has been the case in past studies of preview and frequency and or contextual predictability.

In the present study, the same materials as those in Experiment 2 were used with a different group of 40 participants. Thus, as in Experiment 2, the frequency (low frequency, high frequency) and contextual predictability (low, medium and high predictability) of target words was manipulated, with these words being embedded in a two-line sentence frame. All targets were presented using the boundary change method (Rayner, 1975). That is, a letter string occupied the space of the target word, with it changing to the target itself when the participants' eyes cross a pre-specified location (one letter space before the first letter of the letter string). The letter string was a pseudoword and was a pronounceable (orthographically legal) nonword letter string which shared the same overall shape with its corresponding target word). In this way, parafoveal processing (during fixations on the pre-target word) was inhibited and frequency and predictability effects could be evaluated based only on information obtained from viewing the target for the first time (i.e., foveally viewing the target). From this manipulation, it was assessed how much of the effects of the frequency-predictability interaction observed in Experiment 2 is due to parafoveally processing the target words.

The results will be analysed as in Experiment 2 where mean fixation durations (FFD, SFD, GD, TT) as well as percentage of times the target was skipped (%Skip) across the six experimental conditions was calculated. It was expected that target fixation time measures should reflect the foveal component of lexical access related to the frequency by predictability manipulations.

Inhoff and Rayner's (1986) study showed that when parafoveal previews were denied, there was no frequency effect in that invalid preview condition in first fixation durations but there was in the gaze duration measure. Balota et al.'s (1985) study, in the invalid preview conditions, showed a nonsignificant main effect of contextual predictability in first fixation duration and gaze duration measures. Hand et al.'s (2010) study showed that in single fixation durations, in the invalid preview condition (corresponding to when launch site was 7-9 characters from the target), there was a significant main effect of frequency but a nonsignificant main effect of predictability as well as of the interaction). Thus the results from these past studies are somewhat mixed. In the present study, where participants read targets without seeing a valid preview of the parafoveal word, we regarded first and single fixation measures to reflect lexical access processes. In both measures, we expected a significant main effect of word frequency and of contextual predictability and a nonsignificant interaction. The main effect of word frequency (where readers spend less time on high than low frequency words) was expected because even when the parafoveal access of the target is blocked, the frequency effect has been shown to be robust: many later studies including Hand et al'.s (2010) have shown that that readers spend less time fixating high than low frequency targets (also shown in Reingold et al.'s 2012 study). We expected a main effect of contextual predictability because the experimental passages used in the present study manipulated contextual predictability better than in past studies. Both Balota et al. (1986) and Hand et al. (2010) studies did not find a significant predictability effect (where participants fixate high predictable targets less than low predictable targets). However, both studies had moderate Cloze values for their high predictable contexts: 0.64 in Balota et al.'s study and 0.53 and 0.60 for high and low frequency targets respectively in Hand et al.'s study. In contrast, in the present study, the Cloze values were 0.97 and 0.96 for high and low frequency targets respectively.

In terms of the interaction between word frequency and contextual predictability, the study by Hand et al. (2010) is more similar to the present study since in both studies, both variables were manipulated in the one study. Hand et al.'s (2010) study suggests that when parafoveal access of targets are blocked (that is, read with an invalid preview), additive results between words frequency and contextual predictability should be obtained. That is, word frequency and contextual predictability are unlikely to interact when less of the upcoming (parafoveal) word is viewed. If it is the case that word frequency and contextual predictability only interact when there is sufficient preview information available to the reader (obtained on the prior fixation), as suggested by Hand et al., then a word frequency by contextual predictability interaction is not expected in early fixation measure in the present study because the upcoming target word was not viewed at all. This, in the early fixation duration measures, first fixation and single fixation duration, it was predicted that there would be main effects of word frequency and contextual predictability but no significant interaction.

The gaze duration measure is likely to reflect processing activities between initial lexical access to later processing, whereas the total time measure is likely to capture only later processing activities. In these measures, then which are likely to capture processing subsequent to initial lexical access, we expected that the additive pattern of results would be maintained (expected in the early measures). Thus, it was predicted that there would be main effects of frequency and predictability and a nonsignificant interaction.

We were unsure how results would be in terms of the skipping rates (as indicated by percentage of words which were skipped in each condition) because it is unclear what causes skipping in normal reading alone. One reason that skipping occurs in normal reading is because readers are able to extract information from not only the currently fixated word but also enough information about the parafoveal word such that it is fully identified on the prior fixation which means that there is no reason to fixate it on the next forward going fixation i.e., it is skipped. This reasoning suggests that, in the present study, skipping rates are likely to be very low because the initial parafoveal access of the word means that participants would initially process a nonword letter string which would potentially ‘stand out’ as not being the word which fits in with the current passage they are reading. Thus, readers would want to fixate that parafoveal word on the next forward going fixation. It is possible then that skipping rates would be very low.

Experiment 3

4.2 Method

4.2.1 Participants

Forty participants (30 female, 10 male) took part in the experiment (none of whom had taken part in Experiment 2). Participants were native English speakers and members of the University of Glasgow undergraduate and postgraduate community (mean age 21 years). Participation was voluntary and payment of the normal hourly rate of £6 was given or course credits. None of the participants has a diagnosis of any reading disorder and all the participants had normal or corrected-to-normal vision.

4.2.2 Apparatus

As in Experiment 2.

4.2.3 Materials

Passages of text were the same materials as those presented in Experiment 2 which were 150 passages of text with 25 passages per condition. An invalid parafoveal preview of the target was presented in the location of the target; this letter string changed to the target itself once the participants made a saccade over the pre-defined boundary location (the last letter of the pre-target word). A complete list of the materials is shown in Appendix C (see also Appendix D and Table 3.1 for summary specifications of targets).

4.2.4 Design

The design was identical to that in Experiment 2; a within-subjects 2 (frequency: LF, HF) x 3 (contextual predictability: LP, MP, HP) was used which lead to the creation of 6 conditions: LF-LP, LF-MP, LF-HP and HF-LP, HF-MP, HF-HP. With a total of 150 sentences, there were 25 items per condition where all 40 participants were presented with all the target words. All the targets were presented using the boundary change technique so that participants always viewed an invalid preview of the target prior to viewing the target.

4.2.5 Procedure

When a participant arrived to take part in the eye tracking experiment, the forehead and chin rest were adjusted in order that they were placed at the appropriate height for the participant. They were provided with written as well as verbal instructions about the task they were going to be doing. They were instructed to read the passages as they would read a story in for example a newspaper or magazine i.e., reading for meaning. They were told that in order to ensure this, the computer would randomly present questions requiring a yes or no answer. Following this, the calibration of the eye tracker was carried out which took approximately 5 minutes. The calibration involved the participant following a set of dots presented at various locations on the computer screen; basically the calibration points extended over the full horizontal and vertical range over which the passages were presented.

Participants read five practice trials followed by 150 experimental trials. The procedure for both practice and experimental trials was as follows. Before each passage of text appeared on the screen, participants looked at a black box on the left side of the screen; this was the first character position of the sentence. When participants fixated on this leftmost box, the full passage of text spanning the two lines was presented on the screen. When participants had finished reading the text, they looked away to the bottom-right of the screen and pressed a button on the joypad to clear the screen for the next trial. The leftmost black box re-appeared immediately or, if the passage had a corresponding yes/no comprehension question, then the participant would answer it first before the black box re-appeared. A total of 48 questions appeared on a random basis, half with yes answers and half with no. To record their answer, the joypad was used where a left click recorded the answer ‘no’ and a right click was for a ‘yes’ answer. There was no difficulty in answering the questions. Throughout the duration of the experiment, recalibration was performed whenever necessary.

Since the boundary technique was used to present the targets, on all the trials, an invalid parafoveal preview of the target was presented in the location of the target; this letter string changed to the target itself once the participants made a saccade over the pre-defined boundary location. This was between the last letter of the word preceding the target and the space before the target.

Since the purpose of using an invalid parafoveal preview of a given target word was to deny parafoveal processing of the target (that is, to ensure only foveal processing), and based on the view that whatever stimulus appears in the parafovea is processed, we thought it necessary that the letter string resemble the real target as much as possible, without actually being the target, so for example avoiding letter overlap. Therefore each letter string was constructed based on each letter maintaining visual similarity with the resulting preview letter. Visual similarity of letters was based on ascenders replacing ascenders, descenders replacing descenders and ‘all-other small’ letters replacing ‘all-other small’ letters. Ascenders are letters which extend above the line (b, d, f, h, k, l, t) while descenders are those letters that extend below the line (g, j, q, p, y). All other small letters are (a, c, e, i, m, n, o, r, s, u, v, w, x, z). Therefore each letter was replaced in the target with a visually similar letter. This procedure meant that word shape between preview and target was maintained. However, unlike past studies, we also ensured that the resulting letter string was pronounceable. For example, in English orthography some letters never appear in certain combinations, such as ‘bz’, or ‘tfjk’. In using letter strings with such illegal orthographic combinations of letters, it is possible that they will command attention since they are not natural in the language.

A break was given to participants half way through the experimental trials. In total, the experiment took approximately 1 hour to complete. After the experiment, participants were asked if they had noticed anything unusual. Most reported seeing words change, but no one could say exactly what the letters were prior to changing. Participants were debriefed before leaving the booth.

4.3 Results

We examined first fixation duration (FFD), single fixation duration (SFD), gaze duration (GD) and total time (TT) on the target word. We also examined target skipping rates (%Skip; percentage of times the target was skipped). The target region included the space before the target as well as the target word itself. Data were excluded from analyses for the following reasons: (a) a blink or track loss occurred on the target word; (b) the display change was triggered on the word before the target by a saccade overshooting into the target region.

For each eye movement dependent measure as well as the skipping the target word, a 2 (frequency: LF, HF) x 3 (contextual predictability: LP, MP, HP) within-subjects repeated measures analyses of variance (ANOVAs) was carried, both by subjects (F_1) and by items (F_2). The follow-up tests were Bonferroni multiple comparisons corrections.

4.3.1 Frequency by predictability analyses

The results for each measure (FFD, SFD, GD, TT, as well as %Skip) are presented below, describing main effects (frequency, predictability), follow-up contrasts for Predictability (LP vs. MP vs. HP) when it was significant, the interaction (frequency x predictability), as well as follow-up contrasts of the interaction when it was significant.

Mean fixation times across the six experimental conditions are presented in Table 4.8. Means with standard error bars are displayed in Figure 4.1 for the SFD measure. Since past studies have not analysed SFDs on targets but rather presented the FFDs on targets, Figure 4.2 presents the mean (with standard error bars) for the FFD.

Table 4.8 Mean Fixation Time Results in Milliseconds (Standard Deviation) and Percentage of Times the Target Word was Skipped in the Conditions

Measure		**LF**				**HF**	
	LP	**MP**	**HP**		**LP**	**MP**	**HP**
FFD	247 (26)	250 (22)	235 (28)		238 (27)	236 (26)	228 (19)
SFD	258 (29)	257 (25)	244 (32)		242 (27)	242 (29)	232 (21)
GD	289 (44)	279 (34)	271 (54)		267 (40)	267 (38)	250 (32)
TT	326 (84)	321 (96)	305 (99)		305 (71)	297 (82)	279 (68)
%Skip	15 (10)	10 (9)	10 (7)		13 (9)	14 (9)	11 (8)
Sent. 1	27 (5)	28 (5)	28 (5)		28 (5)	27 (5)	27 (5)

Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; TT = total time including regressions; %Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Figure 4.1. Average single fixation time (SFD) as a function of frequency and contextual predictability (with an invalid preview)

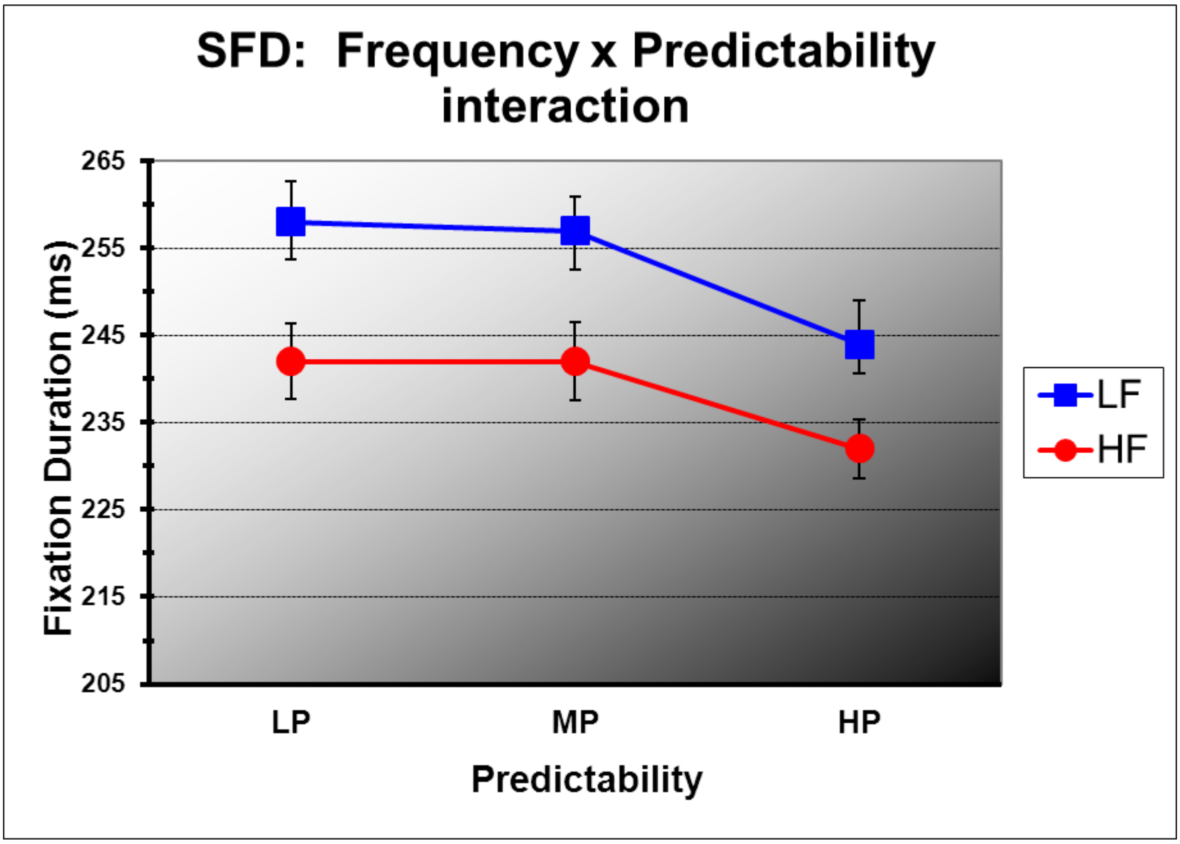
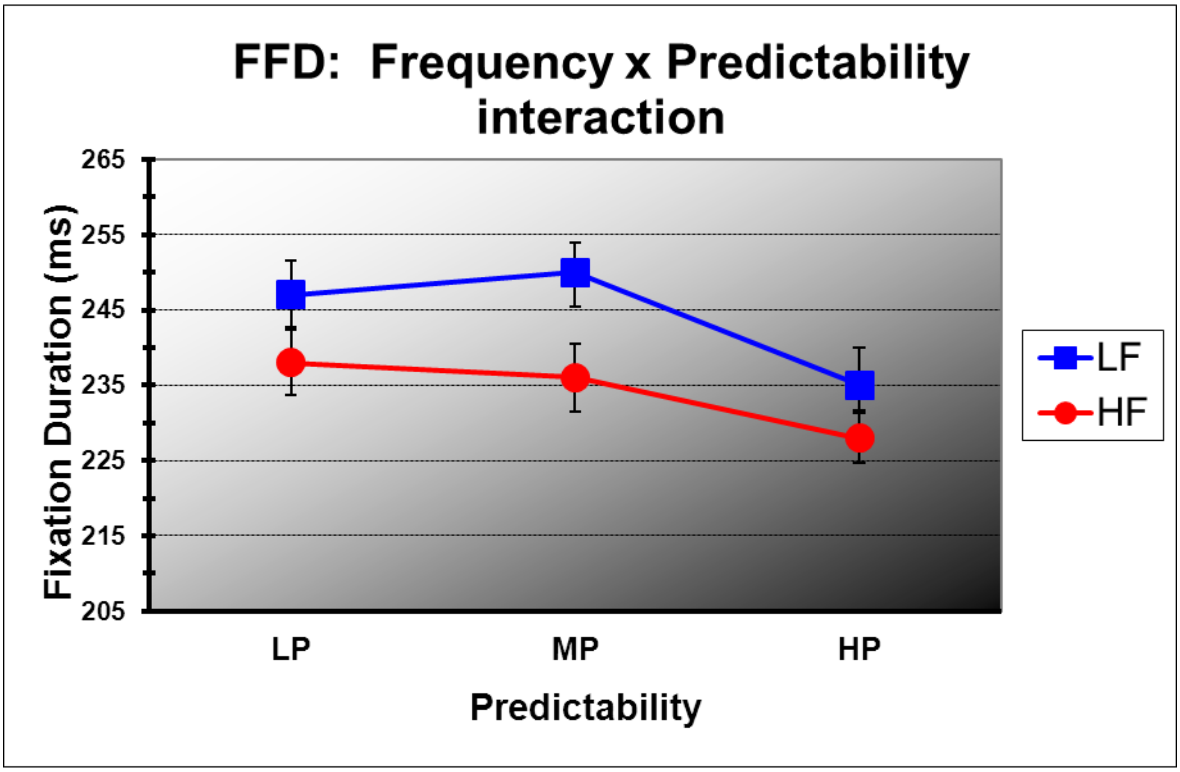


Figure 4.2. Average first fixation duration (FFD) as a function of frequency and contextual predictability (with an invalid preview)



Figures 4.1. and 4.2. Note: ms = milliseconds; FFD = first fixation duration; SFD = single fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

4.3.1.1 First fixation duration

The FFD means are shown in Table 4.8. Figure 4.2 shows the plotted interaction, that is the means (with standard errors). Basically, there were significant main effects of word frequency and contextual predictability and nonsignificant interaction. That is, in the resulting ANOVA, there was a main effect of frequency [$F_1(1,39)=29.63$, $MSE=196$, $p<.001$, $F_2(1,24)=19.60$, $MSE=183$, $p<.001$] . HF targets were fixated for significantly less time than LF targets (234 ms vs. 244 ms). The main effect of contextual predictability was significant [$F_1(2,78)=16.07$, $MSE=216$, $p<.001$, $F_2(2,48)=11.78$, $MSE=180$, $p<.001$]; follow-up Bonferroni contrasts (LP vs. MP; LP vs. HP; and MP vs. HP) showed that of the three comparisons, the latter two were significant. That is, readers' first fixations were significantly shorter on HP vs. LP targets (232 ms vs. 243 ms respectively; $F_1=23.84$, $p<.001$, $F_2=19.46$, $p<.001$] as well as on HP vs. MP targets [232 ms vs. 243 ms; $F_1=24.37$, $p<.001$, $F_2=15.68$, $p<.001$). Readers' had the same first fixations on LP and on MP targets (both 243 ms; all Fs<1). Thus, readers' first fixations were shorter when the target was in a HP (232 ms) context compared to when the target was in a LP (243 ms) and a MP (243 ms) context. There were no differences in readers' first fixations on targets in a LP (243 ms) versus in a MP (243 ms) context. Finally, the interaction of frequency x context was nonsignificant [$F_1(2,78)=1.49$, $MSE=197$, $p>.20$, $F_2(2,48)=1.09$, $MSE=185$, $p>.30$].

4.3.1.2 Single fixation duration

SFD means are also shown in Table 4.8 (displayed in Figure 4.1.). Results in the SFD measure followed the same pattern as the FFD measure, that is, there was also an additive result. There were significant main effects of word frequency and contextual predictability and a nonsignificant interaction. In the ANOVA, there was a significant main effect of frequency [$F_1(1,39)=46.23$, $MSE=258$, $p<.001$, $F_2(1,24)=32.88$, $MSE=224$, $p<.001$]. HF words were fixated for significantly less time than LF words (239 ms vs. 253 ms). There was also a significant main effect of predictability [$F_1(2,78)=15.83$, $MSE=229$, $p<.001$, $F_2(2,48)=9.11$, $MSE=233$, $p<.001$]; follow-up Bonferroni contrasts (LP vs. MP; LP vs. HP; and MP vs. HP) showed the same pattern of results as in the FFD measure in that it was the latter two comparisons which were significant. Specifically, readers' single fixations were significantly shorter on HP (238 ms) versus both LP targets (250 ms) [$F_1=24.13$, $p<.001$, $F_2=15.38$, $p<.001$] and MP targets (250 ms) [$F_1=23.36$, $p<.001$, $F_2=11.69$, $p<.01$). As in the FFD measure, in the SFD measure, readers had the same SFDs on targets appearing in LP and MP contexts (both 250 ms; all *F*s<1). Thus, readers' single fixations were shorter when the target was in a HP (238 ms) context versus in both a LP (250 ms) and a MP (250 ms) context. There were no differences in readers' single fixations on targets in a LP (250 ms) versus in a MP (250 ms) context. Finally, the word frequency x contextual predictability interaction was nonsignificant (this is displayed in Figure 3.1) [all *F*s<1].

4.3.1.3 Gaze duration

The GD data are also shown in Table 4.8. As in the FFD and SFD measures, there was a significant main effect of word frequency and contextual predictability and nonsignificant interaction. That is, there was significant main effect of frequency [$F_1(1,39)=35.92$, $MSE=539$, $p<.001$, $F_2(1,24)=27.49$, $MSE=224$, $p<.001$]. HF targets were fixated for significantly less time than LF targets (261 ms vs. 280 ms). There was a significant main effect of predictability [$F_1(2,78)=12.64$, $MSE=521$, $p<.001$, $F_2(2,48)=12.30$, $MSE=340$, $p<.001$]; follow-up Bonferroni contrasts (LP vs. MP; LP vs. HP; and MP vs. HP) showed that, like the SFD and FFD measures, the latter two contrasts were significant. Specifically, readers' gaze durations were significantly shorter on HP targets (261 ms) versus MP targets (273 ms) [$F_1=12.05$, $p<.001$, $F_2=10.35$, $p<.01$] as well as LP targets (278 ms) [F_1=23.87, p<.001, F_2=23.80, p<.001]. The difference between LP vs. MP targets (278 ms vs. 273 ms) was nonsignificant in the subjects analysis and showed a trend for significance in the items analysis [F_1=2.00, p>.15, F_2=2.76, p=.103]. Thus, gaze durations were shorter when the target was in a HP (261 ms) context versus in both a LP (278 ms) and a MP (273 ms) context. There were no differences in readers' gaze durations on targets in a LP (278 ms) versus in a MP (273 ms) context. In terms of the interaction, as in the two previous measures SFD and FFD, the frequency x predictability interaction was nonsignificant [F_1(2,78)=1.02, MSE=557, p>.35, F_2<1].

4.3.1.4 Total time

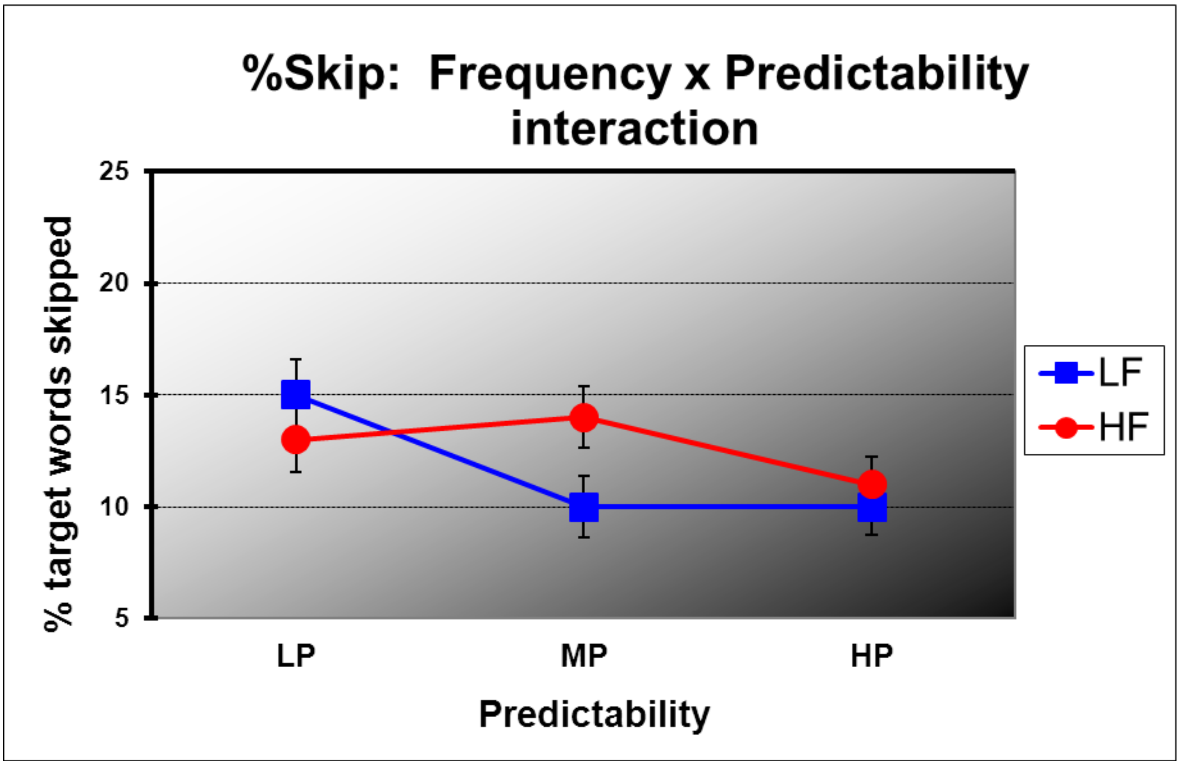
In the total time (TT) measure, there were significant main effects of word frequency and contextual predictability and a nonsignificant interaction. The significant main effect of frequency [F_1(1,39)=16.16, MSE=2074, p<.001, F_2(1,24)=13.41, MSE=1032, p<.01] was that HF targets were fixated for significantly less time than LF targets (294 ms vs. 317 ms). For the significant main effect of predictability [F_1(2,78)=5.55, MSE=2049, p<.01, F_2(2,48)=4.13, MSE=1488, p<.05], follow-up Bonferroni contrasts (LP vs. MP; LP vs. HP; and MP vs. HP) showed an overall similar result as in the previous measures. Readers' TT on HP targets (292 ms) was both significantly less than TT on MP targets (309 ms) though this effect was marginally significant in the items analysis [F_1=5.34, p<.05, F_2=3.93, p=.053] and also significantly less on LP targets (316 ms) [F_1=10.46, p<.01, F_2=7.80, p<.01]. The difference in total time spent on LP vs. MP targets (316 ms vs. 309 ms respectively) was nonsignificant [all *F*s<1]. The frequency x predictability interaction was nonsignificant [all *F*s<1].

4.3.1.5 Percentage of times the target was skipped

The %Skip means are shown in Table 4.8. Basically, the main effect of frequency was not significant (results were marginally significant in subjects analysis and trend in items analysis); there was a significant main effect of contextual predictability and a significant interaction though the interaction effect was trend in items analysis (displayed in Figure 4.3.). That is, the main effect of frequency was marginal in the subjects analysis and showed a trend for significance in the items analysis [F_1(1,39)=3.08, MSE=35, p=.087, F_2(1,24)=2.19, MSE=30, p<.151]. Thus, there were no reliable significant differences in skipping rates in terms of whether a word was low or high frequency. The predictability main effect was significant [F_1(2,78)=6.50, MSE=35, p<.01, F_2(2,48)=3.20, MSE=44, p<.05]; follow-up Bonferroni contrasts showed that LP targets were skipped significantly more than HP targets (14% vs. 11%: F_1=13.01, p<.001, F_2=6.40, p<.05). The other contrasts (LP vs. MP and MP vs. HP) were not significant, at most marginally significant results were observed [LP vs. MP: 14% vs. 12%, F_1=3.16, p=.080, F_2=1.55, p>.20; MP vs. HP: 12% vs. 11%, F_1=3.35, p=.071, F_2=1.65, p>.20]. Therefore, the only reliable significant differences in skipping rates was that readers' skipped a HP target 11% of the time versus a LP target which was skipped 14% of the time – a result in the opposing direction to what we expected, since the usual case is that high predictable targets are skipped more than low ones.

In terms of the interaction (see Figure 4.3.) the interaction of word frequency by contextual predictability was somewhat significant: the interaction was significant in the subjects analysis only, showing only a trend for significance in the items analysis [$F_1(2,78)$=4.38, *MSE*=40, p<.05, $F_2(1,48)$=2.98, *MSE*=53, p=.136].

Figure 4.3. Word frequency by contextual predictability interaction in percentage of times target was skipped (%Skip)



Note: %Skip = percentage of times target was skipped; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

Follow-up contrasts to the significant frequency by predictability interaction examined the effect of frequency (LF vs. HF) at each level of predictability (LP, MP, HP) and also the effect of predictability for LF words and for HF words. There were several results in unexpected directions. The frequency contrasts (LF vs. HF) at MP was the only significant difference [F_1=10.03, p<.01, F_2=4.77, p<.05]; thus in the MP condition, HF targets were skipped more than LF targets (14% vs. 10%). The other two frequency contrasts in the LP and HP conditions were nonsignificant (LP: all Fs<1; HP: all Fs<1). Moreover (though nonsignificant) in the LP condition, the frequency effect was in the unexpected direction: LF words were skipped 15% of the time, more than HF words which were skipped 13% of the time. The predictability contrasts for LF words compared the difference in skipping for LP versus MP targets (15% vs. 10% respectively); LP vs. HP (15% vs. 10% respectively), and MP versus HP (both 10%). It should be noted that the first two contrasts are in the unexpected direction, moreover these differences were significant That is, LP targets were skipped more than MP targets [F_1=10.48, p<.01, F_2=4.99, p<.05] as well as more than HP targets [F_1=10.03, p<.01, F_2=4.77, p<.05]. The percentage of skipping incurred by MP versus HP targets was the same (10%; all Fs<1). Therefore LF targets appearing in LP contexts were skipped 15% of the time and this was significantly more than LF targets appearing in both MP (10%) and HP contexts (10%); MP and HP targets were skipped the same amount. The equivalent comparisons were carried out for HF words appearing in LP, MP and HP contexts: for HF targets we compared percentage skipping of LP vs. MP (13% vs. 14%); LP vs. HP (13% vs. 11%) and MP vs. HP (14% vs. 11%). As can be seen, this time the latter comparison was not only in the unexpected direction but also a significant difference; the first two comparisons were in the expected direction but were nonsignificant. Thus, the results here were there were no significant differences between LP and MP words [all Fs<1]. The difference in skipping between LP versus HP words was only trend for significance in the subjects analysis and nonsignificant in the items analysis [F_1=2.34, p=.126, F_2=1.14, p>.25]. Finally, the difference between MP and HP words was significant in subjects analysis and marginal in items analysis [F_1=6.07, p<.05, F_2=2.89, p=.096]. Therefore, with HF words appearing in LP, MP and HP contexts, the only somewhat reliable difference is that of the last comparison which was in the unexpected direction: MP words were skipped 14% of the time and this was more than how often HP targets were skipped, which was 11% of the time.

4.3.1.6 Sentence 1 first pass time across the conditions

For the first line of the passage (i.e., sentence 1), we also calculated the first pass time (ms per character) across the six conditions (see row 8 in Table 4.8). This analysis showed that the main effect of frequency as well as the main effect of contextual predictability were both nonsignificant [all Fs<1]. Thus, there was a nonsignificant difference in first pass time when the later target was HF (27.3 ms) or LF (27.6 ms). The interaction of word frequency and contextual predictability was also nonsignificant [$F_1(2,78)=1.33$, $p>.25$; $F_2<1$]. This means that one level of one factor did not significantly differ from another level of a second and/or third factor, thus there were no significant differences in fixation time on line 1 across the 6 conditions.

4.3.2 Results summary

The results are summarised in Table's 4.9, 4.10, 4.11 and 4.12. Our hypotheses in the first fixation duration measure (FFD) and single fixation duration measure (SFD) were supported. In these early measures, similar results were observed so both are discussed together here. Basically, in both fixation measures, an additive result was observed (main effects of word frequency and contextual predictability and a nonsignificant interaction). The significant main effect of frequency in all the measures was that high frequency words were processed faster than low frequency words. The significant main effect of contextual predictability was that participants processed HP targets faster than both MP and LP targets. Fixation times on LP and MP targets were the same.

In the gaze duration (GD) measure, our hypotheses were supported and the additive results seen in the first and single fixation measures were maintained. Thus, as in the single and first fixation measures, the GD measure showed significant main effects of word frequency and contextual predictability and a nonsignificant interaction. The contextual predictability main effect was very similar to that obtained in the first and single fixation measures. That is, participants processed HP targets faster than both MP and LP targets. There were nonsignificant differences in the GD's on LP and MP targets.

In the total time measure (TT), an additive result was further maintained. Thus there was a significant main effect of word frequency and contextual predictability and a nonsignificant interaction. The main effect of contextual predictability was in the same directions as that in the gaze duration measure. Specifically, readers processed HP targets faster than MP targets, thought was marginal in items; readers also processed HP targets than LP targets.

In the percentage of times the target was skipped measure (%Skip), results differed to the additive results observed in fixation duration measures, FFD, SFD, GD and TT. That is, in the case of skipping the target word, the word frequency main effect was nonsignificant; the contextual predictability main effect was significant and the interaction of word frequency and contextual predictability was significant, though this was by subjects; in items this effect showed a trend for significance. Follow-up tests to the significant main effect of contextual predictability showed that one comparison was significant and this (unexpected direction) was where participants skipped significantly more LP words than HP words. Follow-ups to the significant interaction showed that in the frequency contrasts (LF vs. HF) for each level of predictability, the only significant contrast was that HF words were skipped more than LF ones in the MP condition. Moreover, even though nonsignificant, in the low predictable condition, the frequency effect was in the unexpected direction where LF targets were skipped more than HF ones. In the HP condition, the frequency effect was in the expected direction but nonsignificant. In the predictability contrasts for the LF words, two results were in the unexpected direction: LP targets were skipped significantly more than MP and HP targets and MP and HP targets incurred the same rate of skipping. Normally, higher predictable words are skipped more often than lower predictable words. For the HF words, this time LP targets displayed the expected direction but were nonsignificant: LP targets were skipped fewer than MP and HP ones. The unexpected direction of result was that the MP targets were skipped significantly more than HP ones (though marginal in items).

Table 4.9 Summary of ANOVA Results in all Fixation Duration Measures and Skipping by Participants (F_1) and by Items (F_2)

Main effect frequency						
Measure			***df***	***F***	***MSE***	***p***
FFD	**LF > HF**	F_1	1,39	29.63	196	<.001
		F_2	1.24	19.6	183	<.001
SFD	**LF > HF**	F_1	1,39	46.23	258	<.001
		F_2	1,24	9.11	233	<.001
GD	**LF > HF**	F_1	1.39	35.92	539	<.001
		F_2	1,24	27.49	224	<.001
TT	**LF > HF**	F_1	1,39	16.16	2074	<.001
		F_2	1,24	13.41	1032	<.01
% Skip	**LF vs. HF**	F_1	1,39	3.08	35	=.087
		F_2	1,24	2.19	30	<.151
Sent. 1	**LF vs. HF**	F_1	1,39	*<1*		*ns*
		F_2	1,24	*<1*		*ns*
Main effect contextual predictability						
FFD		F_1	2,78	16.07	216	<.001
		F_2	2,48	11.78	180	<.001
SFD		F_1	2,78	15.83	229	<.001
		F_2	2,48	9.11	233	<.001
GD		F_1	2,78	12.64	521	<.001
		F_2	2,48	12.3	340	<.001
TT		F_1	2,78	5.55	2049	<.01
		F_2	2,48	4.13	1488	<.05
% Skip		F_1	2,78	6.5	35	<.01
		F_2	2,48	3.2	44	<.05
Sent. 1		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
Interaction word frequency x contextual predictability						
FFD		F_1	2,78	1.49	197	>.20
		F_2	2,48	1.09	185	>.30
SFD		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
GD		F_1	2,78	1.02	557	*ns*
		F_2	2,48	*<1*		*ns*
TT		F_1	2,78	*<1*		*ns*
		F_2	2,48	*<1*		*ns*
% Skip		F_1	2,78	4.38	40	<.05
		F_2	2,48	2.98	53	=.136
Sent. 1		F_1	2,78	1.33	2	>.25
		F_2	2,48	*<1*		*ns*

Note: LF = low frequency; HF = high frequency; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Table 4.10 Follow up Contrasts to Significant Main Effect of Contextual Predictability by Subjects (F_1) and by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP vs. MP		LP > HP		MP > HP	
	F_1	*<1*	*ns*	23.84	<.001	24.37	<.001
	F_2	*<1*	*ns*	19.46	<.001	15.68	<.001
SFD		LP vs. MP		LP > HP		MP > HP	
	F_1	*<1*	*ns*	24.13	<.001	23.36	<.001
	F_2	*<1*	*ns*	15.38	<.001	11.69	<.01
GD		LP vs. MP		LP > HP		MP > HP	
	F_1	2	>.15	23.87	<.001	12.05	<.001
	F_2	2.76	=.103	23.8	<.001	10.35	<.01
TT		LP vs. MP		LP > HP		MP > HP	
	F_1	*<1*	*ns*	10.46	<.01	5.34	<.05
	F_2	*<1*	*ns*	7.8	<.01	3.93	=.053
% Skip		LP > MP		LP > HP		MP > HP	
	F_1	3.16	=.080	13.01	<.001	3.35	=.071
	F_2	1.55	>.20	6.4	<.05	1.65	>.20

Note: LP = low predictability; MP = medium predictability; HP = high predictability; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped

Table 4.11 Frequency Contrasts (LF vs. HF): Follow up to when the Word Frequency x Contextual Predictability Interaction was Significant by Subjects (F_1) and by Items (F_2)

		LP words		**MP words**		**HP words**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***
% Skip		LF vs. HF		LF < HF		LF vs. HF	
	F_1	*<1*	*ns*	10.03	<.01	*<1*	*ns*
	F_2	*<1*	*ns*	4.77	<.05	*<1*	*ns*

Note: LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency; % Skip = percentage of times the target word was skipped

Table 4.12 Contextual Predictability Contrasts: Follow-ups when the Word Frequency x Contextual Predictability Interaction was Significant

		LF words						**HF words**					
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
% Skip		LP > MP		LP > HP		MP vs. HP		LP vs. MP		LP vs. HP		MP > HP	
	F_1	10.48	<.01	10.03	<.01	*<1*	*ns*	*<1*	*ns*	2.34	=.0126	6.07	<.05
	F_2	4.99	<.05	4.77	<.05	*<1*	*ns*	*<1*	*ns*	1.14	>.25	2.89	=.096

Note: LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability; % Skip = percentage of times the target word was skipped

4.4 Discussion

The aim of the current experiment was to investigate the contribution of parafoveal processing to the frequency-predictability interaction observed in Experiment 2. This was in order to explore the possibility that some of the observed interaction in the first and single fixation duration measures in Experiment 2 could have been due to the parafoveal processing aspect of reading high frequency and high contextual predictability targets. During the course of reading printed text, readers process the fixated word in foveal vision as well as some features of the parafoveal word (Rayner, 1998). In particular, early work has shown that parafoveal information is used more efficiently from higher frequency targets than from lower frequency targets (Inhoff & Rayner, 1986) and also from higher contextual predictability targets than from lower contextual predictability targets (Balota et al., 1985). Inhoff and Rayner's (1986) finding suggests that word frequency modulates the use of parafoveal information where the higher the frequency of the parafoveal word, the greater the influences of parafoveal information i.e., the more lexical access is facilitated. Similarly, Balota et al.'s (1985) results suggest that contextual predictability effects, that is top-down sources arising from the sentence context, modulates the use of parafoveal information where the higher the contextual predictability, the greater the influences of parafoveal information i.e., the more lexical access is facilitated. If it is the case that lexical access is facilitated when the parafoveal word is of higher frequency (e.g., Inhoff & Rayner, 1986) and of higher contextual predictability (e.g., Balota et al., 1985), then one way of interpreting this pattern of effects is through an interactive view of lexical access (e.g., McClelland & O'Regan, 1981; McClelland & Rumelhart, 1981).

Therefore, the current experiment was designed to investigate the frequency-predictability relationship without the additional benefits of parafoveal processing obtained from higher frequency and higher contextual predictability targets. This was achieved by having readers read targets embedded in sentences without having viewed the target parafoveally hence removing the parafoveal processing advantage of higher frequency and contextual predictability targets. The boundary technique was used to block the parafoveal access of the target. A boundary location was added to the experimental passages of text used in Experiment 2. Thus, in place of the target, a nonword letter string occupied the target location and only when the readers cross the pre-specified boundary did the nonword letter string change to the target. Hence with this technique, when the reader fixates the target, it was the first time they were viewing it i.e., first time they were processing the target. A secondary aim was that the use of the boundary technique slowed down reading rate (e.g., Rayner et al., 1982) and this was one way to eliminate the floor effect observed in Experiment 2.

It was predicted that fixation duration measures and the skipping measure would reflect the foveal component of lexical access related to the frequency and predictability manipulations. In particular, when readers were subjected to a situation where there the parafoveal preview of the target was denied, it was expected that in first and single fixation durations, there would be significant main effects of word frequency and contextual predictability and a nonsignificant interaction. This additive result was expected to be maintained in gaze duration and total time durations. We were less clear the effect of denying a parafoveal preview would have on skipping but one likely outcome is that skipping rate would be low based on the fact that reading has been shown to depend on both foveal and parafoveal aspects of the text (e.g., Rayner et al., 1982).

Results observed in the two early fixation duration measures were supported (see Tables 4.9-4.12 for a summary of all the results). That is, in both first and single fixation measures, there were significant main effects of word frequency as well as of contextual predictability and also a nonsignificant interaction between the two variables. The main effect of word frequency was in the expected direction where high frequency targets were processed faster than low frequency ones as evidenced by significantly shorter first and single fixation durations on the high frequency targets. Follow-ups to the significant main effect of contextual predictability showed the same pattern of results in both first and single fixation measures. That is, high predictable targets were processed significantly faster than both moderately and low predictable targets. In both first and single fixation measures, readers spent the same amount of time fixating low and medium predictable targets (FFD was 243 ms and SFD was 250 ms). This was unexpected because the equivalent comparison in Experiment 2 showed that medium predictable targets were fixated significantly less than low predictable targets in both first and single fixation measures. However this finding is not necessarily in contradiction to interactive positions. That is, a possible reason that medium predictable targets were not read faster than low predictable targets is that in both cases, it can be argued that the context does not strongly constrain a particular target and this in combination with withholding the parafoveal preview of the target means that readers have relatively few sources of information about the identity of the upcoming word. Hence it could be that both low and medium predictable targets incurred a processing cost in terms of the amount of information usually available to a reader in order to allow them to predict the parafoveal word.

In the first and single fixation duration measures, we had expected the interaction between word frequency and contextual predictability to be nonsignificant because Hand et al.'s (2010) experiment suggests that when less of the parafoveal word is viewed, then word frequency and contextual predictability are unlikely to interact. In the present study in which the parafoveal word was blocked altogether, results showing the nonsignificant interaction between word frequency and contextual predictability support the view (in line with interactive positions) that the frequency-predictability interaction only takes place when there is sufficient preview information of the given target available to the reader. On a theoretical level, these results support interactive views of lexical access (e.g., McClelland & O'Regan, 1981). According to McClelland and O'Regan's (1981) model, information flow is seen to be multidirectional among the various component processors such that lexical access is influenced by parafoveal information. Thus, when parafoveal information is not available, then lexical access would be slowed because there are less information sources provided to the reader. Invalid parafoveal previews do not provide any information to the reader about the identity of the upcoming word thus word frequency and contextual predictability would not be expected to interact according to interactive positions.

The additive pattern of results observed in first and single fixation measures were maintained in the gaze duration and total time measures. Thus, the hypotheses in these measures were supported and the observed results in gaze duration and total time measures were a significant main effect of word frequency, a significant main effect of contextual predictability as well as a nonsignificant interaction.

The skipping measure which was computed was the percentage of times the target was skipped in each of the six conditions. It was less clear what the effect of blocking the parafoveal access of the target would be on skipping rates. However, one likely outcome is that in comparison to when readers have all the available information sources (i.e., information from both foveal and parafoveal sources), in the case of blocking the parafoveal preview of the target, the expectation is that readers would be less likely to skip the target. This is because when they process some aspects of the parafoveal word and given that it was not the actual target, they would want to know what it was, thus causing them to fixate it on the next fixation. Thus, we expected significant main effects of word frequency and contextual predictability and a nonsignificant interaction. Observed results were a nonsignificant main effect of frequency, a significant main effect of contextual predictability and an interaction which was significant in subjects analysis but only trend in items. Follow-ups to the significant main effect of contextual predictability showed that the only reliable significant difference was that LP words were skipped significantly more than HP words (14% versus 11% of the time, respectively) – an effect which is in an unexpected direction. Also in the unexpected direction (but nonsignificant) was that low predictable targets were skipped more than medium predictable ones and also that medium predictable targets were skipped more than high predictable ones. Why would participants have skipped low predictable targets more than high predictable ones? One possible reason could be because of the nature of the preview used. That is, it could be that because the nonword letter strings were so target word-like in terms of overall word shape with the target as well as having avoided 'weird' letter sequences that are likely to command attention. This means that the reader could have processed the nonword as the actual target itself (especially since they did not know they were getting a boundary change experiment). It could be that in low predictable conditions, even though the context was low constraint, readers were able to 'fill in' the target if they thought the preview of the letter string was the actual target.

For the significant skipping interaction (in subjects analysis), frequency and contextual predictability contrasts were conducted. The frequency contrasts showed that it was only in the moderate predictable condition that high frequency targets were skipped significantly more than low frequency ones; a frequency difference was not found in the low and high predictable conditions. Moreover, numerical differences in the low predictable condition showed that the frequency effect was in the unexpected direction where low frequency targets were skipped more than high frequency targets; the high predictability condition displayed numerical results in the expected direction (both numerical differences were nonsignificant). In terms of the predictability contrasts for the low frequency targets, two of the three contrasts were in the unexpected direction and were also significant. The predictability contrasts for the high frequency targets were a bit better in that two of the three contrasts were in the expected direction (but nonsignificant); however the contrast that was in the unexpected direction was partially significant.

First, for the low frequency targets, there were two significant results in the unexpected directions where low predictable targets were skipped significantly more than both medium and high predictable targets. Medium and high predicable targets incurred the same rate of skipping; thus there were no significant differences between these two conditions. The finding that low predictable targets were skipped the most could have been to do with the preview manipulation. That is, even though participants never directly see a parafoveal preview nonword change to the target, perhaps they were aware of the change-over once they has passed it. This could have led to readers adopting unusual reading strategies, of which one consequence is that they fixated longer than normal in medium and high predictable conditions than in low thus overriding the usual effect of contextual predictability. The finding that medium and high predictable conditions had the same rate of skipping could be accounted for by the explanation that the moderate context was sufficient for readers to ‘guess’ or ‘fill in’ the target, especially because the previously viewed parafoveal preview shared the overall word shape with the target – thus being congruent (in terms of overall word shape) with what the reader would expect given the context.

Second, for the high frequency targets, the unexpected direction of significant results was that participants skipped medium predictable targets significantly more than high predictable ones (though this difference was marginal in items). The other two contrasts were in the expected direction but nonsignificant (low predictable targets were skipped less than medium and high predictable ones). Why would there be no significant differences in skipping rates between low and high predictable conditions, as well as low and medium? And why were medium predictable targets skipped significantly more than high predictable targets? Again, as with the low frequency targets, it is possible that these odd results can be attributed to readers adopting different reading strategies if they had become aware of the text changing or ‘shifting’. This is possible as some studies have shown that even though readers never directly fixate the change from parafoveal to target take place, they can be aware of something going on in the text. Since the boundary change-over was happening on every trial, it could be that readers did adopt some strange reading strategies.

A secondary aim of using the boundary technique was that by inhibiting a parafoveal preview of the target i.e., preventing the reader parafoveally processing the target, reading rate should be slowed since readers have no longer pre-processed the target (e.g., Rayner et al., 1982). That is, using the boundary technique ensures that when the reader fixates the target, it is the first time they are processing it. The reason for slowing down reading times was that in Experiment 2, reading times were much faster than those in two past studies which have examined the frequency-predictability interaction. Since factors like participants (university students in all three studies were the population group) did not differ, one possible reason for the floor effect was the difference in the presentation of the materials during the running of the experiment. That is, in the present study, the display font was much clearer because it was less pixelated than in the past studies and this could have led to the much faster times observed in Experiment 2. Thus, in the cross Experiments 2 and 3 comparison, which inhibited parafoveal processing of the target, we expected fixation durations to be slower than in Experiment 2. Moreover, by using the boundary technique we thought that the possible floor effect (see Figure 3.4. for SFD in Chapter 2) could be eradicated – thus perhaps it is the case that targets which are high frequency-high predictable targets are read the fastest than those in other conditions but natural constraints of the visual system means that readers simply cannot fixate any faster on these targets, which are already read so fast because they are high frequency (i.e., more easily accessible from the lexicon) and so highly constrained in the text it appears in (i.e., which narrows the list of potential candidates chosen in the lexicon).

Results in the current experiment for the single fixation duration measure showed that the mean single fixation duration was 246 ms, thus showing that this time had been slowed compared to Experiment 2 (205ms) but was still not as slow as in Hand et al. (2010) and Rayner et al., (2004) (277 ms and 271 ms respectively). Interestingly, Hand et al. (2012), which used a similar clear display as in Experiment 2, had a mean SFD of 192 ms, thus being faster than the SFD in Experiment 2. Table 4.13 shows the single fixation durations in all these studies.

Table 4.13 Single-Fixation Durations in Experiment 3, Experiment 2 and in Rayner et al. (2004) and in Hand et al. (2010, 2012) studies

		LF			HF			
Display		**LP**	**MP**	**HP**	**LP**	**MP**	**HP**	**Mean SFD**
Pixelated	Rayner, Ashby, Pollatsek, & Reichle (2004)	287	269	na	268	258	na	271
Pixelated	Hand, Miellet, Sereno, & O'Donnell (2010)	294	285	na	269	259	na	277
Clear	Hand, O'Donnell, & Sereno (2012)*	202	190	na	193	181	na	192
Clear	Experiment 1	221	210	201	208	194	196	205
*Boundary***	Current study	258	257	244	242	242	232	246

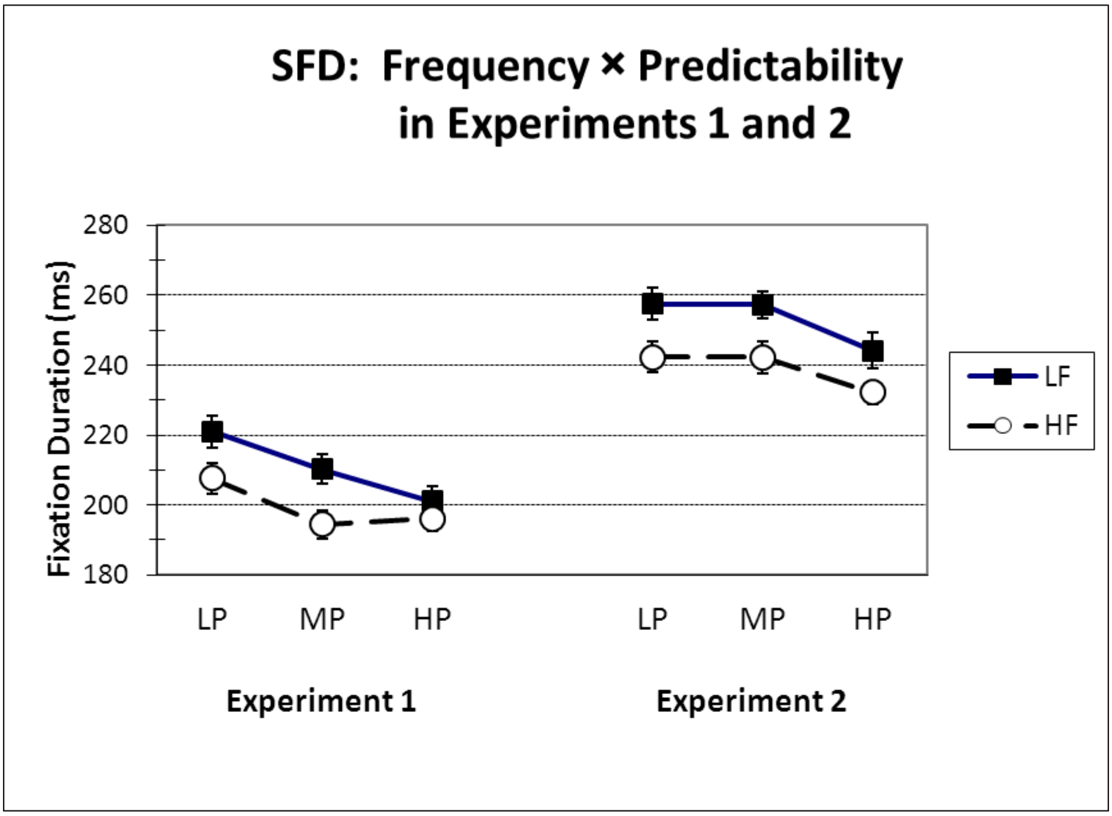
* Collapsed over constraint.
** The present study blocked the parafoveal access of targets to slow reading times rather than using a pixelated font
Note: LF = low frequency; HF = high frequency; LP = low predictable; MP = medium predictable; HP = high predictable; SFD = single fixation duration

As can be seen from Table 4.13, in the current study, high frequency-high predictable targets were fixated for the least amount of time (232 ms) compared to in the other five conditions, thus possibly removing the supposed floor effect in Experiment 2 (196 ms in the high frequency-high predictability condition). Moreover, in the current experiment, we showed a difference in single fixation durations between low and high frequency high predictable targets (12 ms difference between the respective conditions vs. 5 ms in Experiment 2[2]). However, before it can be concluded that the boundary technique was successful in removing the floor effect, when examining results across Experiments 2 and 3, it became apparent that the pattern of results between low and moderate predictable conditions had changed in the current study. This can be seen in Figure 4.4. where for the purposes of clarity, we plotted the single fixation duration data in the current experiment with that in Experiment 2. As can be seen from this Figure, the single fixation times in the current study have been slowed compared to those in Experiment 2. However, the pattern of results between low and medium predictable targets in the current has also altered drastically compared to that in Experiment 2. Namely, in the current study for the low frequency targets, there was only a 1 ms difference between low and moderate predicable words (LF-LP, 258 ms and LF-MP, 257 ms) and for the high frequency targets, low and moderate targets received identical fixation durations (both 242 ms). This differs with the equivalent predictability contrasts in Experiment 2 where for both the low and high frequency targets, low predictable targets had significantly longer first fixations than medium predictable targets. Thus, even though on initial inspection, in the current study, high frequency-high predictable targets incurred the fastest single fixation durations than in the other five conditions (indicating that the floor had been removed in the data), and the results were additive (main effects of frequency and contextual predictability and a nonsignificant interaction) a closer inspection of the pattern of results, as displayed in Figure 4.4., shows that since single fixation durations between low and moderate conditions were about the same respectively for low and high frequency targets, it is difficult to make a firm conclusion regarding the removal of the floor effect. Even though the boundary technique successfully slowed down reading times, at the same time by using this technique, it is possible that many other factors may have been introduced which were likely to have affected normal reading processes. For example, readers were not even getting the real preview of the parafoveal word. It could be that a more reliable comparison with past studies would be to actually use a pixelated font in conjunction with the experimental passages used in Experiment 2 to slow down reading times. However, on a

[2] The nonsignificant interaction in the current study meant that it is not permissible to carry out frequency contrasts; however the numbers indicate that this contrast would be significant.

practical level this would be difficult to implement since it means replacing new technology with older machines.

Figure 4.4. Mean single fixation duration as a function of word frequency and contextual predictability in Experiment 2 and current Experiment 3



Note: SFD = single fixation duration; LF = low frequency; HF = high frequency; LP = low predictable; MP = medium predictable; HP = high predictable

Overall, the boundary technique represents one way to prevent processing of the parafoveal word. This is because when the readers' eyes are fixated on a pre-target location (i.e., where they would normally obtain information parafoveally), instead of the actual target word being there, the technique allows some other 'invalid' letter string (i.e., not conveying the usual useful information) to take its place. Only when the participants' eyes make a fixation over some pre-specified location (normally the last letter of the pre-target word) does the 'invalid' letter string change to the actual target word. Crucially, this change occurs during a saccade so a participant never directly sees the changeover. Thus, when readers fixate the target word, it is the first time they are obtaining useful lexical and semantic information from that word.

The merits of the present boundary study are that it used far better previews in the invalid preview condition were used than in past studies of either frequency or predictability investigations. That is, targets and previews were matched on length and overall word shape. This latter requirement was achieved by replacing ascenders with ascenders, descenders with descenders and so on. In addition, the resulting preview letter strings, whilst being nonword letter strings, followed the orthographic rules of the English language. All of these criteria were fulfilled in order that the nonword letter string did not stand out as being ‘weird’ and thus command attention. Strange letter combinations are likely to be disconcerting and lead to longer fixation durations on the target when it is fixated. However, in making the nonword preview letter strings so word-like, it could be that readers processed the nonword as being a real preview of the parafoveal word. This is quite likely since it was only after taking part in the study that participants were told that the experiment had used a boundary change. One outcome of the nonword being processed as the actual preview of the target is that when readers did fixate the actual target, this could have been disconcerting, ‘wrongly’ leading to longer fixation durations in certain conditions.

4.4.1 Conclusions

The present study was designed to examine the frequency-predictability interactive relationship observed in Experiment 2 without the possible parafoveal processing advantage gained from parafoveal processing high frequency and high contextual predictability targets. The present study was the first to present better parafoveal previews to readers than previous studies have done so. That is, the parafoveal preview of the target was created in order that the letter string was orthographically legal; matched in word length and avoided letter-position overlap. All of these criteria were strictly adhered to in order to prevent the parafoveal preview of the target commanding readers’ attention which could lead the adoption of unnatural reading strategies. To do this, the boundary technique was used to block the parafoveal access of the target. That is, in place of the target, the reader viewed the parafoveal letter string (created as described above). A secondary aim of using the boundary technique was that readers fixation durations were able to be slowed. This was in order to address the issue in Experiment 2 where it was observed that this experiment had seen much faster fixation durations than in comparable previous studies of frequency by predictability. Results in the first, single, gaze and total time duration measures were as hypothesised where the additive result was taken to be line with interactive views of lexical access. That is, since invalid parafoveal previews of target do not provide the reader with useful information regarding the identity of the upcoming word, lexical access would be expected to be slowed and an interaction between word frequency and contextual predictability would not be expected. In terms of the skipping measure, results were far less conclusive. Follow-ups the significant main effect of contextual predictability showed one result in an unexpected direction. For the significant interaction (in subjects analysis), several follow-ups occurred in unexpected directions.

Overall, it was suggested that readers may have adopted unnatural and unusual reading strategies if they had become aware of the text changing in their parafovea. That is, it could be argued that presenting information in the parafovea other than the actual target is ultimately unnatural and unusual – in normal reading, we expect to get the actual word in our parafovea. However when this actual target parafoveal information is withheld, readers ultimately process the 'misinformation' and it is possible that this adds something that takes away from processes which occur in normal reading. Overall, the present study represents a good starting point to investigating the word frequency-contextual predictability relationship without the added benefit obtained from reading high frequency and high contextual predictability targets.

Cross Experiments 2 and 3 comparison

4.5 Additional data analyses

The purpose of the cross Experiments 2 and 3 comparison was to conduct additional data analyses in order to examine whether word frequency and contextual predictability of the parafoveal word affected parafoveal preview benefit. Experiment's 2 set of data were considered to have offered readers a 'valid preview' of the upcoming target and Experiment 3 to have given readers an in 'invalid' preview of the subsequent target. Therefore, in the present secondary analysis, we computed Experiment 2's data as a condition of 'valid preview' and Experiment 3's data as a condition of 'invalid preview' (for the factor of parafoveal preview; between-groups variable). This allowed us to compare the effects of parafoveal processing with word frequency and contextual predictability.

Only one study has examined the joint effects of word frequency and contextual predictability with parafoveal preview in a single study (Hand et al., 2010). Hand et al.'s (2010) study used an alternative method to the 'valid-invalid' preview conditions to index parafoveal processing. Therefore, the present study was the first study to manipulate word frequency and contextual predictability with parafoveal preview, using valid and invalid preview conditions.

In the present analyses a three-way design of word frequency x contextual predictability x parafoveal preview was used. This means that the following main effects were analysed: word frequency, contextual predictability and parafoveal preview. The two-way interactions were: word frequency x predictability; word frequency x preview; and contextual predictability x preview. Finally, the three-way interaction analysed was word frequency x contextual predictability x parafoveal preview. Since the present data analyses was in using data already collected, we hypothesised significance of results in the present analyses based on what had been shown in Experiments 1 and 2 data in fixation duration measures, percentage of times the target was skipped as well as number of fixations on sentence 1 of the passage.

In terms of expected results, all main effects in fixation duration measures and percentage of skipping were expected to be significant in fixation durations and percentage skipping rates. Based on Experiments 2 and 3 results, it appears as if the two-way interaction word frequency x contextual predictability would be nonsignificant in first and single fixation duration measures.

In terms of word frequency x parafoveal preview interaction, Inhoff and Rayner's (1986) study showed that word frequency and parafoveal preview interacted in first fixation duration. However, based on the results shown in Experiments 2 and 3, we thought that this two-way interaction may not reach significance in first and single fixation duration measures. In terms of contextual predictability x parafoveal preview, Balota et al.'s (1985) study showed that contextual predictability and parafoveal preview only interacted in the later gaze duration measure. However the present Experiments used high Cloze values in the highly predictable conditions than in Balota et al.'s (1985) study. Based on Experiments 2 and 3 results, we thought that this two-way interaction may be nonsignificant in first and single fixation duration measures with this pattern being maintained in the other measures. In terms of the three-way interaction, we predicted that in first and single fixation measures the interaction may be significant. In the skipping measure, we expected all two-way and three-way interactions to be nonsignificant. For the final measure, the number of fixations on sentence 1 of the passage, across all calculations, nonsignificant results were expected.

4.6 Method

4.6.1 Design

A mixed group 2 x 3 x 2 design was used: the effects of two types of frequency (low frequency, LF; high frequency, HF) on three types of contextual predictability (low, medium and high predictability; LP, MP, HP respectively) in two types of viewing conditions (a normal parafoveal preview of the target; an invalid parafoveal preview of the target) were investigated. Thus, frequency and contextual predictability were within-subject variables while preview was a between-groups variable. Dependent variables were standard eye movement measures (FFD, SFD, GD, TT) as well as percentage of times the target was skipped across the conditions (%Skip). There were a total of six conditions (LF-LP, LF-MP, LF-HP and HF-LP, HF-MP, HF-HP) viewed either with a normal parafoveal preview of the target (valid preview trials) or with an invalid parafoveal preview of the target (invalid preview trials). A total of 25 passages of text (hence 25 targets) in six conditions gave way to 150 experimental passages of text (where one group read 150 passages in the valid preview condition and another group read the same passages in the invalid preview condition).

4.6.2 Materials

The materials for the valid preview trials can be found in Appendix C. Also shown in this Appendix is the invalid parafoveal preview for every target. Appendix D shows the individual specifications for the 150 target words. Summary specifications of target words as well as information about word frequencies are shown in Table 3.1 in Chapter 3. Details of the two norming tasks used to determine contextual predictability of every target are shown in pre-tests 1 and 2 in Chapter 3.

4.6.3 Apparatus

Details of the apparatus used to monitor participants eye movements is detailed in Section 3.8.2 in Chapter 3.

4.6.4 Participants

Participant information for the valid preview trials are in Section 3.8.1 in Chapter 3 and for participants who were regarded as viewing an invalid preview of the target, please see Section 4.2.1 in this Chapter.

4.6.5 Procedure

Relevant information regarding the procedure for participants viewing a valid preview of the target is shown in Section 3.8.5 in Chapter 2 and in Section 4.2.5 in the current Chapter for the invalid preview trials.

4.7 Results

Table 4.14 shows the mean fixation time in milliseconds as well as percentage of times the target word was skipped in the conditions when viewed in valid and invalid preview conditions. Also shown is the total fixation time spent on sentence 1 as one region. Means with standard error bars are displayed in Figure 4.5. for the SFD measure.

Table 4.14 Mean Fixation Time Results in Milliseconds (Standard Deviation) and Percentage of Times the Target Word was Skipped in the Conditions in Valid and Invalid Preview Conditions

Measure		**LF**				**HF**	
	LP	**MP**	**HP**		**LP**	**MP**	**HP**
FFD							
Valid	219 (28)	210 (27)	200 (26)		207 (26)	195 (26)	196 (23)
Invalid	247 (26)	250 (22)	235 (28)		238 (27)	236 (26)	228 (19)
SFD							
Valid	221 (28)	210 (27)	201 (27)		208 (27)	194 (25)	196 (23)
Invalid	258 (29)	257 (25)	244 (32)		242 (27)	242 (29)	232 (21)
GD							
Valid	234 (32)	219 (31)	213 (36)		217 (29)	201 (27)	200 (25)
Invalid	289 (44)	279 (34)	271 (54)		267 (40)	267 (38)	250 (32)
TT							
Valid	252 (40)	240 (41)	227 (40)		237 (34)	219 (31)	209 (31)
Invalid	326 (84)	321 (96)	305 (99)		305 (71)	297 (82)	279 (68)
%Skip							
Valid	21 (12)	24 (15)	25 (14)		25 (13)	28 (14)	29 (16)
Invalid	15 (10)	10 (9)	10 (7)		13 (9)	14 (9)	11 (8)
Sent. 1							
Valid	28 (4)	29 (5)	29 (5)		29 (5)	29 (5)	29 (5)
Invalid	27 (5)	28 (5)	28 (5)		28 (5)	27 (5)	27 (5)

Note: Valid = trials where targets were read with a valid i.e., normal parafoveal preview of the target; Invalid = trials where targets were read with an invalid parafoveal preview of the target i.e., an orthographically legal letter string which maintained word shape with the actual target; LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; TT = total time including regressions; %Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

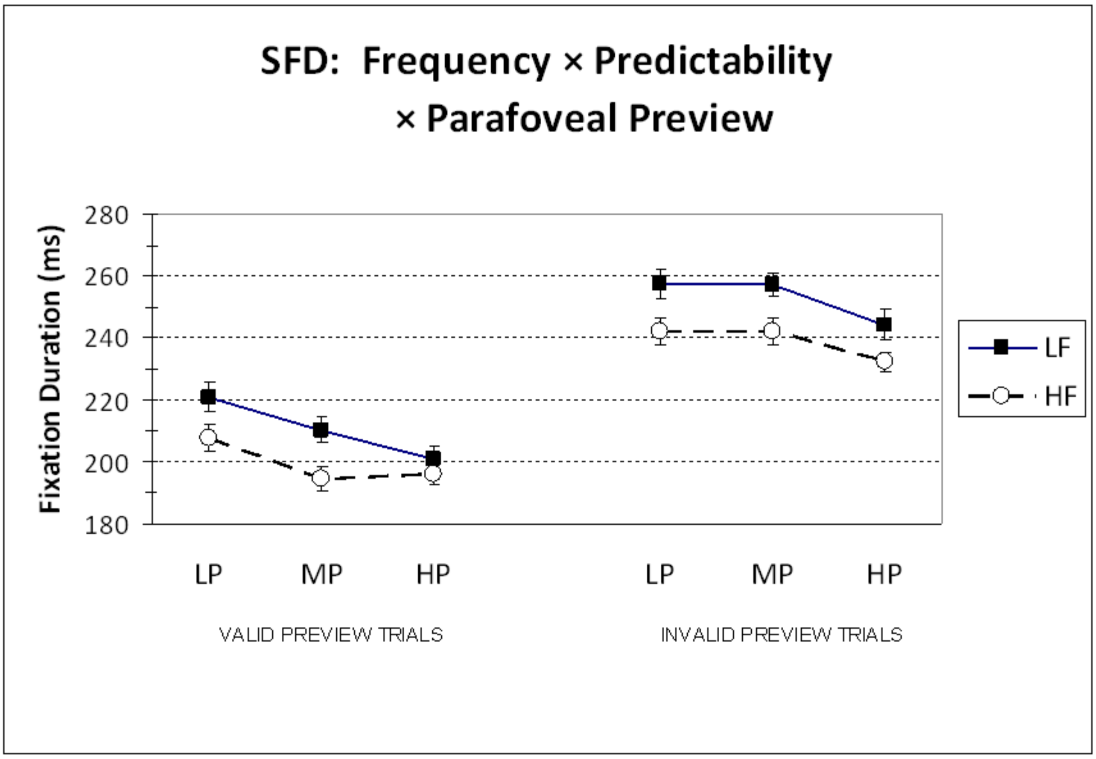
4.7.1 First fixation duration

4.7.1.1 Main effects

All three main effects were significant. First, there was a main effect of word frequency [$F_1(1,78)=67.52$, $MSE=190$, $p<.001$, $F_2(1,48)=39.79$, $MSE=187$, $p<.001$]. As in the initial analyses, HF words were fixated for significantly less time than LF words (217 ms vs. 227 ms). The main effect of contextual predictability was also significant [$F_1(2,156)=32.53$, $MSE=214$, $p<.001$, $F_2(2,96)=21.65$, $MSE=189$, $p<.001$]. Follow-up contrasts (Bonferroni multiple comparisons tests) revealed that LP targets were fixated for significantly longer than both MP targets (228 ms vs. 223 ms; $F_1=10.45$, $p<.01$, $F_2=9.38$, $p<.01$); as well as HP targets (228 ms vs. 215 ms; $F_1=64.26$, $p<.001$, $F_2=43.23$, $p<.001$). Unlike the initial analyses, MP targets were fixated significantly longer than HP targets

(223 ms vs. 215 ms; F_1=22.88, p<.001, F_2=12.33, p<.01). Finally, the main effect of preview was also significant [F_1(1,78)=49.81, *MSE*=2888, p<.001, F_2(1,48)=410.55, *MSE*=224, p<.001]. Thus, participants read targets faster when they saw a valid preview of it (204 ms) than when they saw an invalid preview of it (239 ms).

Figure 4.5. Average single fixation duration (SFD) as a function of frequency, contextual predictability and parafoveal preview

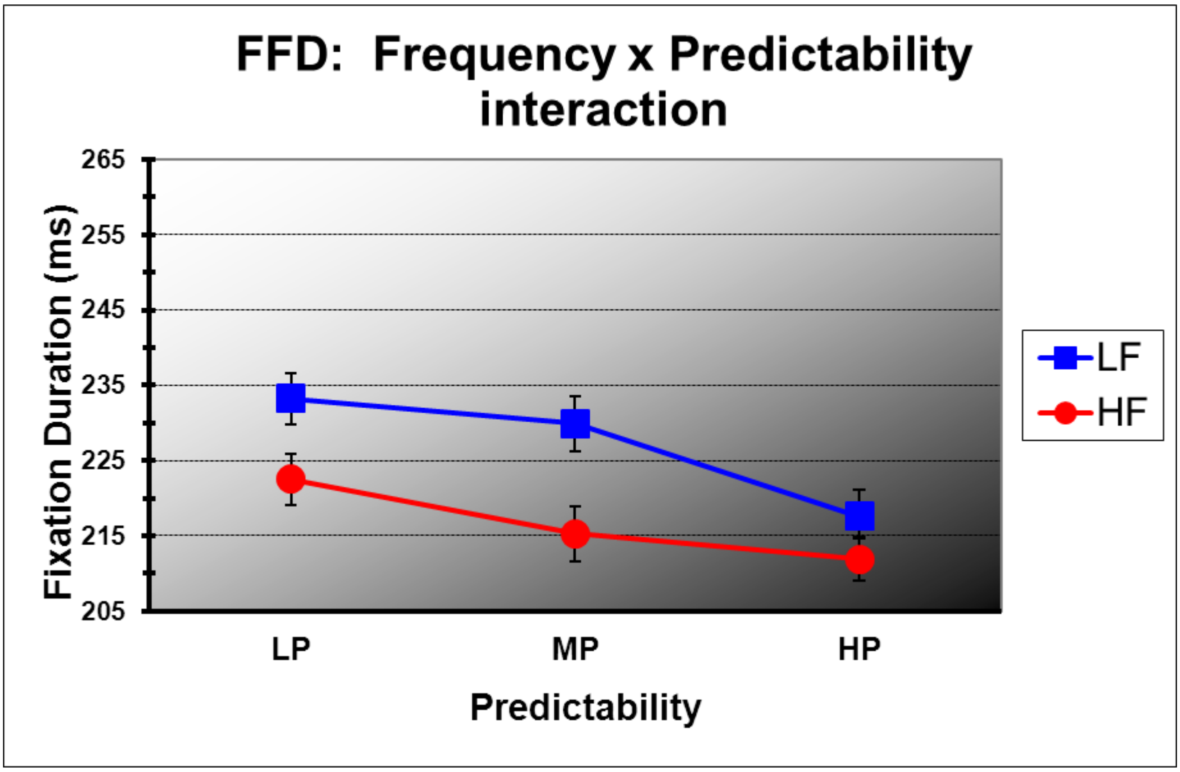


Note: SFD = single fixation duration; LP = low predictable; MP = medium predictable; HP = high predictable; LF = low frequency; HF = high frequency

4.7.1.2 Interactions

In terms of the two-way interactions, all were significant except for the word frequency x preview interaction [all *Fs*<1]. Thus the contextual predictability x preview interaction was significant [F_1(2,156)=5.70, *MSE*=214, p<.05, F_2(2,96)=3.34, *MSE*=189, p<.05] as was the word frequency x contextual predictability interaction [F_1(2,156)=4.57, *MSE*=104, p<.05, F_2(2,96)=3.74, *MSE*=71, p<.05]. Figure 4.6. displays the latter and Figure 4.7. shows the former. The three-way interaction of word frequency x contextual predictability x preview was nonsignificant [all *Fs*<1].

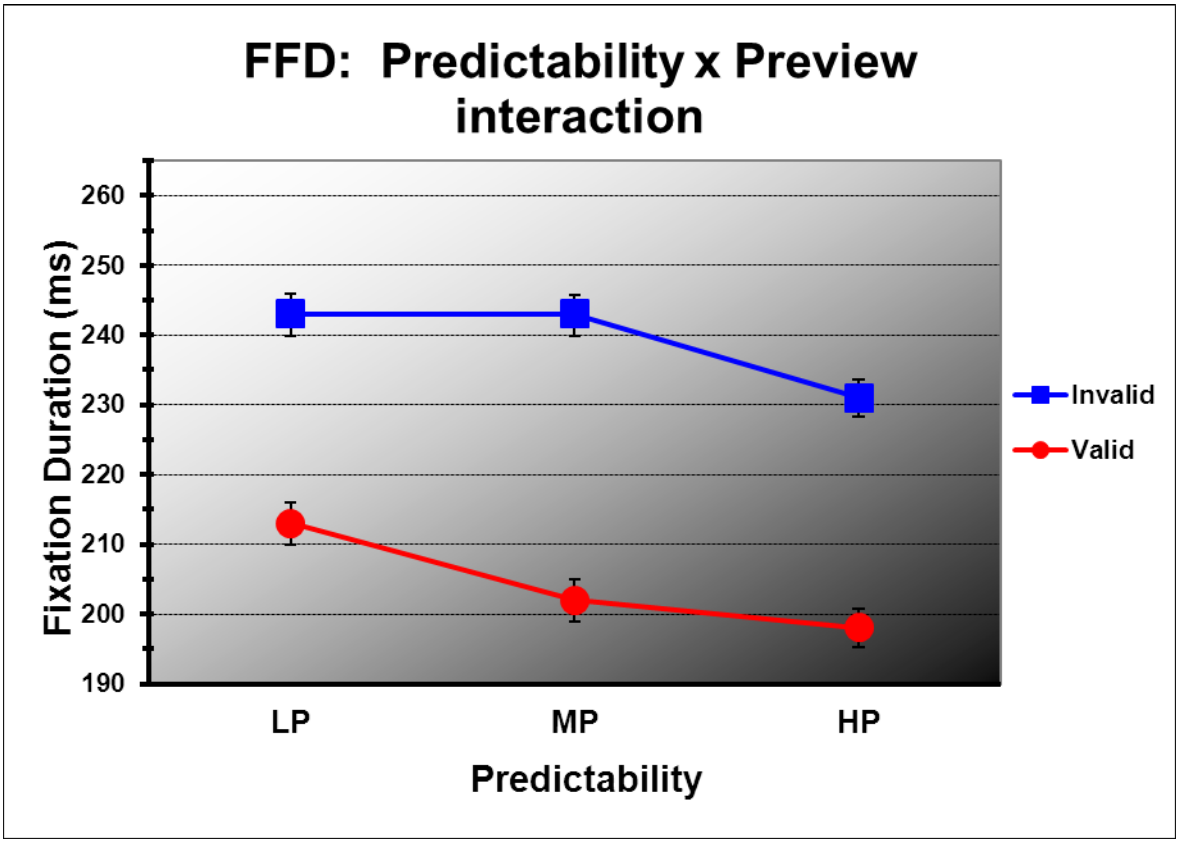
Figure 4.6. Word frequency by contextual predictability interaction in FFD



Note: FFD = first fixation duration; LF = low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability

Follow-ups (Bonferroni multiple comparisons) to the two significant two-way interaction were as follows: for the significant interaction of word frequency x contextual predictability interaction (see Figure 4.6.), follow-ups examined the effect of word frequency (LF vs. HF) at each level of contextual predictability (LP, MP, and HP) as well as the effect of contextual predictability for LF and for HF words. Basically, all numerical differences were in the expected direction. For the frequency contrasts at each level of predictability, all of these contrasts were significant except for the frequency contrast for HP targets (this was only partially significant since the items analysis was nonsignificant) (LP: LF>HF; 233 vs. 223 ms: F_1=25.83, p<.001, F_2=17.58, p<.001, MP: LF>HF; 230 vs. 215 ms: F_1=48.39, p<.001, F_2=26.37, p<.001, HP: LF>HF; 218 vs. 212 ms: F_1=7.24, p<.01, F_2=1.99, p>.15). For the predictability contrasts for LF and for HF targets, most contrasts were significant; however a few were only marginal or showing trend for significance. For the LF words, the contrast was nonsignificant: LP vs. MP (233 ms vs. 230 ms: F_1=2.48, p=.118, F_2=2.79, p=.098). The other two contrasts were significant: LP>HP (233 ms vs. 218 ms: F_1=55.01, p<.001, F_2=35.85, p<001; MP>HP (230 ms vs. 218 ms: F_1=34.21, p<.001, F_2=18.63, p<.001). Thus for LF words, HP targets were fixated for significantly less time (218 ms) than both MP (230 ms) and LP (233 ms) targets and there were nonsignificant differences in FFDs between LP (233 ms) and MP (230 ms) targets. For the equivalent HF words, again most contrasts were significant with the exception of one. Thus the contrasts were as follows: LP>MP (223 ms vs. 215 ms; F_1=11.89, p<.001, F_2=6.83, p<.05), LP>HP (223 ms vs. 212 ms; F_1=25.30, p<.001, F_2=10.30, p<.01), for MP vs. HP (215 ms vs. 212 ms; F_1=2.50, p=.116, F_2<1) there was a nonsignificant difference. Thus, for HF targets, HP targets (212 ms) and MP targets (215 ms) were fixated for significantly less time than LP targets (223 ms) but there were no significant differences in first fixation between MP (215 ms) and HP (212 ms) targets.

Figure 4.7. Contextual predictability by preview interaction in FFD



Note: FFD = first fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability

For the significant interaction of contextual predictability x preview interaction (see Figure 4.7.), when follow-up statistical tests were carried out (the simple main effects looking at valid versus invalid trials) the same pattern of results as those in the earlier Experiment 2 (valid) and Experiment 3 (invalid) were obtained. That is, predictability contrasts compared low versus medium; low versus high and medium versus high for valid and then for invalid preview trials. Therefore, our analysis looked at the follow-ups conducted to the predictability main effect in Experiments 2 and 3 (the simple main effects). These results showed in the valid preview trials, low predictable targets (213 ms) had significantly longer first fixations than medium predictable (203 ms) ones [F_1=21.59, p<.001, and F_2=14.45, p<.001] as well as than high predictable (198 ms) [F_1=41.74, p<.001, and F_2=23.80, p<.001]. There were no significant differences in first fixations on medium (203 ms) and high predictable (198 ms) targets; this contrast was marginally significant in the subjects analysis and nonsignificant in the items analysis [F_1=3.25, p=.074, and F_2=1.16, p>.25].

For the invalid preview trials, the same pairings of contrasts were carried out. There was a different pattern of results here to the valid preview condition: medium predictable versus high predictable was significant as well as low predictable versus high predictable, whereas low predictable versus high predictable was not. That is, in the invalid trial condition, high predictable targets (232 ms) had significantly faster first fixations than both medium predictable (243 ms) and low predictable (243 ms) [MP>HP: F_1=24.37, p<.001, F_2=15.68, p<.001; LP>HP: F_1=23.84, p<.001, F_2=19.46, p<.001]. There was the same duration of first fixations on low and medium predictable targets (both 243 ms; all *F*s<1).

4.7.2 Single fixation duration

4.7.2.1 Main effects

All three main effects were significant. First, there was a main effect of word frequency [$F_1(1,78)$=86.01, *MSE*=226, p<.001, $F_2(1,48)$=53.49, *MSE*=220, p<.001]. As in the FFD measure, HF words were fixated for less time than LF words (219 ms vs. 232 ms). The main effect of contextual predictability was also significant [$F_1(2,156)$=33.83, *MSE*=223, p<.001, $F_2(2,96)$=21.74, *MSE*=206, p<.001]. Follow-up Bonferroni contrasts revealed that LP targets were fixated for significantly longer than both MP targets (232 ms vs. 226 ms; F_1=13.22, p<.001, F_2=9.85, p<.01); as well as HP targets (232 ms vs. 218 ms; F_1=67.37, p<.001, F_2=43.46, p<.001). As in the equivalent FFD measure, MP targets were fixated significantly longer than HP targets (226 ms vs. 218 ms; F_1=20.90, p<.001, F_2=11.93, p<.001). Finally, the main effect of preview was also significant [$F_1(1,78)$=61.37, *MSE*=3267, p<.001, $F_2(1,48)$=421, *MSE*=310, p<.001]. Thus, participants read targets faster when they saw a valid preview of it (205 ms) than when they saw an invalid preview of it (246 ms).

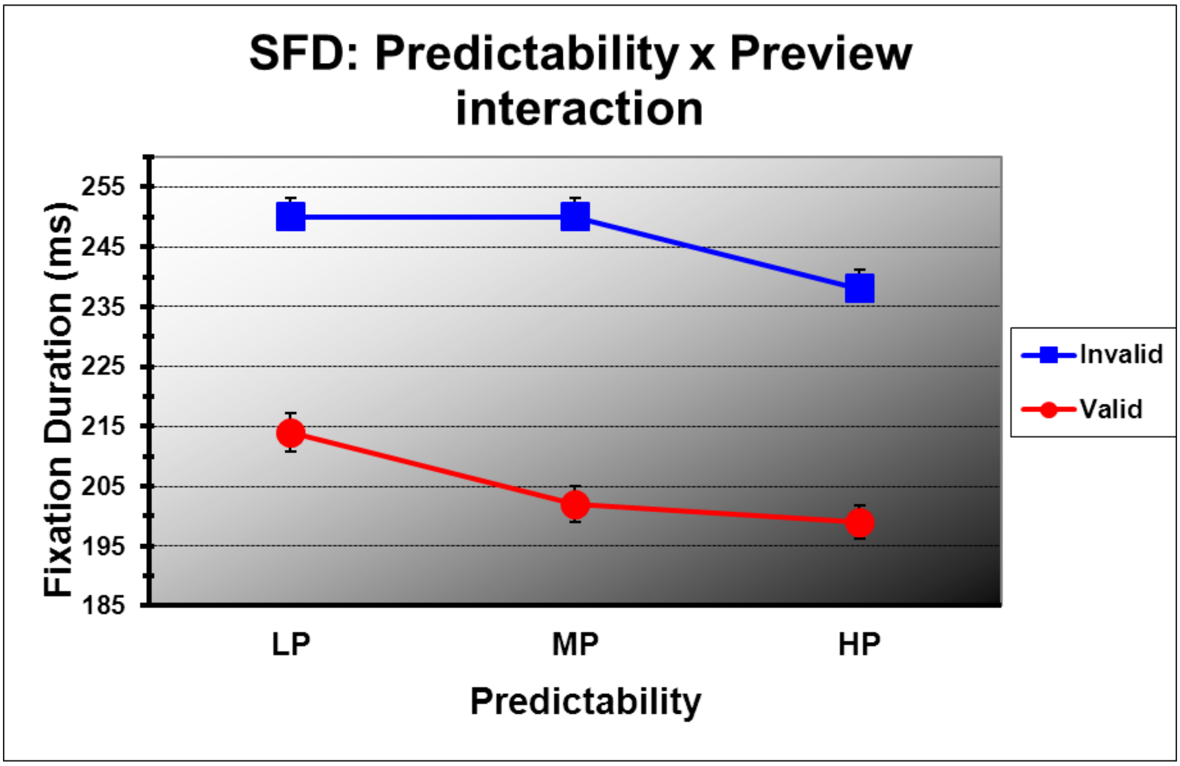
4.7.2.2 Interactions

In terms of the two-way interactions, as in the FFD measure, the word frequency x preview interaction was nonsignificant [$F_1(1,78)=1.01$, $MSE=226$, $p>.30$, $F_2<1$]; the contextual predictability x preview interaction was significant though it was marginal in items analysis [$F_1(2,156)=6.44$, $MSE=223$, $p<.001$, $F_2(2,96)=3.01$, $MSE=206$, $p=.054$] – see Figure 4.8 below. SFD results differed more so with the equivalent FFD results in the word frequency x contextual predictability interaction: results in the SFD measure showed marginal significance in subjects analysis and trend for significance in items analysis [$F_1(2,156)=2.93$, $MSE=192$, $p=.057$, $F_2(2,96)=2.26$, $MSE=280$, $p=.110$]. As in the FFD measure, in the SFD measure, the three-way interaction of word frequency x contextual predictability x preview was nonsignificant [all *Fs*<1].

Thus, in the SFD measure, the only significant two-way interaction was that of contextual predictability x preview interaction (though marginal in items analysis); see Figure 4.8. We looked at the follow-ups conducted to the predictability main effect in Experiments 2 and 3 to give us simple main effects. These results showed that in the valid preview condition, low predictable (215 ms) targets had significantly longer single fixations than medium predictable (202 ms) targets [$F_1=26.36$, $p<.001$, $F_2=17.62$, $p<.001$] as well as on high predictable (199 ms) ones [$F_1=45.19$ $p<.001$, $F_2=30.61$, $p<.001$]. The difference in single fixations between medium (202 ms) and high (199 ms) predictable targets showed a trend for significance in subjects analysis and was nonsignificant in items analysis [$F_1=2.52$, $p=.116$, and $F_2=1.78$, $p>.15$].

In the invalid preview condition, the same pattern of results as in the first fixation invalid preview condition was shown. That is, single fixations on high predictable targets (238 ms) were significantly shorter than those on medium predictable (250 ms) [$F_1=23.36$, $p<.001$, $F_2=11.69$, $p<.01$] as well on low predictable (250 ms) ones [$F_1=24.13$, $p<.001$, $F_2=15.38$, $p<.001$]. Readers had the same single fixations on low and medium predictable targets (both 250 ms; all *F*s<1).

Figure 4.8. Contextual predictability by preview interaction in SFD



Note: SFD = single fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability

4.7.3 Gaze duration

4.7.3.1 Main effects

All three main effects of word frequency, contextual predictability and preview were significant: [word frequency: $F_1(1,78)=73.76$, $MSE=462$, $p<.001$, $F_2(1,48)=45.17$, $MSE=410$, $p<.001$; contextual predictability: $F_1(2,156)=32.77$, $MSE=406$, $p<.001$, $F_2(2,96)=23.88$, $MSE=334$, $p<.001$; preview: $F_1(1,78)=68.33$, $MSE=5661$, $p<.001$, $F_2(1,48)=421$, $MSE=543$, $p<.001$]. Thus, HF words were fixated for less time than LF words (234 ms vs. 251 ms). Also, participants read targets faster when they saw a valid preview of it (214 ms) than when they saw an invalid preview of it (271 ms). Follow-up Bonferroni contrasts to the significant main effect of contextual predictability showed that all three contrasts were significant: LP targets were fixated for significantly longer than both MP targets (252 ms vs. 242 ms; $F_1=20.31$, $p<.001$, $F_2=16.54$, $p<.001$); as well as HP targets (252 ms vs. 234 ms; $F_1=65.24$, $p<.001$, $F_2=47.24$, $p<.001$). MP targets were also fixated significantly longer than HP targets (242 ms vs. 234 ms: $F_1=12.74$, $p<.001$, $F_2=7.87$, $p<.01$).

4.7.3.2 Interactions

As expected, all two-way interactions as well as the three-way interaction were basically nonsignificant: word frequency x preview [all *Fs*<1]; contextual predictability x preview: [F_1(2,156)=3.02, *MSE*=406, *p*=.052, F_2(2,96)=2.02, *MSE*=334, *p*=.138]; word frequency x contextual predictability [all *Fs*<1]; word frequency x contextual predictability x preview [F_1(2,156)=1.08, *MSE*=428, *p*>.30, F_2<1].

4.7.4 Total time

4.7.4.1 Main effects

All three main effects of word frequency, contextual predictability and preview were significant: [word frequency: F_1(1,78)=37.40, *MSE*=1366, *p*<.001, F_2(1,48)=23.56, *MSE*=966, *p*<.001; contextual predictability: F_1(2,156)=18.45, *MSE*=1374, *p*<.001, F_2(2,96)=13.37, *MSE*=1057, *p*<.001; preview: F_1(1,78)=31.37, *MSE*=21208, *p*<.001, F_2(1,48)=204, *MSE*=1861, *p*<.001]. Thus, HF words were fixated for significantly less time than LF words (258 ms vs. 278 ms). Also, participants read targets significantly faster when they saw a valid preview of it (231 ms) than when they saw an invalid preview of it (305 ms). Follow-up Bonferroni contrasts to the significant main effect of contextual predictability showed that all three contrasts were significant: LP targets were fixated for significantly longer than both MP targets (280 ms vs. 269 ms; F_1=7.01, *p*<.01, F_2=5.79, *p*<.05); as well as HP targets (280 ms vs. 255 ms; F_1=36.70, *p*<.001, F_2=26.69, *p*<.001). MP targets were also fixated significantly longer than HP targets (269 ms vs. 255 ms: F_1=11.63, *p*<.001, F_2=7.62, *p*<.01).

4.7.4.2 Interactions

As expected, all two-way interactions as well as the three-way interaction were nonsignificant: word frequency x preview [all *Fs*<1]; contextual predictability x preview: [all *Fs*<1]; word frequency x contextual predictability [all *Fs*<1]; word frequency x contextual predictability x preview [all *Fs*<1].

4.7.5 Percentage of times the target was skipped

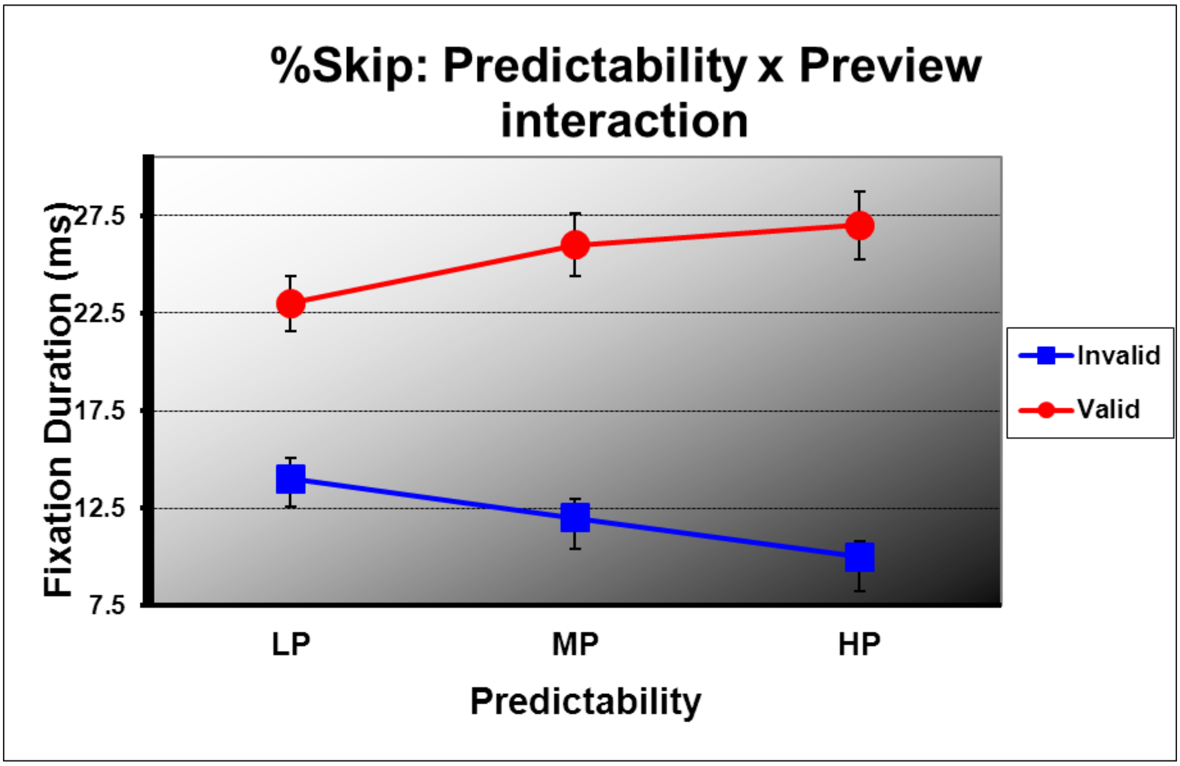
4.7.5.1 Main effects

Of the three main effects, two were significant. That is, the main effect of word frequency was significant [F_1(1,78)=13.83, *MSE*=64, *p*<.001, F_2(1,48)=12.06, *MSE*=46, *p*<.01]. Thus, HF targets were skipped significantly more than LF targets (20% and 17 % respectively). The main effect of preview was also significant [F_1(1,78)=39.45, *MSE*=531, *p*<.001, F_2(1,48)=36.32, *MSE*=361, *p*<.001]. Thus, participants skipped significantly more targets when targets were viewed with a valid preview of the target compared to when they were viewed with an invalid preview of the target (25% vs. 12% respectively). The main effect of contextual predictability was nonsignificant [all *Fs*<1].

4.7.5.2 Interactions

Two of the three two-way interactions were nonsignificant overall: word frequency x preview [$F_1(1,78)=3.59$, $MSE=64$, $p=.062$, $F_2(1,48)=3.13$, $MSE=46$, $p=.083$]; word frequency x contextual predictability [$F_1(2,156)=1.64$, $MSE=57$, $p>.15$, $F_2(2,96)=1.01$, $MSE=58$, $p>.35$]. However, contextual predictability x preview was significant [$F_1(2,156)=10.84$, $MSE=55$, $p<.001$, $F_2(2,96)=7.49$, $MSE=51$, $p<.01$] – see Figure 4.9.

Figure 4.9. Contextual predictability by preview interaction in percentage skipping data



Note: %Skip = percentage of times target was skipped; LP = low predictability; MP = medium predictability; HP = high predictability

As in the first and single fixation durations, we examined the follow-ups previously conducted to the predictability main effect in Experiments 2 (valid condition) and 3 (invalid condition), that is the simple main effects. These showed that in the valid condition, both medium (26%) and high (27%) targets were skipped significantly more than low predictable (23%) targets (though one contrast was marginal in the items analysis) [LP<MP: $F_1=4.49$, $p<.05$, $F_2=3.77$, $p=.058$; LP<HP: $F_1=9.99$, $p<.01$, $F_2=8.39$,

$p<.01$]. There were no significant differences in skipping rates between HP targets (27%) and MP targets (26%) [$F_1=1.08$, $p>.30$, $F_2<1$].

For the invalid trials, a different pattern of skipping was evident than in the valid trials. That is, the numerical differences showed results in opposite directions to what we expected: it was the *LP targets* that had the most skipping (14%), followed by MP (12%) and then HP (11%). The statistical tests showed that the only contrast which was significant was that LP targets were skipped significantly more than HP ones (14% vs. 11%: $F_1=13.01$, $p<.001$, $F_2=6.4$, $p<.05$). The other two contrasts (LP vs. MP and MP vs. HP) were marginally significant in subjects analysis and nonsignificant in items (LP vs. MP: 14% vs. 12%, $F_1=3.16$, $p=.080$, $F_2=1.55$, $p>.20$; MP vs. HP: 12% vs. 11%, $F_1=3.35$, $p=.071$, $F_2=1.65$, $p>.20$). Finally, as expected the three-way interaction was nonsignificant: [$F_1(2,156)=1.47$, $MSE=57$, $p>.20$, $F_2<1$].

4.7.6 Sentence 1 fixation time across the conditions

4.7.6.1 Main effects

The results were as expected for the main effect of frequency and the main effect of contextual predictability: thus, these two main effects were nonsignificant overall: word frequency [all $Fs<1$]; contextual predictability [all $Fs<1$]. However, for the main effect of preview, the result was nonsignificant in subjects analysis but significant in items analysis [$F_1(1,78)=1.27$, $MSE=131$, $p>.25$, $F_2(1,48)=31.88$, $MSE=4$, $p<.001$]. Thus, this preview results was that in the valid preview condition, first-pass time in reading sentence 1 was 28 ms per character and this was significantly faster than that in the invalid preview condition which was 27 ms per character. Since sentence 1 in both Experiments 2 and 3 were identical, this result points to likely participant variability. That is, it could be that in Experiment 2, participants were simply slower readers than participants in Experiment 3.

4.7.6.2 Interactions

As expected, all two-way interactions as well as the three-way interaction were nonsignificant: word frequency x preview [all $Fs<1$]; contextual predictability x preview: [all $Fs<1$]; word frequency x contextual predictability [all $Fs<1$]; word frequency x contextual predictability x preview [all $Fs<1$].

4.7.7 Results summary

The results are summarised in Tables 4.15, 4.16, 4.17, 4.18, 4.19 and 4.20 below. Overall results were in line with past research (Balota et al., 1985; Hand et al., 2010; Reingold et al., 2012). In the first fixation measure, the main effect results were as expected. In terms of the interaction results, most of the results were as predicted with two exceptions. First, contrary to predictions, results obtained significance in the contextual predictability x preview interaction as well as in the word frequency x contextual predictability interaction. In the single fixation measure, a similar pattern of main effect results were shown as in the first fixation measure. There were two differences in the interaction results compared to the first fixation results: in the single fixation measure, the contextual predictability x preview interaction was marginal in items analysis; the frequency x contextual predictability interaction was marginal in subjects analysis and trend in items analysis. Thus it appears that the interaction effects initially observed in the first fixation measure are becoming weaker in the single fixation measure.

In the gaze duration and total time measures, main effect and interaction results were as expected. That is, the additive pattern of results observed in first and single fixations continued to be maintained.

The skipping main effect data were mostly as expected where word frequency and preview were significant. However, one unexpected result was that the main effect of contextual predictability did not reach significance. The interaction of contextual predictability x preview was significant. However this was due to the direction of skipping taking place in the invalid preview trials versus the valid trials: in the former, there was a (unexpected direction) significantly higher rate of skipping for low predictable targets versus high predictable ones whereas in the valid trials, the (expected direction) high predictable targets were skipped significantly more than low predictable ones.

As well as analysing fixation times on the target, as detailed above, we also calculated first pass time on the first line of the passage of text. This was to ensure that there were no significant differences in reading time of line 1 between the conditions. Main effect results and interaction effects were as expected. There was a significant result in items analysis of the main effect of preview; possible reasons for this anomalous result are suggested in the Discussion section.

Main effect frequency						
Measure			***df***	***F***	***MSE***	***p***
FFD	**LF > HF**	F_1	1,78	67.52	190	<.001
		F_2	1,48	39.79	187	<.001
SFD	**LF > HF**	F_1	1,78	86.01	226	<.001
		F_2	1,48	53.49	220	<.001
GD	**LF > HF**	F_1	1,78	73.76	462	<.001
		F_2	1,48	45.17	410	<.001
TT	**LF > HF**	F_1	1,78	37.40	1366	<.001
		F_2	1,48	23.56	966	<.001
% Skip	**LF < HF**	F_1	1,78	13.83	64	<.001
		F_2	1,48	12.06	46	<.01
Sent. 1	**LF vs. HF**	F_1	1,78	*<1*	2	*ns*
		F_2	1,48	*<1*	2	*ns*
Main effect contextual predictability						
FFD		F_1	2,156	32.53	214	<.001
		F_2	2,96	21.65	189	<.001
SFD		F_1	2,156	33.83	223	<.001
		F_2	2,96	21.74	206	<.001
GD		F_1	2,156	32.77	406	<.001
		F_2	2,96	23.88	334	<.001
TT		F_1	2,156	18.45	1374	<.001
		F_2	2,96	13.37	1057	<.001
% Skip		F_1	2,156	*<1*	56	*ns*
		F_2	2,96	*<1*	51	*ns*
Sent. 1		F_1	2,156	*<1*	2	*ns*
		F_2	2,96	*<1*	3	*ns*
Main effect preview						
FFD	**Invalid > Valid**	F_1	1,78	49.81	2888	<.001
		F_2	1,48	410.55	224	<.001
SFD	**Invalid > Valid**	F_1	1,78	61.37	3267	<.001
		F_2	1,48	421	310	<.001
GD	**Invalid > Valid**	F_1	1,78	68.33	5661	<.001
		F_2	1,48	421	543	<.001
TT	**Invalid > Valid**	F_1	1,78	31.37	21208	<.001
		F_2	1,48	204	1861	<.001
% Skip	**Invalid < Valid**	F_1	1,78	39.45	531	<.001
		F_2	1,48	36.32	361	<.001
Sent. 1	**Invalid vs. Valid**	F_1	1,78	1.27	131	>.25
		F_2	1,48	31.88	4	<.001

Table 4.15 Summary of ANOVA Main Effect Results in all Fixation Duration Measures and Skipping by Subjects (F_1) and by Items (F_2)

Note: LF = low frequency; HF = high frequency; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Table 4.16 Follow up Contrasts to Significant Main Effect of Contextual Predictability by Subjects (F_1) and by Items (F_2)

Measure		***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP > MP		LP > HP		MP > HP	
	F_1	10.45	<.01	64.26	<.001	22.88	<.001
	F_2	9.38	<.01	43.23	<.001	12.33	<.01
SFD		LP > MP		LP > HP		MP > HP	
	F_1	13.22	<.001	67.37	<.001	20.90	<.001
	F_2	9.85	<.01	43.46	<.001	11.93	<.001
GD		LP > MP		LP > HP		MP > HP	
	F_1	20.31	<.001	65.24	<.001	12.74	<.001
	F_2	16.54	<.001	47.24	<.001	7.87	<.01
TT		LP > MP		LP > HP		MP > HP	
	F_1	7.01	<.01	36.70	<.001	11.63	<.001
	F_2	5.79	<.05	26.69	<.001	7.62	<.01

Note: FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; LP = low predicable; MP = medium predictable; HP = high predictable

Table 4.17 Summary of ANOVA Interaction Results in all Fixation Duration Measures and Skipping by Subjects (F_1) and by Items (F_2)

Frequency x preview					
Measure		***df***	***F***	***MSE***	***p***
FFD	F_1	1,78	<*1*	190	*ns*
	F_2	1,48	<*1*	187	*ns*
SFD	F_1	1,78	1.01	226	*ns*
	F_2	1,48	<*1*	220	*ns*
GD	F_1	1,78	<*1*	462	*ns*
	F_2	1,48	<*1*	410	*ns*
TT	F_1	1,78	<*1*	1366	*ns*
	F_2	1,48	<*1*	966	*ns*
% Skip	F_1	1,78	3.59	64	=.062
	F_2	1,48	3.13	46	=.083
Sent. 1	F_1	1,78	<*1*	2	*ns*
	F_2	1,48	<*1*	2	*ns*
Contextual predictability x preview					
FFD	F_1	2,156	5.70	214	<.05
	F_2	2,96	3.34	189	<.05
SFD	F_1	2,156	6.44	223	<.001
	F_2	2,96	3.01	206	.054
GD	F_1	2,156	3.02	406	=.052
	F_2	2,96	2.02	334	=.138
TT	F_1	2,156	<*1*	1374	*ns*
	F_2	2,96	<*1*	1057	*ns*
% Skip	F_1	2,156	10.84	55	<.001
	F_2	2,96	7.49	51	<.01
Sent. 1	F_1	2,156	<*1*	2	*ns*
	F_2	2,96	<*1*	3	*ns*
Frequency x contextual predictability					
FFD	F_1	2,156	4.57	104	<.05
	F_2	2,96	3.74	71	<.05
SFD	F_1	2,156	2.93	192	=.057
	F_2	2,96	2.26	280	=.110
GD	F_1	2,156	<*1*	428	*ns*
	F_2	2,96	<*1*	433	*ns*
TT	F_1	2,156	<*1*	599	*ns*
	F_2	2,96	<*1*	755	*ns*
% Skip	F_1	2,156	1.64	57	>.15
	F_2	2,96	1.01	58	>.35
Sent. 1	F_1	2,156	<*1*	2	*ns*
	F_2	2,96	<*1*	2	*ns*
Frequency x contextual predictability x preview					
FFD	F_1	2,156	<*1*	177	*ns*
	F_2	2,96	<*1*	193	*ns*
SFD	F_1	2,156	<*1*	192	*ns*
	F_2	2,96	<*1*	280	*ns*
GD	F_1	2,156	<*1*	428	*ns*
	F_2	2,96	<*1*	433	*ns*
TT	F_1	2,156	<*1*	599	*ns*
	F_2	2,96	<*1*	755	*ns*
% Skip	F_1	2,156	<*1*	58	*ns*
	F_2	2,96	<*1*	57	*ns*
Sent. 1	F_1	2,156	<*1*	2	*ns*
	F_2	2,96	<*1*	2	*ns*

Note: FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Table 4.18 Simple Main Effects: Follow-up Contrasts to the Significant Main Effect of Contextual Predictability in Experiment 2 (Valid Preview) and in Experiment 3 (Invalid Preview) by Subjects (F_1) and by Items (F_2)

		Valid preview						**Invalid preview**					
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP > MP		LP > HP		MP vs. HP		LP vs. MP		LP > HP		MP > HP	
	F_1	21.59	<.001	41.74	<.001	3.25	=.074	<1	n.s.	23.84	<.001	24.4	<.001
	F_2	14.45	<.001	23.8	<.001	1.16	>.25	<1	n.s.	19.46	<.001	15.7	<.001
SFD		LP > MP		LP > HP		MP vs. HP		LP vs. MP		LP > HP		MP > HP	
	F_1	26.36	<.001	45.19	<.001	2.52	=.116	<1	n.s.	24.13	<.001	23.4	<.001
	F_2	17.62	<.001	30.61	<.001	1.78	>.15	<1	n.s.	15.38	<.001	11.7	<.01
% Skip		LP < MP		LP < HP		MP vs. HP		LP vs. MP		LP > HP		MP vs. HP	
	F_1	4.49	<.05	9.99	<.01	1.08	>.30	3.16	=.080	13.01	<.001	3.35	=.071
	F_2	3.77	=.058	8.39	<.01	*<1*	*ns*	1.55	>.20	6.4	<.05	1.65	>.20

Note: FFD = first fixation duration; SFD = single fixation duration; % Skip = percentage of times the target word was skipped; LP = low predictability; MP = medium predictability; HP = high predictability

Table 4.19 Follow-up Frequency Contrasts (LF vs. HF) when the Word Frequency x Contextual Predictability Interaction was Significant by Subjects (F_1) and by Items (F_2)

		LP words		**MP words**		**HP words**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***
FFD		LF > HF		LF > HF		LF > HF	
	F_1	25.83	<.001	48.39	<.001	7.24	<.01
	F_2	17.58	<.001	26.37	<.001	1.99	>.15

Note: FFD = first fixation duration; LP = low predictability; MP = medium predictability; HP = high predictability; LF = low frequency; HF = high frequency

Table 4.20 Follow-up Contextual Predictability Contrasts when the Word Frequency x Contextual Predictability Interaction was Significant by Subjects (F_1) and by Items (F_2)

		LF words						**HF words**					
Measure		***F***	***p***	***F***	***P***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
FFD		LP vs. MP		LP > HP		MP > HP		LP > MP		LP > HP		MP vs. HP	
	F_1	2.48	=.118	55.01	<.001	34.21	<.001	11.89	<.001	25.30	<.001	2.5	=.116
	F_2	2.79	=.098	35.85	<.001	18.63	<.001	6.83	<.05	10.3	<.01	*<1*	*ns*

Note: FFD = first fixation duration; LF= low frequency; HF = high frequency; LP = low predictability; MP = medium predictability; HP = high predictability

4.9 Discussion

The aim of the current analyses was to examine the effects of manipulating word frequency and contextual predictability and parafoveal preview in order to examine the single and combined effects of these variables on fixation duration and skipping measures. For the variable of parafoveal preview, data previously collected in Experiments 2 and 3 were used for the condition of the 'valid' and 'invalid' parafoveal previews of the target, respectively. The parafoveal preview benefit is calculated by subtracting fixation durations in a valid preview condition from those in an invalid preview condition. Inhoff and Rayner's (1986) study suggested that parafoveal preview benefit is greater for high frequency than low frequency words. Balota et al.'s (1985) study suggested that the extraction of information from the parafoveal word is more efficient when aided by sentential context. Both findings suggest that word frequency and contextual predictability respectively affect lexical access rather than having their effect on post-lexical stages. Only one study has examined the effects of both word frequency and contextual predictability with parafoveal preview (Hand et al., 2010), however using launch site to index the extent of parafoveal processing. Therefore, it was reasoned that the relationship between frequency, contextual predictability and parafoveal preview needs further investigation. This was the first study to ensure several criteria were met when creating the nonword parafoveal preview letter strings so that they did not have odd letter sequences which could command attention.

Since the purpose of the current analyses was to carry out analyses with existing data, predictions were made based on the results of the data analysed in Experiments 2 and 3. Based on these results, it was hypothesised that all main effects would be significant in fixation duration measures as well as in the skipping measure. It was expected that the interaction between word frequency by contextual predictability would be nonsignificant in first and single fixation durations. The word frequency by parafoveal preview interaction as well as the contextual predictability by parafoveal preview interactions was also hypothesised to be nonsignificant in first and single fixation measures. This pattern of results was expected to be maintained in gaze duration and total time as well as shown in the skipping measure. The three-way interaction was hypothesised to be significant in first and single fixation measures and nonsignificant in gaze duration, total time and skipping.

4.9.1 Main effects in first and single fixation durations

In first and single fixations, it was expected that there would be significant main effects of word frequency, contextual predictability and parafoveal preview. All main effects were significant as expected. Thus, participants had significantly shorter first fixations when reading high than low frequency words, replicating many other studies (e.g., Inhoff & Rayner, 1986; Hand et al., 2010; Sereno & Rayner, 2000). For the significant main effect of contextual predictability, follow-ups showed that all three contrasts were significant (LP>MP; LP>HP and MP>HP). Hand et al. (2010) also showed a significant main effect of contextual predictability (in first, single and gaze duration measures). In addition the significant main effect of preview showed that participants first and single fixation durations were significantly shorter when reading targets they had previously viewed with a valid preview compared to when they had viewed a target with an invalid preview. Other studies which showed a significant main effect of preview were those by Balota et al. (1985) in the first fixation duration (as well as gaze duration); Inhoff and Rayner (1986) in the first fixation duration and Hand et al. (in first, single and gaze durations).

4.9.2 First fixation duration interactions

In the first fixation measure, we expected the two way interaction of word frequency by contextual predictability to be nonsignificant. However, actual results showed that this interaction was significant. Follow-up frequency contrasts (LF vs. HF) were significant at the three levels of predictability. Direction of results was as expected: participants' first fixations were shorter on high frequency targets than on low frequency targets; this was for low, medium and high predictability targets (the last contrast for HP target was just beyond trend in items analysis). Follow-up predictability contrasts for the low frequency targets were predicted to show no significant differences in first fixations between low and medium predictable targets; actual results showed that this difference was approaching significance being trend and marginal in subjects and items analyses respectively. Also predicted was that for the low frequency targets, participants would have significantly shorter first fixations on high predictable targets versus both low and medium predictable targets; results were as expected here. Overall, these results substantiated the underlying pattern of results observed in Experiments 2 and 3.

In terms of the two-way interaction of word frequency by preview, it was hypothesised that this interaction would not reach significance based on the pattern of results in Experiments 2 and 3. Actual results were as hypothesised (see Table 4.21).

Numerical differences in this Table show that when the parafoveal target was high frequency, participants' first fixations were 199 ms when there was a valid preview of the target versus when 234 ms when there was a valid preview of the target. Thus, the size of the parafoveal preview benefit when reading high frequency targets was 35 ms, however when reading low frequency targets the preview benefit was 34 ms. Kennison and Clifton (1995) and Sereno & Rayner (2000) also showed a nonsignificant interaction between word frequency and preview in first fixation durations.

Table 4.21 FFD Frequency x Preview Interaction*: Size of the Parafoveal Preview Benefit for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HF	199 ms	234 ms	35 ms
LF	210 ms	244 ms	34 ms

*Interaction was nonsignificant
Note: FFD = first fixation duration; HF = high frequency; LF = low frequency

Some previous studies have shown that word frequency and preview are interactive in first fixations (Inhoff & Rayner, 1986; Hand et al., 2010; Reingold et al. 2012). Specifically, this interaction is that a preview of high frequency words leads to shorter first fixations than a preview of low frequency words. Also if one looks at the parafoveal preview benefit (calculated by subtracting the time taken in invalid from valid conditions for high and low frequency targets), the typical finding is that readers obtained larger parafoveal preview benefits from high than low frequency targets. Evidence such a parafoveal preview benefit in Inhoff and Rayner's (1986) study, led to the authors suggesting that word frequency modulates parafoveal information where readers acquire more information from high frequency versus low frequency targets. In Hand et al.'s (2010) study in general, the preview condition of 'near' can be considered as offering a valid preview of the parafoveal word and 'far' as an invalid preview. Hand et al.'s results (see Table 4.22) showed that the size of the preview benefit was 45 ms for high frequency targets, versus 7 ms for low frequency ones (in first fixation durations). In Reingold et al.'s (2012) study, in first fixations, the preview benefit was 37 ms for high frequency targets and 26 ms for low frequency ones.

Furthermore, if one consider frequency effects (the finding that high frequency words are processed faster than low frequency ones), in Hand et al.'s study, there was a larger frequency effect in the valid preview condition compared to the invalid condition (50 ms vs. 12 ms respectively). Also, in Inhoff and Rayner's (1986) study, there was also a larger frequency effect in the valid preview condition than the invalid (valid preview: 18 ms vs. invalid preview: 0ms). This was also the case in Reingold et al.'s (2012) study (valid preview: 20 ms vs. invalid preview: 9 ms). The equivalent comparison in the present study was different to all these studies in that the size of the frequency effect was not larger in the valid preview condition: in the valid preview condition, the size of the frequency effect was 11 ms vs. 10 ms in the invalid preview condition, most probably reflecting in part the nonsignificant interaction.

Table 4.22 FFD and SFD Frequency x Preview Interaction**:** Size of the Parafoveal Preview Benefit for High and Low Frequency targets in Hand et al.'s (2010) Study

	Valid preview*	**Invalid preview****	**Size of parafoveal preview benefit**
FFD			
HF	226 ms	271 ms	45 ms
LF	276 ms	283 ms	7 ms
SFD			
HF	227 ms	278 ms	51 ms
LF	275 ms	293 ms	18 ms

* Valid preview was the 'near' (1-3 characters) condition
** Invalid preview was the 'far' (7-9 characters) condition
Note: FFD = first fixation duration; HF = high frequency; LF = low frequency; SFD = single fixation duration

In terms of the two-way interaction of contextual predictability by preview, it was expected that results would be nonsignificant because of the pattern of results observed in Experiments 2 and 3. However, this interaction reached significance. One previous study has shown a significant interaction between contextual predictability and preview in first and single fixation measures (Hand et al., 2010). Previously Balota et al. (1985) showed this significant interaction in gaze duration only. The significant interaction is that there is a greater predictability effect (where higher predictability words are processed faster than low predictability ones) with greater preview of the parafoveal word (valid preview conditions, or near launch distance) than lesser preview of the parafoveal word (invalid preview, or far launch distance).

In the present study, to find out which predictability groups significantly differed, follow-up contextual predictability contrasts were carried out (which were actually contextual predictability contrasts conducted in Experiments 2 and 3 for valid and invalid preview respectively to give the simple main effects, see Table 4.18). Results are summarised in Table 4.23.

Table 4.23 FFDs in Valid and Invalid Parafoveal Preview Conditions with High, Medium and Low Predictable Targets

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HP	198 ms	232 ms	34 ms
MP	203 ms	243 ms	40 ms
LP	213 ms	243 ms	30 ms

Note: FFDs = first fixation durations; HP = high predictable; MP = medium predictable, LP = low predictable

As can be seen from Table 4.23, first fixations when reading targets in the invalid preview condition were always numerically longer than those in the valid preview condition. Predictability contrasts for valid preview and for invalid preview showed that results were in the expected direction but some comparisons were nonsignificant. Basically, for the valid trials, both moderate and high predictable targets incurred faster first fixations then low predictable targets; the difference between first fixations when reading moderate and high predictable targets was nonsignificant: this could be because of a floor effect, as suggested in Experiment 2. Conversely, for the invalid trials, there was no difference in first fixations between low and moderate predictable targets however high predictable targets incurred faster first fixations than both low and moderate predictable targets. The lack of a significant effect between low and moderate predictable targets could be because of when context is not highly predictable coupled with no useful preview of the upcoming word, then readers are slowed considerably. An interactive-activation type process would suggest that when context is not highly predictable and there is no preview of the parafoveal word, then readers do not have all the necessary information to limit the candidates that are chosen, hence readers will take more time to read the sentence in these conditions. The implication is that when the context is highly predictable, even without a parafoveal preview of the target, readers can guess or fill in the upcoming target.

As can be seen in Table 4.23 (column 4), the parafoveal preview benefit is occurring in an unexpected direction: the moderate predictable targets incurred the greatest preview benefit. However, in the right direction was that low predictable targets incurred the least parafoveal preview benefit compared to moderate and high predicable targets. In the invalid preview condition, the first fixation was 232 ms whereas in the valid preview condition, it was 198 ms. Previously we had said that using a boundary technique is likely to slow down reading times, thus removing the floor effect seen in Experiment 2. The longer time of 232 ms suggests that this has occurred – particularly since the medium-high predictability contrast was significant in the invalid preview condition. However the fact that the high predictable condition did not incur the largest preview benefit could be due to the possibility that even though the boundary technique did slow down reading time, they were not sufficiently slowed. For example, earlier we suggested that in terms of single fixations, when participants read targets with an invalid preview, they incurred mean single fixations of 246 ms, but this was still not as slow as those in past experiments (Hand et al., 2010: 277 ms and Rayner et al., 2004: 271ms). If this same logic is applied here, this indicates a longer first fixation than 232 ms in the invalid trials would have led to a greater parafoveal preview benefit for the high predictable words.

If just the low and moderate predictable conditions are considered, then the pattern of results resembles those in Hand et al. (2010; see Table 4.24). The faster fixation times in the present experiment are likely to be due to the clearer display than in Hand et al.'s (2010) study.

Table 4.24 FFDs and SFDs in Medium and High predictable Targets in Hand et al.'s (2010) Study

	Valid preview*	**Invalid preview****	**Size of parafoveal preview benefit**
FFD			
HP	n/a	n/a	
MP	237 ms	276 ms	39 ms
LP	264 ms	278 ms	14 ms
SFD			
HP	n/a	n/a	
MP	238 ms	284 ms	46 ms
LP	264 ms	287 ms	23 ms

* Valid preview was the 'near' (1-3 characters) condition
** Invalid preview was the 'far' (7-9 characters) condition

Table 3.24 Note: FFD = first fixation duration; SFD = single fixation duration; HP = high predictable; MP = medium predictable; LP = low predictable

The three-way interaction of word frequency by contextual predictability by preview was nonsignificant. It was expected that significance might be reached given the differing pattern of results of the individual variables in Experiments 2 and 3. However, both *F*-ratios were less than 1 indicating weak effects failing to manifest in significance. The study by Hand et al. (2010) offers the most similar comparison to the present study since only this study has manipulated all three factors in a single study. Hand et al.'s (2010) results were that the three-way interaction of these variables was significant. To follow this up, they conducted separate frequency by predictability ANOVAs (the relationship between these variables was the focus of the study) at each of the three levels of preview, near, middle and far (where 'near' and 'far' can generally be considered as valid and invalid trials). Overall, results showed that the frequency by predictability interaction was significant in valid preview trials. Such a result suggests that this interaction was modulated by launch distance. In the present study, the three-way interaction was nonsignificant in all fixation duration measures. This was unexpected given that two of the three two-way interactions reached significance in first fixations. A possible reason for this result could be due to the use a mixed design, specifically where preview was manipulated between-groups. In Hand et al.'s (2010) study, all the variables were within-subjects, which is recognised as a stronger design than between-groups since it minimises participant variables. A stronger manipulation would be to test the effects of frequency, predictability and preview in a within-subjects design. Potential issues with implementing parafoveal preview across two experiments is that the between-groups design is likely to introduce more subject variability and hence could be weakening effects. This would certainly explain some of the nonsignificant effects, as discussed above. Based on these results, it is difficult to make firm conclusions regarding the time course of parafoveal lexical and semantic processing. Nonetheless, the present study represents a good starting point in researching parafoveal preview benefit when reading high and low frequency targets in low, medium and high predictable contexts given the scarcity of research which has done so. As the next step, it was decided to manipulate parafoveal preview in a within-subjects design where participants read low and high frequency targets embedded in low and high predictable contexts (**Experiment 4**).

We now turn to results obtained in the single fixation duration measure. We expected results to be in the same direction as those observed in the first fixation measure. However, there were two notable differences between the two measures. This was in terms of the word frequency by contextual predictability interaction and the contextual predictability by preview interaction, where results were no longer reaching full significance in both the interactions.

4.9.3 Single fixation duration interactions

In the single fixation measure, as in the first fixation measure, we expected the two way interaction of word frequency by contextual predictability to be nonsignificant. However, since this interaction was shown to be significant in the first fixation measure, it could be expected that this interaction would also be significant in single fixations. Actual results showed a hint of an interaction, without actually reaching significance: the interaction was marginally significant in the subjects analysis and showed a trend for significance in the items analysis. Thus no follow-ups were conducted here.

In terms of the word frequency by preview interaction, as was shown in the first fixation duration, we expected this to be nonsignificant. Actual results showed this to be case (see Table 4.25). The numerical differences in this Table show that low frequency targets read with an invalid parafoveal preview had the longest single fixations, and this time was longer than low frequency targets read with a valid preview. However, the size of the parafoveal preview benefit was in an unexpected direction: when reading high frequency targets, this was 40 ms which is shorter than when reading low frequency targets, which was 42 ms. Only two previous studies have analysed results in single fixations (Hand et al., 2010; Reingold et al., 2012) and both studies showed a significant interaction between frequency and preview (for Hand et al.'s single fixation results, see Table 4.22). We calculated the parafoveal preview benefit in both these studies and found that there that there was a larger preview benefit from high frequency parafoveal targets than from low frequency ones (in Reingold et al. in the SFD measure, the high frequency preview benefit was 51 ms and the low frequency preview benefit was 37 ms). In the present results, numerical differences show an opposite direction, with low frequency targets incurring a slightly greater preview benefit than high frequency targets.

Table 4.25 SFDs Frequency x Preview Interaction*: Size of the Parafoveal Preview Benefit for High and Low Frequency Targets in Valid and Invalid Parafoveal Preview Conditions

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HF	199 ms	239 ms	40 ms
LF	211 ms	253 ms	42 ms

*Interaction was nonsignificant
Note: SFDs = single fixation durations; HF = high frequency; LF = low frequency

The contextual predictability by preview interaction could be expected to be significant in single fixations as it was shown to be interactive in the first fixation measure. Actual results showed significance in subjects analysis but marginal in items. Follow-ups were designed to see which predictability groups differed (see Table 4.26).

Table 4.26 SFDs in Valid and Invalid Parafoveal Preview Conditions with High, Medium and Low Predictable Targets

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HP	199 ms	238 ms	39 ms
MP	202 ms	250 ms	48 ms
LP	215 ms	258 ms	43 ms

Note: SFDs = single fixation durations; HP = high predictable; MP = medium predictable, LP = low predictable

Follow-ups showed that for valid preview trials, both medium and high predictable targets had faster single fixations than low predictable targets; there were no significant differences between medium and high predictable targets: this was the same pattern in first fixations. For invalid preview trials, the pattern of results were different to the valid condition in that high predictable targets had faster single fixations that medium predictable targets (as well as low predictable) and low and medium targets has no significant differences: again this was the same pattern observed in first fixations. The numerical differences show that single fixation durations in the invalid preview condition for low, medium and high predictable were always slower than those respectively in the valid preview condition. As in the first fixation measure, the size of the parafoveal preview benefit was numerically greatest for moderate predictable targets, followed by low predictable and least for high predictable.

The two differences between first and single fixation measures were in the word frequency by contextual predictability interaction and the contextual predictability by preview interaction. Basically, results were no longer reaching full significance in single fixations. This suggests that in general effects were weakening in the single fixation measure. This could be because of the manipulation of preview as a between-groups factor; a stronger design would be to manipulate preview in one study. Finally as in the first fixations, the three-way interaction of word frequency, contextual predictability and preview remained nonsignificant.

4.9.4 Gaze duration

4.9.4.1 Main effects

In the gaze duration, main effects resembled those in the first and single fixation. Thus word frequency, contextual predictability and preview were all significant. That is, participants had significantly shorter gaze durations on high frequency versus low frequency targets; and significantly shorter gaze durations on targets when they were presented in the valid preview versus the invalid preview. Follow-ups to contextual predictability showed that high predictable targets has significantly shorter gaze durations than both medium and low predictable targets and also that medium predictable targets had shorter gaze durations than low predictable targets.

4.9.4.2 Interactions

The word frequency by contextual predictability interaction ceased to be significant in gaze durations. Word frequency by preview remained nonsignificant as in the first and single fixation measures. Contextual predictability by preview was approaching significance where subjects analysis was marginally significant and trend in items. As with the first and single fixations, the three-way interaction remained nonsignificant.

4.9.5 Total time

4.9.5.1 Main effects

All three main effects were significant. Thus participants spent significantly less time on high frequency than low frequency targets; as well as significantly less time on targets when they were presented in the valid preview condition compared to the invalid preview condition. In addition, participants spent significantly less time on high predictable targets than on both medium and low predictable targets, as well on medium predictable targets compared to low predictable targets.

4.9.5.2 Interactions

The word frequency by contextual predictability interaction remained nonsignificant; as did word frequency by preview. Contextual predictability by preview was now nonsignificant. As before, the three-way interaction remained nonsignificant.

4.9.6 Percentage of times the target was skipped

4.9.6.1 Main effects

Main effects of frequency and preview were significant and in the expected directions. This means that participants skipped significantly more targets when they were high frequency versus low frequency and also when they were presented in the valid preview versus invalid preview. However, the main effect of contextual predictability was nonsignificant despite there being significant main effects of contextual predictability in both Experiments 2 and 3. A closer inspection of Experiments 2 and 3 skipping results shows why this might have been the case. In Experiment 3, follow-ups to the significant contextual predictability main effect showed that low predictable targets were skipped significantly more than high predictable targets – a result which is in the unexpected direction. Since skipping in Experiments 2 and 3 are going in opposite directions, the between-groups comparison effectively cancels out the differences leading to the nonsignificant result.

4.9.6.2 Interactions

For all two-way and three-way interaction, based on the pattern of results observed in Experiments 2 and 3, it was hypothesised that all interactions would be nonsignificant. Actual results showed that word frequency by contextual predictability was nonsignificant; word frequency by preview was approaching significance where both subjects and items analyses were marginally significant (Table 4.27) and contextual predictability by preview was reliably significant (Table 4.28); and the three-way interaction was nonsignificant.

Table 4.27 Percent Skipping Frequency x Preview Interaction*: Valid and Invalid Parafoveal Preview Conditions with High and Low Frequency Targets

	Valid preview	Invalid preview	Size of parafoveal preview benefit
Target			
HF	27%	13%	14%
LF	23%	12%	11%

* Frequency x preview interaction was marginally significant
Note: HF = high frequency; LF = low frequency

As can be seen in Table 4.27, the patterns of results numerically were all in the expected direction: low frequency targets in the invalid preview condition were skipped the least and high frequency targets in the valid preview were skipped the most. The results also show that there was a greater preview benefit from high frequency targets than low frequency ones. Since the interaction was at best marginally significant, it could only be concluded that these differences were approaching significance.

Table 4.28 Percent Skipping Predictability x Preview Interaction: Valid and Invalid Parafoveal Preview Conditions with High, Medium and Low Predictable Targets

	Valid preview	Invalid preview	Size of parafoveal preview benefit
Target			
HP	27%	11%	16%
MP	26%	12%	14%
LP	23%	14%	9%

For the contextual predictability by preview interaction, as can be seen in Table 4.28, in the valid preview condition, high, medium and low predictable targets were skipped 27%, 26% and 23% of the time. Therefore, the pattern of results is in the expected direction. The significant differences were that high predictable targets were skipped significantly more than low predictable targets; and medium predictable targets were skipped significantly more than low predictable targets (though this was marginal in items analysis). There were no significant differences in skipping between moderate and high predictable targets, which is likely to do with the floor effect. For the invalid preview condition, results were in the unexpected direction: low predictable targets were skipped the most followed by medium predictable with high predictable targets incurring the least skipping. The significant differences were that low predictable targets were skipped significantly more than high predictable targets (unexpected direction). The other two contrasts (low vs. moderate and moderate vs. high) and were not fully significant: results were marginally significant in subjects analysis and nonsignificant in items. Why would low predictable targets be skipped significantly more than high predictable targets? This unexpected result could be because of the issue of what stimuli to present in the 'invalid preview' condition. Specifically, we used a nonword letter string. However, this invalid parafoveal preview is not the same as a 'no preview' condition in which case there would be no parafoveal information at all. That is, participants could have parafoveally processed aspects of the nonword or 'misinformation' and this could have added something to the reading process. The size of the parafoveal preview benefit shows numerically these were in the expected direction: the preview benefit was 16% for high predictable targets, which was more than for medium predictable targets (14%) and both were more than the benefit for low predictable targets (9%).

4.9.7 Conclusions

The aim of the current data analyses were to use the data previously collected in Experiments 2 and 3 in order to investigate the combined effects of the variables word frequency, contextual predictability and preview. Only one study has previously manipulated all three variables in the one study (Hand et al., 2010) however using launch distance to index parafoveal processing. Therefore, the present study was the first to investigate these three factors by using valid-invalid previews conditions, implemented in a between-groups design. Thus, preview was manipulated across two separate studies. In fact, this method is unusual, with existing studies investigating either word frequency by parafoveal preview (e.g., Inhoff & Rayner, 1986; Reingold et al., 2012) or contextual predictability by parafoveal preview (e.g., Balota et al., 1985) using a within-subjects design.

This present analyses showed the same pattern of results in main effects across first and single fixations where there were significant main effects of word frequency, contextual predictability and preview. All main effects were in the expected direction. In both measures, there was a significant interaction of word frequency and contextual predictability, though in the single fixation measure this interaction was approaching significance (marginal in subjects and trend in items). Word frequency by preview was nonsignificant in both first and single fixations. Therefore, it could not be concluded that word frequency modulates the use of parafoveal information (e.g., Inhoff & Rayner, 1986; Hand et al., 2010; Reingold et al., 2012). Contextual predictability by preview was hypothesised to be nonsignificant in both first and single fixations, but due to the unexpected direction of results in the invalid preview condition, this interaction was significant in both measures (though marginal in items for single fixations). We were not able to conclude that contextual predictability effects, that is, top-down sources arising from the sentence context, modulates the use of parafoveal information where the higher the contextual predictability, the more lexical access is facilitated. The three-way interaction was expected to be significant in first and single fixations but was nonsignificant. This could be because of loss of power because preview was a between-groups factor.

For the skipping data, it was hypothesised that there would be significant main effects of all three variables and that all two-way interactions as well as the three-way interaction would be nonsignificant. The main effects of word frequency and preview were significant but contextual predictability was not. It was suggested that this was likely to be because in Experiment 3, predictability contrasts were in the unexpected direction, namely that low predictable targets were skipped significantly more than high predictable ones leading to a pattern of skipping where differences between groups were cancelled out. Word frequency by contextual predictability was nonsignificant, as was frequency by preview (though results were approaching significance here). Contextual predictability by preview was unexpectedly significant, however a closer inspection of results showed some predictability contrasts were in the unexpected direction. Overall, a stronger design would be to manipulate the factor of preview (valid versus invalid) within-subjects. However, the present study was well designed in that the nonword preview letter strings were constructed so as they did not contain ‘weird’ letter sequences y could have commanded attention.

The implication of the above results in relation to the additive-interactive debate is not firmly conclusive. The two-way interaction effects (frequency x contextual predictability; word frequency by preview; contextual predictability by preview) were not expected to show significance in first and single fixations given the pattern of results in Experiments 2 and 3. Frequency x contextual predictability was significant only in first fixations, failing to reach significance in single fixations. The contextual predictability by preview interaction was significant but likely to be due to results in the unexpected direction in the invalid preview condition. Contextual predictability by preview did not manifest in the three-way interaction; this interaction was expected to be significant. This could have been due to manipulating preview between-groups, which is less powerful than within-subjects. It was decided that the relationship between frequency, predictability and preview in a within-subjects design was necessary (reported in **Experiment 4**). However, the current study was a good starting to point to investigating the single and combined effects of word frequency, contextual predictability and preview. Of particular merit was the use of preview nonwords that did not command attention in the text.

Chapter 5

Word frequency, contextual predictability and parafoveal preview effects in an eye movement reading experiment

Pre-test 3

Cloze task

5.1 Method

5.1.1 Participants

Twenty participants took part in the Cloze word generation task (5 male, 15 female; mean age 24 years). All participants Native English speakers and attending the University of Glasgow as undergraduate or postgraduate students. Participation was voluntary. None of the participants took part in any of the eye tracking experiments reported in this thesis. Participants did not have any learning or reading disorders which could have inhibited performance on a normal reading task.

5.1.2 Apparatus

The Cloze task was administered using the pencil-paper method. The passages of text were typed onto A4 pages and stapled together to make a booklet. The instructions were typed onto the first page of this booklet.

5.1.3 Materials

The presentation of the materials followed the same format as previous Cloze pencil-paper tasks in earlier experiments in this thesis. That is, the passages were presented up to but not including the target or the post-target context. Participants were instructed to read the lines of text and then to write in what they think the next word in the passage should be. They were instructed that this should be one word, and was a continuation of the passage rather than the final word. A blank underlined space was provided after each passage in order for the participant to write down what they think they word should be given what they had just read. Figure 5.1. below shows the presentation of four example passages. Conditions are labelled in this Figure only; participants were presented with a random order of passages with no reference to which condition the experimenter thought it should be in. Subsequent to Results analyses, these passages were assigned to one of the eight conditions.

Figure 5.1. Example presentation used in the Cloze task

[HF-HP]

Robbie and his dad were getting ready to play catch.

They finally found the _______________

[HF-LP]

It was getting dark and we made our final approach with great care.

In the distance we saw a _______________

[LF-HP]

The triumphant King arranged a sumptuous and lavish banquet.

It was a delightful _______________

[LF-LP]

Miss Dearborn lived at number 31 in Alder Crescent.

The older kids said that she was a _______________

5.1.4 Design

The design of the task reflected the desired design of the eye tracking experiment. That is, a within-subjects design was used (with the exemption of preview, since it is not possible to have participants write in a nonword). Therefore for the purpose of the Cloze task, the design was a 2 x 2 design: frequency (HF, LF) on predictability (HP, LP). The dependent variable was the word written in the blank space with responses scored as “1” if the desired target word (according to the experimenter’s intuition) was written down and “0” for any other word given. The final set of chosen 240 passages is shown in Appendix E. Appendix F displays the final Cloze probability obtained for each target in its high and low predictable context (as well as the target word characteristics).

5.1.5 Procedure

Participants who had agreed to take part in the experiment were given the task along with the instructions and consent form. When they had read the latter and further questions were clarified, they were ready to begin the task. The experiment took around 45 minutes to complete. After completion, participants were fully debriefed.

5.2 Results

The mean Cloze probability for HF-HP targets was 0.67 (0.2); for HF-LP targets 0.07 (0.1), for LF-HP targets 0.60 (0.3) and for LF-LP 0.05 (0.1). The ANOVA was carried out by items (F_2) analysis. Follow up contrasts made use of Bonferroni tests.

The results showed that the main effect of frequency was significant [$F_2(1,119)=23.52$, $MSE=.008$, $p<.001$]. This means that HF targets had a significantly higher Cloze probability than LF targets (HF: 0.37 vs. LF: 0.33). We would not want this; however numerically this was a very small difference. The results also showed a significant main effect of predictability [$F_2(1,119)=680.07$, $MSE=.058$, $p<.001$]. This means that high predictable targets had a significantly higher Cloze probability than low predictable ones (HP: 0.63 vs. LP: 0.06).

The interaction of frequency and predictability was also significant [$F_2(1,119)=8.34$, $MSE=.006$, $p<.01$]. The first two follow-ups to this interaction compared HF-HP vs. HF-LP and LF-HP vs. LF-LP. Since these contrasts compared low with high predictable targets for, first high frequency targets and second, low frequency targets, the expected direction was that both these contrasts would be significant. Actual results were as desired (HF-HP vs. HF-LP; 0.67 vs. 0.07: $F_2=3602.95$, $p<.001$ and LF-HP vs. LF-LP; 0.60 vs. 0.05: $F_2=3129.3$, $p<.001$).

The second set of contrasts compared high with low frequency targets at the high predictable condition and then at the low predictable condition. Since the design of the experiment was that high (and low) predictable Cloze probabilities were a clearly defined categories, the expected result would be that these contrasts would be nonsignificant. Indeed there were small numerical differences in these contrasts. However, HF-HP versus LF-HP was significant (0.67 vs. 0.60: $F_2=37.28$, $p<.001$) as was HF-LP versus LF-LP (0.07 vs. 0.05: $F_2=4.09$, $p<.05$).

5.2.1 Results summary

The results obtained in the Cloze task are summarised in Tables 5.1 and 5.2 below. Participants were presented with the complete passage of text prior to a blank underlined space which denoted that they should generate the next word in the passage. Basically, all effects were significant when only the main effect of predictability was expected to be. The main effect of frequency was significant showing that there were significant differences in Cloze probabilities of high and low frequency targets. The main effect of predictability showed that targets in the high predictable condition had significantly higher Cloze probabilities than those in the low predictable condition. For the interaction of frequency by predictability, despite small numerical differences in the Cloze probabilities between respective conditions, results reached significance when comparing Cloze probabilities obtained for low and high frequency targets at high and low predictable conditions respectively.

Table 5.1 Summary of ANOVA Results in Cloze Task by Items (F_2)

Main effect frequency					
Measure		***df***	***F***	***MSE***	***p***
Cloze probability	F_2	1,119	23.52	.008	<.001
Main effect contextual predictability					
Cloze probability	F_2	1,119	680.07	.058	<.001
Interaction word frequency x contextual predictability					
Cloze probability	F_2	1,119	8.34	.006	<.01

Table 5.2 Follow-ups when the Frequency by Predictability Interaction was Significant

		HF-HP vs. HF-LP		**LF-HP vs. LF-LP**		**HF-HP vs. LF-HP**		**HF-LP vs. LF-LP**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Cloze probability	F_2	3602.95	<.001	3129.3	<.001	37.28	<.001	4.09	<.05

Note: HF = high frequency; LF = low frequency; HP = high predictability; LP = low predictability

5.3 Discussion

The Cloze task was used to present participants with a sentence frame with instructions to generate the next word in the passage, given the context they had read up to that point. Care was taken to ensure that these participants had not taken part in earlier experiments reported in this thesis or the later eye tracking experiment reported in this Chapter. Participants were also instructed to read the passage of text carefully before writing in one word in the underlined space. Participants were also informed that their chosen word would be a continuation of the passage rather than the last word in the passage. The order of passages of text was arranged randomly and there was no reference made to the experimenter's intuition as to what condition each passage best fitted.

The results of the experiment showed an unexpected significant main effect of frequency. This means that high frequency targets had a significantly higher Cloze probability than low frequency ones. The numeric difference here was so small that it would not have been expected to be significant (HF: 0.37; LF: 0.33). It should be recalled that in the rating task reported in pre-test 1, the small numeric differences in word rating given to targets also reached significance in the frequency main effect. The possible reasons here could be due to when words are higher in frequency, they also become more predictable for readers.

The main effect of predictability, as expected, was significant. This effect was in the desired direction where high predictable targets obtained a significantly higher Cloze probability than low predictable ones (HP: 0.63 vs. LP: 0.06). This result suggests that two predictability conditions were manipulated appropriately and confirms the validity of our predictability variable.

The interaction of frequency and predictability was unexpectedly significant. Four follow-up contrasts were carried out to examine where there were significant differences in Cloze probabilities. The first two predictability contrasts compared high with low predictability Cloze probabilities for high frequency targets, and then for low frequency targets. We would expect both these sets of contrasts to be significant because the desired design of this experiment was that predictability conditions should be different from each other – high should differ from low predictable in terms of respective Cloze probabilities. In fact, both these contrasts were very highly significant as indicated by the very large *F* ratios. This result indicates that for high (and low) frequency targets, there were clear differences in Cloze probabilities between high and low predictable conditions.

The second set of contrasts were the frequency contrasts which compared high with low frequency targets at the high predictable condition and then at the low predictable condition. There were small numeric differences in these comparisons (in particular for the low predictable condition) however both sets of contrasts were significant. These results raise interesting questions about the task of generating a word for a given context. A reader is likely to start constructing a context about what is going on in a passage of text – and it could be that passages of text that subsequently go on to predict a high predictable target have different factors versus passages of text that go on to predict a low predictable target. There is a preference for high frequency targets, for example if a reader is asked to generate synonyms for a given word, they are likely to begin with words they happen to know and these words are likely to be high frequency ones. One result in Experiment 4 possibly points different strategies by the reader dependent on whether they are reading a high predictable or a low predictable context – this was the significant interaction of frequency by predictability in the reading of line 1 of the sentence frame. Most likely the reason for the significant interaction was that over the full set of materials, high frequency targets were always just higher in their mean Cloze completion in comparison to a low frequency target.

5.3.1 Conclusions

In the Cloze task, participants were given passages of text and were asked to generate the next word once they had read the sentence frame. We expected a nonsignificant main effect of frequency which would indicate that Cloze probability was equally matched over low and frequency targets. Also expected was a significant main effect of predictability which would indicate that targets in the high predictable condition contained targets with a higher Cloze probability than targets in the low predictable condition. Finally, a nonsignificant interaction between frequency and predictability would indicate that Cloze probability between respective conditions was equally controlled.

The main effect of predictability was significant. This indicates that our manipulation of the predictability variable was as desired where there were clear statistical differences in the Cloze probability obtained for high predictable targets and for low predictable targets. Both the main effect of frequency and interaction of frequency by predictability were unexpectedly significant. Despite there being small numerical differences in respective conditions that were compared, it is likely that the interaction was significant because high frequency words were always slightly higher in their Cloze probability than low frequency targets – the significant main effect of frequency shows that high frequency targets had a higher Cloze probability than low frequency ones. Over the full set of materials, this difference was probably large enough to be significant.

Once all targets had their associated Cloze probabilities for the two possible presentations i.e., a high predictable version of the sentence and a low predictable version of the sentence, passages were validated as belonging to one of two predictability conditions (see Appendices E and F for full material set and target word characteristics.

Experiment 4

5.4 Introduction

The final experiment in this thesis re-examined the effects of word frequency, contextual predictability and parafoveal preview. This was in order to address various issues which had arisen in the previous experiments, mainly to do with the manipulation of the preview variable. The overall aim of the present experiment was to design a study that allowed better manipulation of preview. Previously, parafoveal preview was manipulated as a between-groups factor in Experiment 3. The data collected in Experiment 2 was considered to form a condition of valid preview of the target and the eye tracking data collected in Experiment 3 was considered a condition of invalid parafoveal preview of the target. However this method of manipulating preview across two separate studies is unusual and the between-groups manipulation potentially introduced extraneous participant variables. Past studies have used a within-subjects design to investigate the effects of word frequency and parafoveal preview (e.g., Inhoff & Rayner, 1986; Reingold et al., 2012) or the effects of contextual predictability and parafoveal preview (e.g., Balota et al., 1985).

5.4.1 Present study

The present experiment used a three factor within-subjects design where the variables of word frequency, contextual predictability and parafoveal preview were manipulated. Materials were constructed so that targets were either low or high frequency and placed in sentence contexts that made them neutral, or low predictable and biasing, or high predictable. Two versions of materials were constructed so that each second line of text could be preceded by both a high predictable context and a low predictable context – participants read one version. Four sets of materials were made up (for four different participant groups) which sampled the whole set of materials (see Table 5.5).

Preview was manipulated by participants being presented with a valid preview of the target which was where the target appeared as usual in the text (since readers would be able to obtain the usual parafoveal information from the target on the pre-target fixation). For the invalid condition, we made sure that the same criteria we set out in Experiment 3 were fulfilled. That is, the nonword was pronounceable (orthographically legal), of equal character length to the target and had the same overall word shape to the target. This nonword occupied the space of the target until the readers eyes crossed the pre-specified boundary (just after the last character of the pre-target word).

We were interested in the main effects of frequency, predictability and preview. The three two-way interactions were frequency x predictability, frequency x preview and predictability x preview. The three-way was the interaction of frequency, predictability and preview. It was hypothesised that there would be significant main effects of the three variables in all fixation duration measures as well as skipping. We also expected the three two-way interactions to be significant along with the three-way interaction in first and single fixation duration measures and skipping.

5.5 Method

5.5.1 Participants

Eighty participants (42 female, 38 male) in total took part in the experiment. This was by having four groups of twenty participants read one of the four versions of the materials. All participants were both right-handed and native English speaking students attending the University of Glasgow (mean age 24 years old). Participation was entirely voluntary. Payment was made at the usual rate of £6 per hour. Participants had normal or corrected-to-normal vision and had not been diagnosed with any reading or learning disorder. The predictability of target words was determined by administering the Cloze task to a separate group of 20 participants (see pre-test 3). Participants who took part in this Cloze task did not partake in the main eye tracking experiment since they would have been familiar with the materials.

5.5.2 Apparatus

Participant’s eye movements were recorded using an EyeLink® 2K eye tracker manufactured by SR Research Ltd. (Mississauga, Ontario, Canada). The sampling rate was 1000 Hz using corneal reflection and pupil tracking and the spatial resolution was 0.01˚. The gaze tracking range is 32˚ horizontally and 25˚ vertically. All 240 passages of text, as well as the practice trials, were presented using the 14-point Bitstream Vera Sans Mono font (black characters on a white background) on a Dell P1130 19" flat screen CRT (1024 x 768 resolution; 100 Hz). A monocular desktop mount incorporated the camera and sat just below the CRT display. At a viewing distance of approximately 72cm, approximately 4 characters equalled 1˚ of visual angle. Viewing was binocular and eye movements were recorded from the right eye. A forehead/chin rest were used to minimise head movements.

5.5.3 Materials

There were 120 high frequency targets and 120 low frequency targets. The complete set of 240 passages of text is presented in Appendix E. Appendix F shows the individual specifications for every target – for example, a given word's frequency; age-of acquisition score and Cloze neutral and biasing probability. Table 5.3, shown below, presents the average of these specifications.

Table 5.3 Summary Specifications of 240 Low and High Frequency Targets

Condition	Mean lett. (SD)	Mean freq. per mill. (SD)	Mean no. syll. (SD)	Mean no. phon. (SD)	Mean no. morph. (SD)	Mean ortho. neigh. (SD)	Mean phono. neigh. (SD)	Mean AoA. (SD)	Mean img. rating (SD)	Mean cnc. rating (SD)	Mean Cloze Neutral (SD)	Mean Cloze Biasing (SD)
HF targets (items 1-30)	4.77 (.43)	102.29 (86.35	1.13 (.35)	3.83 (.7)	1.17 (.38)	7.43 (4.96)	17.43 (11.84)	270.19 (84.34)	537.09 (108.54)	499.59 (125.45)	.07 (.12)	.65 (.25)
HF targets (items 31-60)	4.77 (.43)	103.71 (58.63	1.00 (.00)	4.03 (.81)	1.2 (.41)	6.77 (5.49)	14.33 (12.87)	307.00 (99.65)	538.34 (98.1)	502.67 (118.15)	.07 (.12)	.66 (.25)
HF targets (items 61-90)	4.77 (.43)	105.44 (90.97	1.3 (.53)	3.73 (.69)	1.13 (.69)	5.4 (4.92)	15.67 (12.52)	291.45 (67.28)	541.85 (90.28)	483.82 (133.21)	.06 (.12)	.67 (.24)
HF targets (items 91-120)	4.77 (.43)	103.7 (122.9	1.17 (.38)	3.9 (.76)	1.1 (.31)	6.07 (5)	16.63 (12.39)	249.06 (68.82)	541.11 (120.36)	497.72 (141.04)	.09 (.13)	.68 (.22)
LF targets (items 121-150)	4.77 (.43)	7.78 (5.51)	1.33 (.48)	4 (.64)	1.2 (.41)	5.13 (4.73)	12.40 (11.26)	322.77 (88.93)	531.55 (101.77)	531.8 (104.33)	.04 (.07)	.61 (.27)
LF targets (items 151-180)	4.77 (.43)	8.06 (4.89)	1.3 (.47)	3.83 (.79)	1.23 (.43)	4.53 (3.62)	13.37 (11.98)	345.06 (91.19)	568.76 (97.08)	554.55 (101.26)	.04 (.07)	.60 (.28)
LF targets (items 181-210)	4.77 (.43)	8.31 (4.95)	1.43 (.57)	4.14 (.79)	1.28 (.45)	5.83 (4.77)	13.41 (12.67)	371.81 (113.29)	531.25 (90.83)	508.69 (125.94)	.05 (.08)	.61 (.25)
LF targets (items 211-240)	4.77 (.43)	7.51 (6.71)	1.33 (.55)	4 (.71)	1.21 (.41)	6 (4.99)	14.79 (11.7)	305.81 (74.67)	555.35 (96.21)	539.6 (101.43)	.07 (.12)	.61 (.25)

Note: SD = standard deviation; HF = high frequency; LF = low frequency; lett.= number of letters; freq. per mill.= frequency per million; no. syll.= number of syllables; no. phon. = number of phonemes; no. morph.=number of morphemes; ortho. neigh.= orthographic neighbours; phono. neigh.= phonological neighbours; AoA = age-of acquisition, which ranges from 100 (low AoA; learned earlier in life) to 700 (high AoA; learned later in life); img.= imageability, which ranges from low 100 (low imageability) to 700 (high imageability); cnc = concreteness, which ranges from 100 (low concreteness) to 700 (high concreteness); Cloze values are the mean probability of participants guessing the correct target presented in their respective contexts; the higher the Cloze value, the greater the probability of a given word being guessed in its particular context

As can be seen in Table 5.3, items 1-120 were high frequency targets and items 121-240 were low frequency. Each high and low frequency target was preceded by a high and low predictable context. Moreover, the high frequency target corresponded to its low frequency target in terms of word length (matched exactly) and the other ratings which were matched (number of syllables; number of morphemes; orthographic neighbours; Cloze neutral and biasing etc.). High frequency items 1-30 corresponded with low frequency items 121-150; high frequency items 31-60 with low frequency 151-180; high frequency 61-90 with low frequency 181-210 and high frequency 91-120 with low frequency 211-240.

As in previous experiments in this thesis, word frequencies were acquired from the British National Corpus (BNC), a database of 90 million written words (http://natcorp.ox.ac.uk), frequency values for HF were occurrences of 27 million or above (range: 27.89-524.78 occurrences per million); and LF were occurrences of 24 million or below (range: 0.17-23.11 occurrences per million). Across the eight conditions, word length was matched exactly (average word length 4.77 characters; range: 4-5 characters; SD = .43). The number of syllables did not differ across the eight conditions, and the mean ranged from 1 to 1.43 syllables.

Age of acquisition ratings were collected by presenting participants with the list of targets, and giving them written instructions to rate each word according to how early in life they think they learned the word and its meaning (in either written or spoken form). A 7-point scale was used (2 year age bands starting from age 0-2, up to 13+), where the lower the age-of-acquisition, the earlier in life the word was acquired.

Imageability ratings were collected were obtained from five norms – Cortese and Fugett's (2004) imageability norms; the Bristol Norms (Stadthagen-Gonzalez & Davis, 2006; http://language.psy.bris.ac.uk/); Gilhooly and Logie's (1980) imageability norms; Morrison, Chappell, & Ellis's (1997) imageability norms and finally the MRC Psycholinguistic Database (Wilson, 1988; http://www.psych.rl.ac.uk/bristol_norms.html). These norms reflect participants' ratings of words on a scale of 100 (low imageability) to 700 (high imageability). Concreteness ratings were collected from the MRC Psycholinguistic Database (Wilson, 1988; http://www.psych.rl.ac.uk/bristol norms.html). If words did not have imageability and/or concreteness rating, then a group of participants were recruited to acquire these. The same 7-point scale was used as in previous studies (e.g., Paivio, Yuille, & Madigan, 1968), where the lower the rating, the less the word was judged to be in its imageability or concreteness.

We accessed the English Lexicon Project (ELP) website (http://elexicon.wustl.edu) where our selected target words were input to obtain the following: mean number of phonemes; morphemes; orthographic and phonological neighbours. As can be seen from Table 5.3 the range of these values remained small: mean number of phonemes range 3.73-4.14; mean number of morphemes range 1.1-1.28; mean orthographic neighbours range: 4.53-7.43 and mean phonological neighbours range: 12.40-17.43. There were only two targets which were not entries in the ELP ('salsa' and 'chewy').

The Cloze task was used to determine the Cloze probability of high and low predictable passages of text. As can be seen from Table 5.3, the mean Cloze probability for ‘high predictable’ passages of text was between 60% to 68% completion and for ‘low predictable’ from 4% to 9% completion. Pre-test 3 on p. 245 at shows details of the Cloze task experiment used in this study.

5.5.4 Design

HF and LF targets were manipulated so that they were accommodated in high predictable and low predictable contexts, as well as with valid and invalid previews of the target. This was for four different groups of participants, with each set containing 240 passages (thus, 240 targets), manipulated on frequency, predictability and preview. Thus, the design of the experiment was a 2 (Frequency: high frequency, HF; low frequency, LF) x 3 (Predictability: high predictability, HP; low predictability, LP) x 2 (Preview: Valid; Invalid). This lead to the creation of 8 experimental conditions: HF-HP Valid; HF-HP Invalid; HF-LP Valid; HF-LP Invalid; LF-LP Valid; LF-LP Invalid; LF-LP Valid and LF-LP Invalid. Targets were all 4 and 5 letters in length and the majority were used predominantly as nouns in the language. Table 5.4 below shows the design of the experimental passages of text over four groups of participants.

5.5.5 Procedure

A participant arrived to take part in the experiment at the pre-specified time. They were given a consent form to read and sign. Once this was done, they were seated for the experiment and the forehead/chin rest were adjusted so that they were at the correct height for the participant. The participant was given written and verbal instructions about the task. They were told to read the passages of text for comprehension and that they would sometimes see questions about the content of the passages which would require a yes-no answer. The experiment was then calibrated which took about 5 minutes: the participant followed a series of dots around the screen; these calibration points extended over the full horizontal and vertical range over which the passages were presented.

Table 5.4 Design of Experiment

		Condition							
		HF				LF			
		HP		LP		HP		LP	
		V	I	V	I	V	I	V	I
Participant group (N=80)	Group 1 (n=20)	1-30	31-60	61-90	91-120	121-150	151-180	181-210	211-240
	Group 2 (n=20)	31-60	1-30	91-120	61-90	151-180	121-150	211-240	181-210
	Group 3 (n=20)	61-90	91-120	1-30	31-60	181-210	211-240	121-150	151-180
	Group 4 (n=20)	91-120	61-90	31-60	1-30	211-240	181-210	151-180	121-150

Note: Appendix E shows items 1-240 and Appendix F displays target word characteristics of items 1-240

Eight practice trials were presented (and were the same for every participant) and contained all possible variations that the participant would encounter in the actual trials. For example, valid and invalid targets were presented as well as questions. This was then followed by 240 experimental trials. Recalibration was performed whenever necessary. Each passage began when participants fixated the upper left-most calibration point which corresponded to the first character of text – this was a black square box. When they had read the passage, participants looked away to the bottom-right of the screen and pressed a button on a hand-held device to clear the screen. The black box re-appeared immediately or, if the passage had a question, the participant was required to answer it. There was no difficulty in answering the questions and participants were correct over 90% of the time.

The boundary technique was used in the invalid preview condition. The purpose of the invalid preview condition was to prevent parafoveal processing of the target on a prior fixation. This was achieved by replacing the target with a nonword, and once participants made a saccade over the last letter of the pre-target word, this nonword changed to the actual target. The nonword fulfilled several criteria in order to ensure that it did not command attention by, for example, have strange letter sequences or being much shorter or longer in length compared to the eventual target. These criteria were that ascenders, descenders and small letters in the target were replaced with ascenders, descenders and small letters in the nonword. Thus, word shape between target and nonword were highly similar. In addition, the resulting nonword letter string was pronounceable given the rules of English orthography and there was no letter overlap between respective letters in target and nonword.

Participants were given a break half way through the experiment. The experiment took approximately 50 minutes to complete. At the end of the experiment, participants were asked if they had any comments they wanted to make regarding the trials. Most reported noticing changes taking place in the text but were not able to say exactly what letters or words they had seen prior to this change. Participants were fully debriefed before leaving the lab.

5.6 Results

The target region included the space before the target word and the target word itself. Data were eliminated from analyses for the following reasons: (a) a blink or track loss occurred on the target word; (b) the fixation on the target was either the first or the last fixation on the line; or (c) the duration of any individual fixation shorter than 100 ms or longer than 750 ms.

Eye movement dependent measures on the target word which were analysed were: (a) first fixation duration (FFD; the duration of the very first fixation on the target word, provided that the word was not skipped, regardless of whether it is the only fixation or the first of many fixations on that word); (b) single fixation duration (SFD; the duration of the first fixation on the target word if it received only one fixation on the first pass); (c) gaze duration (GD; the sum of all fixations on a target word on the first pass prior to making an eye movement to another word); and (d) total time spent fixating the target word (TT; the sum of all fixations on the target word including any re-reading of the target word i.e., regressions back to it). Also computed was the amount of skipping the target as indicated by the percentage of skips (%Skip). Two-way within-subjects repeated measures analyses of variance (ANOVAs) as well as the three-way ANOVA were carried out on each measure both by subjects (F_1) and by items (F_2).

Table 5.5 shows the mean fixation time in milliseconds on the target (which was always in line 2) and also the percentage of times the target was skipped i.e. when it was not fixated as well as the first pass reading time of sentence 1 of the passage of text when computed as one region (in milliseconds per character). The mean SFD with standard error bars are displayed in Figure 5.2 below.

Table 5.5 Mean Fixation Time Results in Milliseconds (Standard Deviation) and Percentage of Times the Target was Skipped

Measure	**HF**			**LF**	
	HP	**LP**		**HP**	**LP**
FFD					
Valid	204 (28)	210 (28)		205 (30)	213 (30)
Invalid	244 (42)	251 (46)		254 (45)	258 (46)
SFD					
Valid	205 (30)	211 (28)		206 (31)	215 (31)
Invalid	252 (50)	262 (60)		265 (56)	272 (60)
GD					
Valid	212 (33)	219 (33)		213 (36)	224 (36)
Invalid	269 (59)	276 (70)		285 (62)	289 (68)
TT					
Valid	234 (43)	246 (46)		237 (48)	255 (47)
Invalid	303 (77)	317 (77)		320 (79)	339 (92)
%Skip					
Valid	32 (15)	28 (14)		29 (14)	25 (15)
Invalid	14 (14)	16 (14)		14 (11)	13 (12)
Sent 1					
Valid	31.4 (9.13)	31.8 (9.21)		32.1 (9.05)	31.5 (8.89)
Invalid	31.4 (9.08)	31.9 (8.8)		32.1 (9.22)	31.4 (9.08)

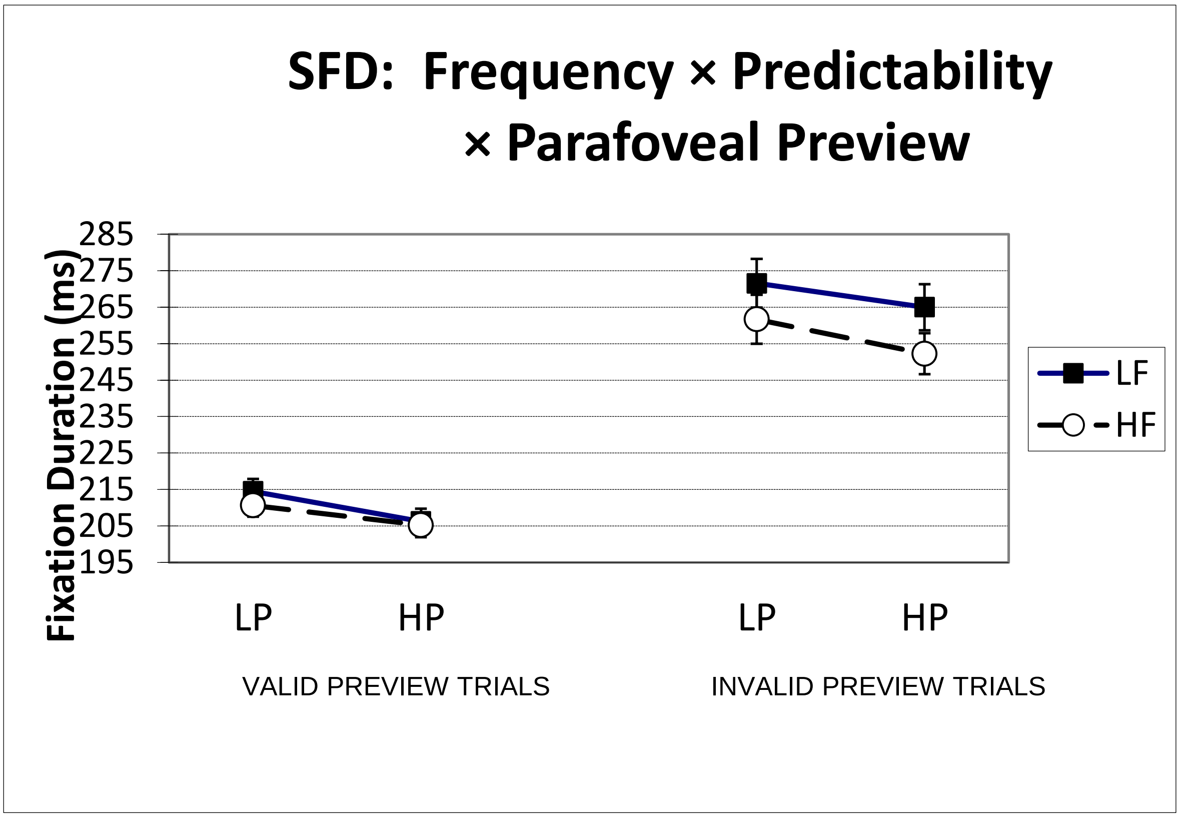
Note: HF = high frequency; LF = low frequency; HP = high predictability; LP = low predictability; FFD = first fixation duration; SFD = single fixation duration; GD = gaze duration; TT = total time; %Skip = percentage of times the target was skipped; Sent 1 = sentence 1 first pass reading, in milliseconds per character

5.6.1 First fixation duration

5.6.1.1 Main effects

All three main effects were significant in both subjects and items analyses. There was a significant main effect of frequency [F_1(1,79)=16.62, *MSE*=247, *p*<.001, F_2(1,119)=10.05, *MSE*=691, *p*<.01]; HF targets were fixated for less time than LF words (227 ms vs. 232 ms respectively). Also significant was the main effect of predictability [F_1(1,79)=18.83, *MSE*=287, *p*<.001, F_2(1,119)=13.38, *MSE*=527, *p*<.001]; HP targets were fixated for less time than LP targets (227 ms vs. 233 ms respectively). Finally, the main effect of preview was significant [F_1(1,79)=142.21, *MSE*=2155, *p*<.001, F_2(1,119)=674.05, *MSE*=662, *p*<.001], that is, valid preview targets were fixated for less time than those viewed with an invalid preview (208 ms vs. 252 ms respectively).

Figure 5.2. Average single fixation duration (SFD) as a function of frequency, predictability and parafoveal preview

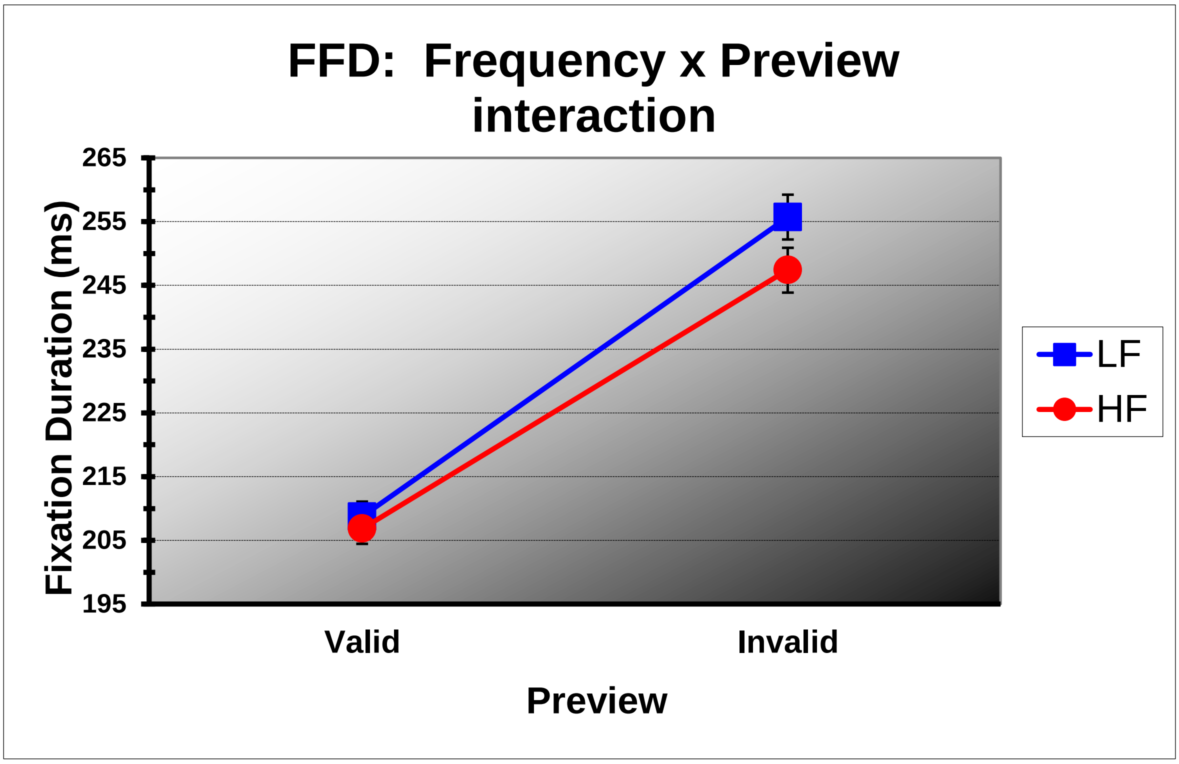


Note: SFD = single fixation duration; LF = low frequency; HF = high frequency; LP = low predictability; HP = high predictability; ms = milliseconds

5.6.1.2 Interactions

Regarding the interactions, only frequency x preview was significant and only in subjects analysis; for items analysis the interaction was nonsignificant [F_1(1,79)=4.82, *MSE*=338, *p*<.05, F_2(1,119)=1.79, *MSE*=656, *p*>.15]. This interaction is displayed in Figure 5.3. below. The frequency x predictability interaction was nonsignificant as was the predictability x preview interaction [all *Fs*<1]. The three-way frequency x predictability x preview interaction did not reach significance [F_1(1,79)=1.03, *MSE*=250, *p*>.30, F_2<1].

Figure 5.3. Frequency by preview interaction in first fixation duration



Note: FFD = first fixation duration; LF = low frequency; HF = high frequency; ms = milliseconds

For the frequency x preview significant interaction (in subjects), three out of four Bonferroni multiple comparisons follow-ups were significant: HF targets viewed with a valid preview were fixated for significantly less time than HF targets in an invalid preview (207 ms vs. 247 ms respectively: F_1=389.57, p<.001; similarly for the LF targets, those viewed in a valid preview, was for significantly less time than LF targets in an invalid preview (209 ms vs. 256 ms respectively: F_1=521.74, p<.001). For the final two contrasts, one was significant: for targets presented in an invalid preview, HF targets were fixated for significantly less time than LF targets (247 ms vs. 257 ms: F_1=16.13, p<.001). However for the valid presentation, there was no significant difference between high and low frequency targets (207 ms vs. 209 ms: F_1<1). This means that there was no significant frequency effect for targets viewed with a valid preview; only for those viewed with an invalid preview.

5.6.2 Single fixation duration

5.6.2.1 Main effects

As in the first fixation duration measure, in the SFD measure, all three main effects of frequency, predictability and preview were significant in both subjects and items analyses. For the main effect of frequency the results were: [$F_1(1,79)=20.27$, $MSE=373$, $p<.001$, $F_2(1,119)=11.72$, $MSE=1011$, $p<.001$]. Thus, HF targets were fixated for significantly less time than LF targets (232 ms vs. 239 ms respectively). For the main effect of predictability, the results were: [$F_1(1,79)=18.87$, $MSE=468$, $p<.001$, $F_2(1,119)=14.7$, $MSE=663$, $p<.001$]. This means that HP targets were fixated for significantly less time than LP targets (232 ms vs. 240 ms respectively). Finally the main effect of preview was: [$F_1(1,79)=125.65$, $MSE=3639$, $p<.001$, $F_2(1,119)=680.79$, $MSE=829$, $p<.001$]; valid preview targets were fixated for significantly less time than invalid preview targets (209 ms vs. 263 ms).

5.6.2.2 Interactions

The results for the two-way and three-way interactions followed the same pattern as in the first fixation. Only frequency x preview was significant and was so in subjects analysis; in items, the interaction was just beyond trend [$F_1(1,79)=8.63$, $MSE=365$, $p<.01$, $F_2(1,119)=1.89$, $MSE=853$, $p>.15$]. This interaction is displayed in Figure 5.4. below. The frequency x predictability interaction was nonsignificant [all $Fs<1$]. The predictability x preview interaction was nonsignificant [all $Fs<1$] as was the three way frequency, predictability and preview interaction [all $Fs<1$].

Bonferroni follow-ups conducted to the significant (in subjects) frequency x preview interaction showed that three out of four contrasts were significant. First, HF valid targets were fixated for significantly less time than HF invalid targets (208 ms vs. 257 ms respectively: $F_1=526.4$, $p<.001$; also LF valid targets were fixated for significantly less time than LF invalid targets (210 ms vs. 268 ms respectively: $F_1=724.24$, $p<.001$. Also significant was the difference between HF and LF targets – in the invalid preview condition (HF: 257 ms vs. LF: 268 ms, $F_1=26.04$, $p<.001$). The same frequency comparison for valid preview was nonsignificant [$F_1=1.31$, $p>.25$]. Thus, the frequency effect was only significant for targets viewed with an invalid preview where HF targets were viewed for significantly less time than LF ones.

Figure 5.4. Frequency by preview interaction in single fixation duration

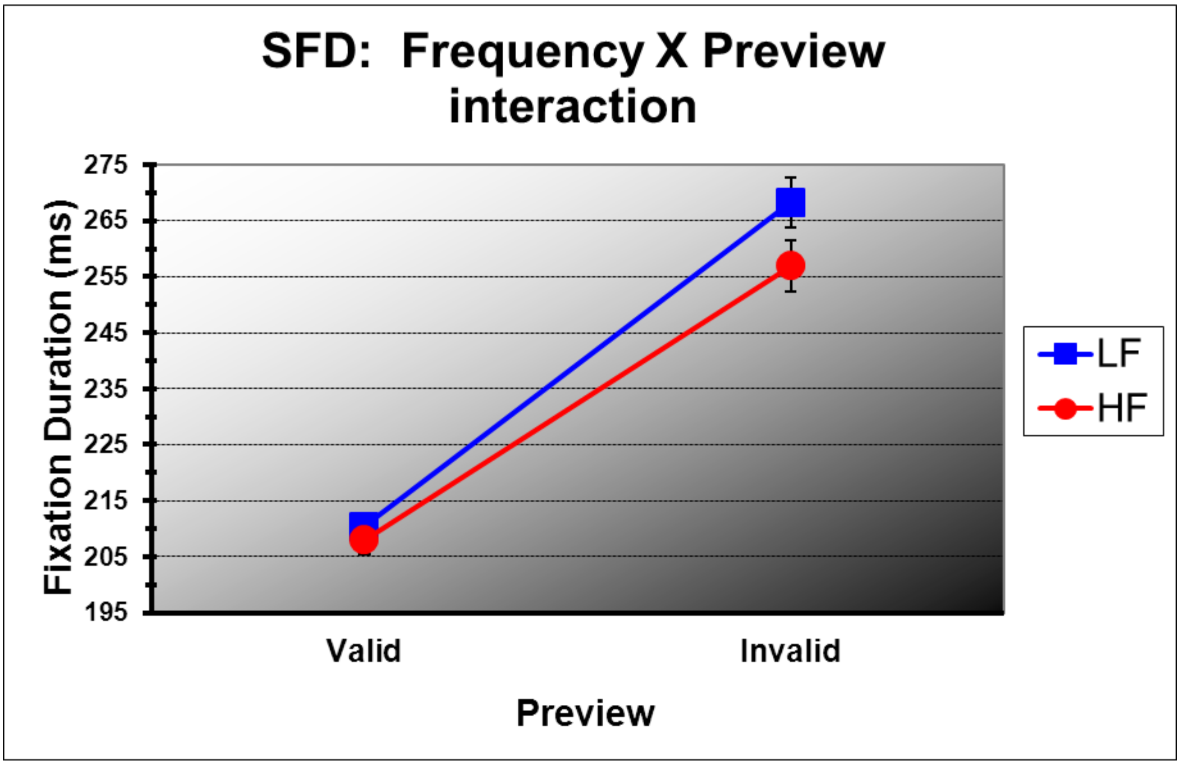


Figure 5.3. Note: SFD = single fixation duration; LF = low frequency; HF = high frequency; ms = milliseconds

5.6.3 Gaze duration

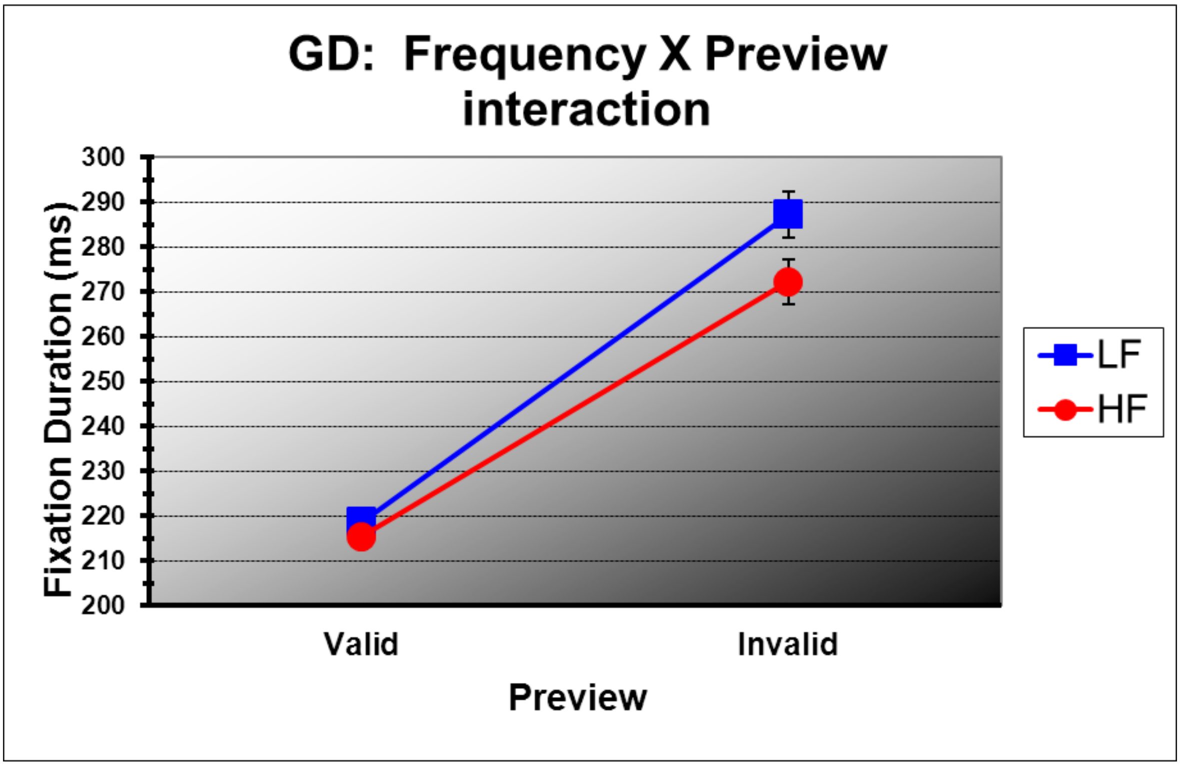
5.6.3.1 Main effects

All three main effects were significant: [frequency: $F_1(1,79)=34.17$, $MSE=389$, $p<.001$, $F_2(1,119)=2.18$, $MSE=1072$, $p<.001$; predictability: $F_1(1,79)=12.83$, $MSE=633$, $p<.001$, $F_2(1,119)=11.66$, $MSE=845$, $p<.001$; preview: $F_1(1,79)=136.95$, $MSE=4579$, $p<.001$, $F_2(1,119)=704.09$, $MSE=1283$, $p<.001$]. HF targets were fixated for significantly less time than LF ones (244 ms vs. 253 ms respectively). HP targets were fixated for significantly less time than LP ones (245 ms vs. 252 ms respectively). Valid preview targets were fixated for significantly less time than invalid preview ones (217 ms vs. 280 ms respectively).

5.6.3.2 Interactions

The results for the three two-way and three-way interactions followed the same pattern as in the first and single fixation measures except that the frequency x preview interaction was now significant in both subjects and items analyses [$F_1(1,79)=13.01$, $MSE=423$, $p<.001$, $F_2(1,119)=6.31$, $MSE=926$, $p<.05$]; this is shown in Figure 5.5. The frequency x predictability interaction as well as the predictability x preview interactions were nonsignificant [all *Fs*<1]. The three way interaction of frequency, predictability and preview did not reach significance [$F_1(1,79)=1.46$, $MSE=459$, $p>.20$, $F_2(1,119)=1.19$, $MSE=588$, $p>.25$].

Figure 5.5. Frequency by preview interaction in gaze duration



Note: GD = gaze duration; LF = low frequency; HF = high frequency; ms = milliseconds

For the significant interaction of frequency x preview, the follow-up contrasts followed the same patterns of significance as in first and single fixation measures. First, HF valid targets were fixated for significantly less time than HF invalid targets (215 ms vs. 272 ms: $F_1=608.10$, $p<.001$, $F_2=412.56$, $p<.001$); in addition, LF valid targets were also fixated for significantly less time than LF invalid ones (219 ms vs. 287 ms: $F_1=885.73$, $p<.001$, $F_2=569.47$, $p<.001$). Third, HF targets were fixated for significantly less time than LF ones – for the invalid preview condition (272 ms vs. 287 ms: $F_1=42.43$, $p<.001$, $F_2=26.95$, $p<.001$). However, the same frequency contrast for the valid preview condition was nonsignificant ($F_1=2$, $p>.15$, $F_2=2.69$, $p=.104$).

5.6.4 Total time

5.6.4.1 Main effects

The three main effects were all significant in subjects and items analyses [frequency: $F_1(1,79)=33.96$, $MSE=758$, $p<.001$, $F_2(1,119)=11.06$, $MSE=3533$, $p<.01$;

predictability: $F_1(1,79)=43.62$, $MSE=938$, $p<.001$, $F_2(1,119)=35.36$, $MSE=1670$, $p<.001$; preview: $F_1(1,79)=122.63$, $MSE=7723$, $p<.001$, $F_2(1,119)=759.71$, $MSE=1805$, $p<.001$]. Thus, participants spent a significantly shorter amount of total time on HF targets than on LF ones (275 ms vs. 288 ms). Participants also showed significantly shorter total time on HP targets compared to LP ones (273 ms vs. 289 ms). Finally, when targets were presented in a valid preview, participants total time was significantly shorter (243ms) than when targets were read in the invalid preview condition (320 ms).

5.6.4.2 Interactions

The frequency x preview interaction was significant in subjects and items analyses [$F_1(1,79)=10.34$, $MSE=634$, $p<.01$, $F_2(1,119)=5.01$, $MSE=1465$, $p<.05$]. Follow-ups to this showed that all four contrasts were significant in both subjects and items analyses. First, HF valid targets were fixated for significantly less time than HF invalid targets (240 ms vs. 310 ms: $F_1=628.1$, $p<.001$, $F_2=402$, $p<.001$); LF valid targets were also fixated for significantly less time than LF invalid ones (246 ms vs. 329 ms: $F_1=876.71$, $p<.001$, $F_2=538.96$, $p<.001$). Finally, for the valid preview trials and also for the invalid preview trials, participants spent significantly less time on HF targets than on LF ones (Valid preview trials: 240 ms vs. 246 ms; $F_1=4.98$, $p<.05$, $F_2=4.28$, $p<.05$; Invalid preview trials: 310 ms vs. 329 ms; $F_1=45.96$, $p<.001$, $F_2=27.4$, $p<.001$).

The frequency x predictability interaction was nonsignificant [$F_1(1,79)=1.22$, $MSE=885$, $p>.25$, $F_2<1$) as was predictability x preview [all $Fs<1$]. The three way interaction of frequency, predictability and preview was nonsignificant [all $Fs<1$].

5.6.5 Percentage of times the target was skipped

5.6.5.1 Main effects

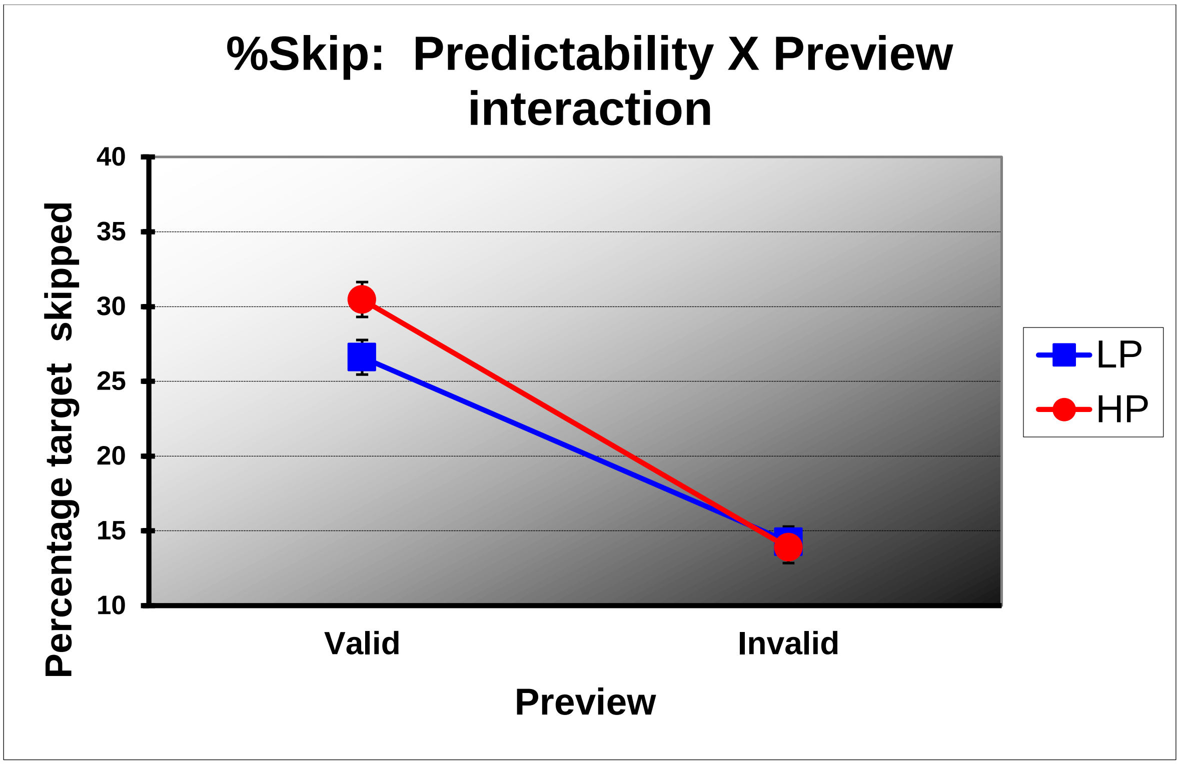
The mean percentage of skipping across the conditions is displayed in Table 5.3. All main effects were significant. The frequency main effect was significant [$F_1(1,79)=8.76$, $MSE=59$, $p<.01$, $F_2(1,119)=7.35$, $MSE=106$, $p<.01$]. Participants skipped significantly more targets when they were HF (22%) than when they were LF (20%). The predictability main effect was also significant [$F_1(1,79)=8.6$, $MSE=56$, $p<.01$, $F_2(1,119)=7.83$, $MSE=93$, $p<.01$]. Thus, participants skipped HP targets (22%) significantly more than LP ones (20%). The significant preview main effect [$F_1(1,79)=279.94$, $MSE=120$, $p<.001$, $F_2(1,119)=38$, $MSE=383.01$, $p<.001$] showed that participants skipped significantly more targets when they appeared in the valid preview condition (29%) than those that appeared in the invalid preview condition (14%).

5.6.5.2 Interactions

There was a significant interaction in predictability x preview [$F_1(1,79)=15.16$, $MSE=47$, $p<.001$, $F_2(1,119)=12.13$, $MSE=88$, $p<.001$]. This is shown in Figure 5.6. Bonferroni tests compared valid with invalid preview trials for, first, targets presented in a HP context; and then for targets presented in LP context. The third and fourth comparisons compared HP and LP targets in a valid preview, and then in an invalid preview. Basically all contrasts except the last one were significant. Thus, participants skipped significantly more HP-Valid targets (30%) than HP-Invalid ones (14%) [$F_1=466.3$, $p<.001$, $F_2=372.95$, $p<.001$] as well as more LP-Valid targets (27%) than LP-Invalid ones (14%) [$F_1=258.8$, $p<.001$, $F_2=206.98$, $p<.001$]. Also, participants skipped significantly more HP-Valid targets (30%) than LP-Valid ones (27%) [$F_1=25.18$, $p<.001$, $F_2=20.15$, $p<.001$]. There were no significant difference in skipping rates between HP and LP targets (both incurred 14% skipping) in an Invalid context [all *Fs*<1].

In terms of the other interactions in the skipping data, frequency x predictability was nonsignificant [$F_1(1,79)=1.8$, $MSE=46$, $p>.15$, $F_2(1,119)=1.62$, $MSE=77$, $p>.20$] as was frequency x preview [$F_1(1,79)=1.25$, $MSE=33$, $p>.25$, $F_2<1$] and so was the three-way interaction [$F_1(1,79)=1.55$, $MSE=34$, $p>.20$, $F_2<1$].

Figure 5.6. Predictability x preview interaction in percentage of times target was skipped



Note: %Skip = percent skipping; LP = low predictability; HP = high predictability

5.6.6 Sentence 1 fixation time across the conditions

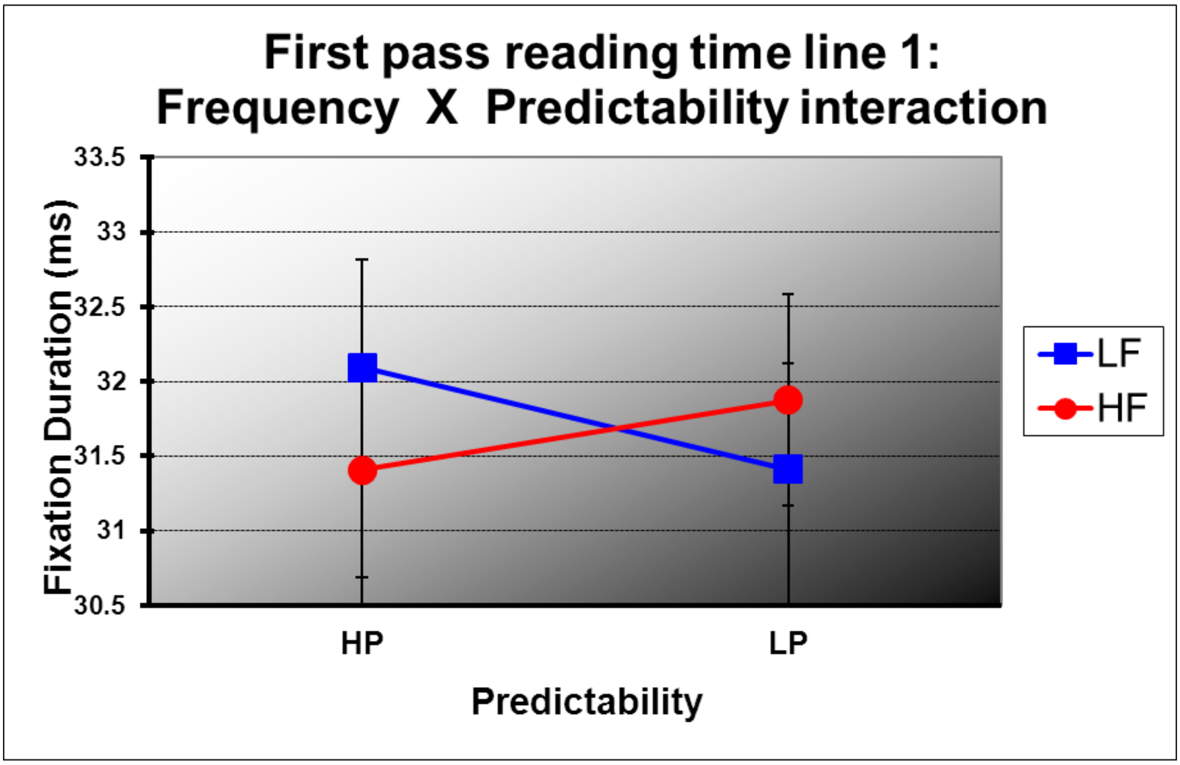
The mean number of characters in line 1 of the materials was always 59, 60 or 61 characters. Specifically, in the eight conditions (HF-HP Valid; HF-HP Invalid; HF-LP Valid; HF-LP Invalid; LF-HP Valid; LF-HP Invalid; LF-LP Valid and LF-LP Invalid), the mean number of characters in line 1 of the passages was (SD in parentheses): 59 (6); 59 (6); 61 (6); 61 (6); 59 (6); 59 (6); 60 (7) and 60 (7). Table 5.6 shows the mean first pass time (milliseconds per character) when line 1 was computed as one region. That is, across the 8 conditions, the first pass time in the reading of line 1 was 31.4 (9.13); 31.4 (9.08); 31.8 (9.21); 31.9 (8.8); 32.1 (9.05); 32.1 (9.22); 31.5 (8.89) and 31.4 (9.08). The resulting ANOVA showed that all three main effects of frequency, predictability and preview were nonsignificant [main effects of frequency and of predictability; all $Fs<1$; preview $F_1<1$, $F_2(1,119)=2.74$, $MSE=10$, $p=.1004$].

The interaction effects, except for frequency x predictability, were all as expected with nonsignificant results (frequency x preview; predictability x preview and frequency x predictability x preview, all $Fs<1$]. For the interaction of frequency x predictability (see Figure 5.7.), the result was unexpected in that subjects analysis were significant and items analysis were marginally significant [$F_1(1,79)=13.05$, $MSE=4$, $p<.001$, $F_2(1,119)=3.26$, $MSE=22$, $p=.074$]. Four Bonferroni follow-ups to this significant interaction were conducted: HF-HP versus HF-LP (31.4 ms vs. 31.9 ms); LF-HP versus LF-LP (32.1 ms vs. 31.4 ms); HF-HP versus LF-HP (31.4 ms vs. 32.1 ms) and HF-LP versus LF-LP (31.9 ms vs. 31.4 ms). We also looked at the items follow-ups since the interaction was marginally significant there. Basically, the results showed that all four contrasts were significant in subjects and nonsignificant in items: (HF-HP vs. HF-LP, $F_1=4.31$, $p<.05$, $F_2=1.49$, $p>.20$); (LF-HP vs. LF-LP, $F_1=9.21$, $p<.01$, $F_2=1.78$, $p>.15$); (HF-HP vs. LF-HP, $F_1=9.39$, $p<.01$, $F_2=2.05$, $p>.15$) and (HF-LP vs. LF-LP, $F_1=4.19$, $p<.05$, $F_2=1.26$, $p>.25$). Since the interaction was only significant in subjects analysis and all four contrasts were only significant in subjects, it is limited what conclusions can be drawn from contrasts that were all nonsignificant in items analysis.

5.6.7 Results summary

The results are summarised in Tables 5.6 to 5.10. Basically in FFD and SFD, the overall same pattern of results was observed. In both measures, there were significant main effects of frequency, predictability and preview. The interaction of frequency x preview was significant in subjects analysis and nonsignificant in items analysis. In the GD and TT

Figure 5.7. Frequency x predictability interaction in first pass reading time of line 1



Note: LF = low frequency; HF = high frequency; HP = high predictability; LP = low predictability

measures, a similar of results showed that all three main effects were significant. Only frequency x preview was significant in both subjects and items; follow-ups in GD showed that for the HF and LF targets, there were significant differences between valid and invalid preview conditions, as well as a significant frequency effect for targets in the invalid but not valid preview condition. The latter contrast was perhaps showing hints at reaching significance in the GD measure since this difference was trend in subjects and marginal in items. In the TT measure, all four contrasts were significant in subjects and items. The skipping data showed the three significant main effects. The significant interaction was the predictability x preview one. Direction of follow-up results were all as expected and significance was observed in three out of four contrasts. Finally the first pass time in the reading of line 1 was as expected with nonsignificant three main effects. All interactions were nonsignificant, except for the frequency x predictability but it was only so in subjects analysis and marginal in items. Furthermore, all follow-up items contrasts were nonsignificant.

Table 5.6 Summary of ANOVA Main Effect results in all Fixation Duration Measures and Skipping by Subjects (F_1) and by Items (F_2)

Main effect frequency						
Measure			***df***	***F***	***MSE***	***p***
FFD	**HF < LF**	F_1	1,79	16.62	247	<.001
		F_2	1,119	10.05	691	<.01
SFD	**HF < LF**	F_1	1,79	20.27	373	<.001
		F_2	1,119	11.72	1011	<.001
GD	**HF < LF**	F_1	1,79	34.17	389	<.001
		F_2	1,119	20.18	1071	<.001
TT	**HF < LF**	F_1	1,79	33.96	758	<.001
		F_2	1,119	11.06	3533	<.01
% Skip	**HF > LF**	F_1	1,79	8.76	59	<.01
		F_2	1,119	7.35	106	<.01
Sent. 1	**HF vs. LF**	F_1	1,79	*<1*	4	*ns*
		F_2	1,119	*<1*	36	*ns*
Main effect predictability						
FFD	**HP < LP**	F_1	1,79	16.62	247	<.001
		F_2	1,119	13.38	527	<.001
SFD	**HP < LP**	F_1	1,79	18.87	468	<.001
		F_2	1,119	14.7	663	<.001
GD	**HP < LP**	F_1	1,79	12.83	633	<.001
		F_2	1,119	11.66	845	<.001
TT	**HP < LP**	F_1	1,79	43.62	938	<.001
		F_2	1,119	35.36	1670	<.001
% Skip	**HP > LP**	F_1	1,79	8.6	56	<.01
		F_2	1,119	7.83	93	<.01
Sent. 1	**HF vs. LF**	F_1	1,79	*<1*	5	*ns*
		F_2	1,119	*<1*	43	*ns*
Main effect preview						
FFD	**Valid < Invalid**	F_1	1,79	142.21	2155	<.001
		F_2	1,119	674.05	662	<.001
SFD	**Valid < Invalid**	F_1	1,79	125.65	3639	<.001
		F_2	1,119	680.79	829	<.001
GD	**Valid < Invalid**	F_1	1,79	136.95	4579	<.001
		F_2	1,119	704.09	1283	<.001
TT	**Valid < Invalid**	F_1	1,79	122.63	7723	<.001
		F_2	1,119	759.71	1805	<.001
% Skip	**Valid > Invalid**	F_1	1,79	279.94	120	<.001
		F_2	1,119	383	131	<.001
Sent. 1	**Valid vs. Invalid**	F_1	1,79	*<1*	4	*ns*
		F_2	1,119	2.74	10	=.1004

Note: HF = high frequency; LF = low frequency; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Frequency x Predictability					
Measure		***df***	***F***	***MSE***	***p***
FFD	F_1	1,79	<*1*	350	*ns*
	F_2	1,119	<*1*	473	*ns*
SFD	F_1	1,79	<*1*	611	*ns*
	F_2	1,119	<*1*	590	*ns*
GD	F_1	1,79	<*1*	593	*ns*
	F_2	1,119	<*1*	640	*ns*
TT	F_1	1,79	1.22	885	>.25
	F_2	1,119	<*1*	1853	*ns*
% Skip	F_1	1,79	1.8	46	>.15
	F_2	1,119	1.62	77	>.20
Sent. 1	F_1	1,79	13.05	4	<.001
	F_2	1,119	3.26	22	=.074
Frequency x Preview					
FFD	F_1	1,79	4.82	338	<.05
	F_2	1,119	1.79	656	>.15
SFD	F_1	1,79	8.63	365	<.01
	F_2	1,119	1.89	853	>.15
GD	F_1	1,79	13.01	423	<.001
	F_2	1,119	6.31	926	<.05
TT	F_1	1,79	10.34	634	<.01
	F_2	1,119	5.01	1465	<.05
% Skip	F_1	1,79	1.25	33	>.25
	F_2	1,119	<*1*	109	*ns*
Sent. 1	F_1	1,79	<*1*	6	*ns*
	F_2	1,119	<*1*	8	*ns*
Predictability x Preview					
FFD	F_1	1,79	<*1*	279	*ns*
	F_2	1,119	<*1*	454	*ns*
SFD	F_1	1,79	<*1*	436	*ns*
	F_2	1,119	<*1*	594	*ns*
GD	F_1	1,79	<*1*	469	*ns*
	F_2	1,119	<*1*	747	*ns*
TT	F_1	1,79	<*1*	834	*ns*
	F_2	1,119	<*1*	1275	*ns*
% Skip	F_1	1,79	15.16	47	<.001
	F_2	1,119	12.13	88	<.001
Sent. 1	F_1	1,79	<*1*	4	*ns*
	F_2	1,119	<*1*	15	*ns*

Table 5.7 Summary of ANOVA Interaction Results in all Fixation Duration Measures and Skipping by Subjects (F_1) and by Items (F_2)

Note: FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration; TT = total fixation time; % Skip = percentage of times the target word was skipped; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

Table continued overleaf

Table 5.7 continued

Frequency x Predictability x Preview					
FFD	F_1	1,79	1.03	250	>.30
	F_2	1,119	*<1*	501	*ns*
SFD	F_1	1,79	*<1*	430	*ns*
	F_2	1,119	*<1*	526	*ns*
GD	F_1	1,79	1.46	459	>.20
	F_2	1,119	1.19	588	>.25
TT	F_1	1,79	*<1*	708	*ns*
	F_2	1,119	*<1*	1211	*ns*
% Skip	F_1	1,79	1.55	34	>.20
	F_2	1,119	*<1*	85	*ns*
Sent. 1	F_1	1,79	*<1*	4	*ns*
	F_2	1,119	*<1*	6	*ns*

Table 5.8 Follow-up Contrasts when the Frequency x Preview Interaction was Significant by Subjects (F_1) and by Items (F_2)

		HF Valid < HF Invalid		**LF Valid < LF Invalid**		**HF Valid vs. LF Valid**		**HF Invalid < LF Invalid**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
FFD	F_1	389.57	<.001	521.74	<.001	*<1*	*ns*	16.13	<.001
	F_2	305.98	<.001	375.8	<.001	1.84	>.15	10.55	<.01
SFD	F_1	526.4	<.001	734.24	<.001	1.31	>.25	28.04	<.001
	F_2	296.47	<.001	367.24	<.001	2.77	0.099	13.02	<.001
GD	F_1	608.1	<.001	885.73	<.001	2	>.15	42.43	<.001
	F_2	412.56	<.001	569.47	<.001	2.69	=.104	26.95	<.001
TT	F_1	628.1	<.001	876.71	<.001	4.98	<.05	45.96	<.001
	F_2	401.96	<.001	538.96	<.001	4.26	<.05	27.4	<.001

Note: HF = high frequency; LF = low frequency; FFD = first fixation duration; SFD = single fixation duration; GD; gaze duration

Table 5.9 Follow-up Contrasts when the Predictability x Preview Interaction was Significant by Subjects (F_1) and by Items (F_2)

		HP-Valid > HP Invalid		**LP-Valid > LP Invalid**		**HP Valid > LP Valid**		**HP Invalid vs. LP Invalid**	
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
% Skip	F_1	466.3	<.001	258.8	<.001	25.18	<.001	*<1*	*ns*
	F_2	372.95	<.001	206.98	<.001	20.15	<.001	*<1*	*ns*

Note: %Skip = percentage of times target was skipped, HP = high predictable; LP = low predictable

Table 5.10 Follow-up Contrasts when the Frequency x Predictability Interaction was Significant by Subjects (F_1) and by Items (F_2)

		HF-HP vs. HF-LP		LF-HP vs. LF-LP		HF-HP vs. LF-HP		HF-LP vs. LF-LP	
Measure		***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Sent. 1									
	F_1	4.31	<.05	9.21	<.01	9.39	<.01	4.19	<.05
	F_2	1.49	>.20	1.78	>.15	2.05	>.15	1.26	>.25

Note: HF = high frequency; LF = low frequency; HP = high predictable; LP = low predictable; Sent. 1 = sentence 1 first pass reading time (in milliseconds per character)

5.7 Discussion

The aim of the current experiment was to examine the effects of manipulating frequency, predictability and preview in a within-subjects design on first, single, gaze duration measures as well as total time and examine the amount of skipping of the target. The design of the experiment was such that four different groups of participants saw one of four versions of materials; a total of 80 participants took part. Targets were either high or low frequency and were embedded in the middle of the second line of a two-line passage of text. The predictability of the passage of text was determined via the Cloze task, details of which at the beginning of this Chapter in pre-test 3. Predictability was deemed to be high predictable or low predictable, given the context presented prior to the target. Two previous studies have examined the effects of frequency, predictability and preview on fixation duration and skipping measures (cross Experiments 2 and 3 comparison in Chapter 4, present thesis; Hand et al., 2010). Regarding the cross Experiment comparison, it was suggested that manipulating preview using a between-groups design was a useful start in the investigation of these variables but could be bettered by being manipulated so that the same participants took part in both valid and invalid preview conditions. The merits of this previous study were replicated in the current experiment. That is, we made sure that the nonword used in the invalid preview condition provided no useful information to the reader about the actual target. Nonwords were created so that they had the same overall word shape and were of equal character length as its corresponding target. Letter overlap between each position of the target and the nonword was avoided since we wanted to make sure that readers were prevented from acquiring lexical information from the parafoveal word. The present experiment was the first study to use parafoveal previews which tried to be less disruptive as possible compared to past studies which have examined frequency with preview (e.g., Kennison & Clifton, 1995) or predictability with preview (e.g., Balota et al., 1985) which have, for example used orthographically illegal letter strings (Kennison

& Clifton, 1995) or scrambled letters (Schroyens et al., 1999). Hand et al., (2010) also investigated the effects of frequency, predictability and preview. The effect of preview was analysed in post-hoc where launch distance to the target was used to index the amount of parafoveal processing.

In the present study, we analysed main effects of frequency, predictability and preview and expected significance in all fixation duration measures along with the skipping measure. Also expected was significance in the interactions of frequency x predictability, frequency x preview and predictability x preview as well as the three-way of frequency, predictability and preview in first and single fixations and skipping.

5.7.1 Main effects in first and single fixation durations

In the first and single fixations, results showed significant main effects of frequency, predictability and preview. This means that participants had shorter first fixations when reading high frequency versus low frequency targets. The significant main effect of predictability was that participants fixated high predictable targets for less time than low predictable ones as shown by shorter first and single fixations on the high predictable targets compared to the low predictable ones. The significant main effect of preview was that readers had shorter first and single fixations on targets in the valid preview condition than targets in the invalid preview condition.

These significant main effects are well documented in the eye tracking literature. For example, Hand et al. (2010) showed significant main effects of frequency, predictability and preview in first and single fixation duration measures. In particular, for the predictability variable, their Cloze manipulation was similar to the present study high predictable targets where had a mean Cloze probability of 0.66 (0.24) and 0.60 (0.26) and low predictable targets were 0.07 (0.12) and 0.05 (0.09) for high and low frequency targets respectively. In Hand et al. (2010) high predictability targets had a mean Cloze probability of 0.60 (0.31) and 0.53 (0.31) and 0.02 (0.06) and 0.02 (0.06) for high and low frequency targets respectively.

5.7.2 First fixation duration interactions

Regarding the two-way interactions, only frequency x preview was significant (in subjects analysis). Follow-ups showed that three out of contrasts were significant. Basically, there was an effect of preview for high and low frequency targets. That is, for high frequency targets (and low frequency targets) participants had shorter first fixations when the target appeared in a valid preview condition than in invalid preview. When looking at the effect of frequency, there was a difference between high and low frequency targets where high frequency targets has shorter first fixations than low frequency ones – but this was only for targets in the invalid preview condition. The other two interactions, frequency x predictability and predictability x preview as well as the three-way were nonsignificant in the first fixation measure.

Our earlier study (cross Experiments 2 and 3 comparison) which manipulated frequency, predictability and preview, did not find a significant effect of frequency x preview in first fixation duration for possible reasons of preview being manipulated in a weaker between-groups design. The current experiment appears to have somewhat successfully implemented the within-subjects manipulation of preview (having obtained significance in subjects analysis) and in doing so, is in line with several other studies which have shown a significant interaction of frequency x preview in first fixation duration (Inhoff & Rayner, 1986; Hand et al., 2010; Reingold et al., 2012). The interaction in these studies was that a preview of high frequency targets led to shorter first fixations than when participants were given a preview of low frequency targets. The present study showed this since in the invalid trials only where high frequency targets were read significantly faster than low frequency ones (see Table 5.13).

In terms of the preview effect (or parafoveal preview benefit), the typical finding is that high frequency targets have a larger preview benefit than low frequency targets (e.g., Inhoff & Rayner, 1986). This has led to the researchers stating that more information is extracted from a high frequency parafoveal word than from a low frequency one. In the present study, we calculated the size of the parafoveal preview benefit (see Table 5.11).

Table 5.11 FFD Frequency x Preview Interaction: Size of the Parafoveal Preview Benefit for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HF	207 ms	247 ms	40 ms
LF	209 ms	257 ms	48 ms

Note: HF = high frequency; LF = low frequency

As can be seen in Table 5.11, the size of the preview benefit was reversed compared to what has been traditionally reported in the literature – *low frequency* targets had a larger preview benefit than high frequency targets. Before discussing possible reasons for this, we thought it useful to examine results in the single fixation duration measure (see Section 5.7.3).

The size of the frequency effect is shown in Table 5.12. It should be noted that the frequency contrast in the valid preview condition was nonsignificant. When an examination of the size of the frequency effect was carried out, results were again reversed. That is, the frequency effect in the valid preview condition (2 ms) was smaller than that in the invalid condition (10 ms) (though it should be noted that the frequency contrast in the valid preview was nonsignificant).

Table 5.12 FFD Frequency x Preview Interaction: Size of the Frequency Effect for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	HF targets	**LF targets**	**Size of frequency effect**
Preview			
Valid preview condition *	207 ms	209 ms	2 ms
Invalid preview condition	247 ms	257 ms	10 ms

* The frequency contrast (HF vs. LF) in this condition was nonsignificant
Note: FFD = first fixation duration; HF = high frequency; LF = low frequency

An examination of past studies shows that the frequency effect is larger in the valid than in the invalid preview condition (see Table 5.13) in past studies which have manipulated frequency and preview. A comparison of first fixation durations in high versus low frequency showed that the frequency effect was larger in the valid preview condition than in the invalid preview condition (Hand et al. 2010: Valid = 50 ms; Invalid = 12 ms, Inhoff & Rayner, 1986: Valid = 18 ms; Invalid = 0ms, Reingold et al. 2012, Valid = 20 ms; Invalid = 9ms). The next section discusses results obtained in the single fixation duration measure before discussing why this might have been the case in the present study.

Table 5.13 Size of Frequency Effect in FFD in past studies which have examined Frequency and Preview

Past studies which have manipulated frequency and preview	Preview	Size of frequency effect in first fixations
Hand, Miellet, O'Donnell and Sereno (2010) *	Valid preview condition	50 ms
	Invalid preview condition	12 ms
Inhoff and Rayner (1986)	Valid preview condition	18 ms
	Invalid preview condition	0 ms
Reingold, Reichle, Glaholt & Sheridan (2012)	Valid preview condition	20 ms
	Invalid preview condition	9 ms

* Collapsed over the predictability factor

5.7.3 Single fixation duration interactions

In the single fixation duration measure, the only significant interaction was again just the frequency x preview one (and again, only in subjects analysis). Follow-ups to this interaction showed significance in the same sets of contrasts as in the first fixation duration measure. Thus, there was a significant effect of preview (valid vs. invalid) for both high and low frequency targets. Specifically, participants had shorter single fixations (read faster) when the target was in the valid preview condition than in the invalid preview condition. The effect of frequency (high vs. low) was significant for invalid preview targets only. That is, single fixations on high frequency targets were shorter than on low frequency ones only in the invalid preview condition. The other two interactions, frequency x predictability and predictability x preview as well as the three-way were nonsignificant in the single fixation measure.

In regards to the frequency x preview interaction, we first examined the size of the preview effect/parafoveal preview benefit (see Table 5.14) and second, the size of the frequency effect (see Table 5.15) in the present study.

Table 5.14 SFD Frequency x Preview Interaction: Size of the Parafoveal Preview Benefit for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HF	208 ms	257 ms	49 ms
LF	210 ms	268 ms	58 ms

Note: SFD = single fixation duration; HF = high frequency; LF = low frequency

Table 5.15 SFD Frequency x Preview Interaction: Size of the Frequency Effect for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	HF targets	**LF targets**	**Size of frequency effect**
Preview			
Valid preview condition *	208 ms	210 ms	2 ms
Invalid preview condition	257 ms	268 ms	11 ms

* The frequency contrast (HF vs. LF) in this condition was nonsignificant
Note: SFD = single fixation duration; HF = high frequency; LF = low frequency

Table 5.14 shows that the size of the parafoveal preview benefit was 49 ms for high frequency targets and 58 ms for low frequency ones. Therefore, as with the first fixation duration measure, we find that that the parafoveal preview benefit was in reverse to what has been reported in the literature. In the present study, we find that low frequency targets had a larger preview benefit than high frequency targets. Table 5.15 shows the results of the two frequency contrasts to the frequency x preview interaction. The frequency effect in valid preview condition was 2ms and 11 ms in the invalid preview condition – though because of the nonsignificance of the frequency contrast in the valid preview trials, these results are not firmly conclusive regarding a true comparison of the size of the frequency effect in valid versus invalid preview conditions.

Only three past investigations of frequency x preview have analysed results in the single fixation measure (cross Experiments 2 and 3 comparison, present thesis; Hand et al., 2010; Reingold et al., 2012). In our cross comparison of Experiments 2 and 3, the interaction was nonsignificant in the single fixation duration measure possibly due to between-groups manipulation of the preview factor. Therefore, we compared our parafoveal preview results with the two other studies (see Table 5.16).

Table 5.16 Hand et al. (2010) and Reingold et al. (2012) SFDs Frequency x Preview Interaction: Size of the Parafoveal Preview Benefit for High and Low Frequency Targets in Valid and Invalid Preview Conditions

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Hand, Miellet, O'Donnell and Sereno (2010)*			
HF	227 ms	278 ms	51 ms
LF	275 ms	293 ms	18 ms
Reingold, Reichle, Glaholt & Sheridan (2012)			
HF	216 ms	267 ms	51 ms
LF	239 ms	276 ms	37 ms

Note: HF = high frequency; LF = low frequency

As can be seen from Table 5.16, both previous studies which have shown a significant frequency x preview interaction in the single fixation measure have shown a parafoveal preview benefit where the size of the benefit is larger for high frequency targets than low frequency ones (Hand et al. HF = 51 ms; LF = 18 ms; Reingold et al. HF = 51 ms; LF = 37 ms). In the present study, we found reverse results where high frequency preview benefit was 49 ms and low frequency preview benefit was 58 ms.

The second comparison we carried out to the significant frequency x preview interaction was calculating the size of the frequency effect in single fixation durations in the valid and invalid preview conditions. The frequency effect sizes in Hand et al.'s (2010) and Reingold et al.'s (2012) studies are shown in Table 5.17.

Table 5.17 Size of Frequency Effect in SFD in past studies which have examined Frequency and Preview

Past studies which have manipulated frequency and preview	Preview	Size of frequency effect in single fixations
Hand, Miellet, O'Donnell and Sereno (2010) *	Valid preview condition	49 ms
	Invalid preview condition	15 ms
Reingold, Reichle, Glaholt & Sheridan (2012)	Valid preview condition	23 ms
	Invalid preview condition	9 ms

* Collapsed over the predictability variable

Table 5.17 shows that in the two past investigations of frequency x preview, both studies have shown a larger frequency effect in single fixations for targets viewed in a valid preview condition that in an invalid preview condition (Hand et al. 2010, Valid = 49 ms, Invalid = 15 ms, Reingold et al. 2012, Valid = 23 ms, Invalid = 9 ms).

The reason for reversed direction of results in the parafoveal preview benefit and frequency effects in first and single fixations could be to do with the fast reading times evident in the present study. The mean single fixation duration in the present study across the normal reading conditions (i.e., the valid preview trials) was 209 ms. If targets had been viewed in the more pixelated displays used in some past studies (e.g., Rayner et al. 2004; Hand et al., 2010), then it is highly likely that all the fixation durations in the present study would be a lot more slowed. For example, the mean single fixation duration in Rayner et al.'s (2004) was 271 ms and in Hand et al. (2010) was 277 ms. In the present study, a clear display could have led to targets in the high frequency-valid preview condition (already read the fastest given the preferential conditions of words being highly frequent to the reader as well being given a parafoveal preview of the target) hitting a floor, that is, it just was not possible for readers to fixate the targets any faster than they were already doing. It is possible then that if times had been slowed given a more pixelated presentation, we would have seen both parafoveal preview benefit and frequency effects in the expected direction.

This issue of floor effects (where reading times just cannot get faster because of the constraints of the visual system) was raised in past experiments in this thesis – it could be that clear display presentations leads to faster reading times. In fact, in Experiment 2 of this thesis, the mean single fixation duration was 205 ms. In another eye tracking study which used a similar display to that in our Experiment 2, Hand et al.'s (2012) study had an even faster mean single fixation duration of 192 ms. In Experiment 3 (Chapter 4) where we used the boundary technique, the mean single fixation was slowed to 246 ms and similarly, in the present experiment, the boundary trials had a slowed mean single fixation duration of 263 ms.

5.7.4 Gaze duration

5.7.4.1 Main effects

In the gaze duration measure, all three main effects were significant. All effects were in the desired direction. High frequency (and high predictable) targets had shorter (i.e., read faster) gaze durations than low frequency (and low predictable) ones; as well as targets in valid preview having shorter gaze durations than those in the invalid preview.

5.7.4.2 Interactions

The pattern of significance across the two-way and three-way interaction was the same as in the first and single fixation measures except for the frequency x preview interaction being significant in both subjects and items. Follow-ups showed the same three contrasts being significant as in the first and single fixations. Neither the frequency x predictability or the predictability x preview interactions was significant. In the three-way interaction, despite *F*-ratios being greater than one, subjects and items analyses were nonsignificant.

5.7.5 Total time

5.7.5.1 Main effects

All three main effects were significant. Participants spent less time on high than low frequency targets; on high than low predictable targets and on valid preview than on invalid preview targets.

5.7.5.2 Interactions

Interaction results were as follows: frequency x preview was significant in both subjects and items; predictability x preview and frequency x predictability were nonsignificant as was the three-way of these factors. Follow-ups to frequency x preview showed that all four contrasts were significant – the effects of preview (valid vs. invalid) was significant for both high and low frequency targets; the effects of frequency (high versus low) were significant for both valid and invalid preview conditions.

5.7.6 Percentage of times the target was skipped

5.7.6.1 Main effects

Regarding the skipping results, all main effects were significant as well as direction of effects occurring in the expected way. That is, participants skipped significantly more high frequency targets than low frequency ones; as well as high predictable ones more than low predictable ones and also significantly more targets when they appeared in the valid preview condition than in the invalid preview condition.

5.7.6.2 Interactions

Only predictability x preview was significant. Follow-up preview contrasts were both significant; we calculated the size of the parafoveal preview benefit (see Table 5.18). The size of the parafoveal preview benefit was larger for high predictable targets (16%) than for low predictable ones (13%). This result then is consistent with interactive approaches to lexical access. However, this cannot be firmly concluded because in terms of the predictability contrasts, there were no significant differences between high and low predictability in the invalid preview condition. The size of the predictability effect (see Table 5.19) was greater for valid than for invalid preview trials (though it should be noted that the predictability contrast in the invalid preview condition was nonsignificant).

Table 5.18 Percent Skipping Predictability x Preview Interaction: Size of the Parafoveal Preview Benefit Valid and Invalid Parafoveal Preview Conditions with High and Low Predictable Targets

	Valid preview	**Invalid preview**	**Size of parafoveal preview benefit**
Target			
HP	30%	14%	16%
LP	27%	14%	13%

Note: HP = high predictable; LP = low predictable

Table 5.19 Percent Skipping Predictability x Preview Interaction: Size of the Predictability Effect for High and Low Predictable Targets In Valid and Invalid Preview Conditions

	HP	**LP**	**Size of predictability effect**
Preview			
Valid	30%	27%	3%
Invalid*	14%	14%	0

* The predictability contrast (HP vs. LP) was nonsignificant

In terms of the nonsignificant predictability contrast (see row 4, Table 5.19) when readers saw targets with an invalid preview, both high and low predictable targets were skipped 14% of the time. In the earlier cross comparison of Experiments 2 and 3, the corresponding condition to the current HP-Invalid was the MP-Invalid, which incurred 12% skipping. The LP-Invalid condition in the cross comparison saw 14% skipping. In the cross comparison, then, results were numerically in an unexpected direction where *medium predictabl*e *were skipped less than low predictable targets* (this contrast was marginal in subjects and nonsignificant in items). It could be that in the present study, that low predictable targets should have been skipped far fewer times than they actually were. Perhaps the stimuli used in 'invalid' preview condition meant that readers processed this 'misinformation' and this then adds something to the reading process.

Additionally, it would not be expected that participants would skip a target when they had been given an invalid parafoveal preview of it. Given that invalid preview trials did show some skipping, it could be useful to analyse the skipping data in the invalid preview condition to see what participants were doing. It could be that readers intended to land on the target, but landed on the space or a few characters just after the target and kept on reading. In this case, they likely picked up the target given the small span of attention to the left. Another possibility is that the reader landed several characters past the target but kept on reading but they did come back to the target. The third possibility is the most intriguing where if readers landed several characters past the target, so did not pick it up from the left attention span, but they never came back to the target. The latter would be more likely in the high predictable contexts compared to low predictable because the highly predictable contexts are likely to be guiding our word recognition. Given that the nonwords were pronounceable, they would not have stood out in the text as being 'strange'. Given the percent skipping in the invalid preview conditions was relatively low (mean skipping was 14%) this kind of analysis would mean splitting an already small data set into even smaller sets.

5.7.7 Sentence 1 first pass reading time main effect and interaction results

Results analysed for Sentence 1 (computed as one region) showed expected results where the main effects of frequency, predictability and preview were all nonsignificant. Thus, there were no significant differences between high and low frequency targets; high and low predictable targets; valid and invalid preview conditions on line 1. This is an important finding because all the respective manipulations were on line 2 of the passage of text – there should be no effects of frequency, predictability or preview appearing on line 1.

Therefore, we would also expect the two-way and three-way interactions to be nonsignificant in order to reflect this lack of manipulation. Indeed, frequency x preview; predictability x preview and frequency x predictability x preview were all nonsignificant. Even though the frequency x predictability interaction was significant in subjects and marginally so in items, all four follow-up contrasts were nonsignificant in items analyses. Even though subjects follow-ups were significant, the difference in reading time was really small at around 40 ms. Therefore, at best, results are inconclusive about how much importance can be given to this significant frequency x predictability in subjects analysis.

5.7.8 Cloze values used in high and low predictable conditions

Table 5.20 shows that the mean Cloze completion for high predictable context 0.66 and 0.60 for high and low frequency targets; for low predictable contexts, this was 0.07 and 0.05 respectively. These probabilities are acceptable when it comes to the conventional literature on Cloze completion. That is, in Experiment 2 of this thesis (Chapter 3), it was seen that past studies of frequency and predictability have used Cloze completions (for high predictable) from 0.53 to 0.94 (see Table 3.8 in Chapter 3). For low predictable, Cloze completions have been from 0.01 to 0.06. Therefore, the present study's use of around 60% completion for high predictable and 5% for low predictable is congruent with this prior work.

Table 5.20 Mean Cloze Probability in High and Low Frequency Conditions across High and Low Predictable Contexts

Present study Mean Cloze probability		
HF	mean	SD
mean Cloze HP	0.66	0.24
mean Cloze LP	0.07	0.12
LF		
mean Cloze HP	0.60	0.26
mean Cloze LP	0.05	0.09

Note: HF = high frequency; LF = low frequency; HP = high predictable; LP = low predictable

5.7.9 Conclusions

The aim of the present study was to implement a large, well-controlled investigation of the single and combined effects of frequency, predictability and preview in a within-subjects design. The design of the experiment meant that four different groups of participants saw one of four versions of materials and a total of 80 participants took part. Previously Hand et al. (2010) investigated the single and combined effects of frequency, predictability and preview in an eye tracking reading study. The preview variable was indexed by launch distance. In the present study, we used the method of presenting targets in 'valid preview' and 'invalid preview' conditions.

Results in the present study showed similar effects in the first, single and gaze duration measures. First, all main effects were significant, direction of effects were all in the expected direction. All main effects in skipping were also in the expected direction and were all significant.

In first and single fixation durations, there was a significant frequency x preview interaction (in subjects only in both measures). Direction of numerical differences were as expected and significance was seen in three out of four follow-up contrasts, the exception being that there was no frequency effect for valid preview trials. Thus, when we calculated the size of the parafoveal preview benefit and the size of the frequency effect, in both cases, results showed there was a larger parafoveal preview benefit for low than high frequency targets; and a larger frequency effect in the invalid preview condition than in the valid preview condition. Since these results did not replicate the past studies which have shown that the preview benefit is greater for high than low frequency targets (Hand et al., 2010; Inhoff & Rayner, 1986; Reingold et al., 2012) and that the frequency effect is greater in the valid preview than in the invalid preview condition (Hand et al., 2010; Reingold et al., 2012) it was not possible to conclude that word frequency modulates parafoveal information where readers obtain more information from a high than a low frequency word (e.g., Inhoff & Rayner, 1986). Our results were explained by possible reasons of a floor effect in the high frequency-valid condition.

In the skipping data, predictability x preview was significant. Three out of four contrasts were significant: high predictable (and low predictable) targets in valid preview trials were skipped significantly more than high (and low) predictable in the invalid trials. For the predictability contrasts, high predictable targets were skipped significantly more than low predictability ones – but only in valid preview trials. When readers saw targets with an invalid preview, both high and low predictable targets were skipped 14% of the time. The size of the parafoveal preview benefit was greater for high than low predictable targets. The lack of a significant three way interaction means that it is not possible to reliably conclude that contextual predictability effects, that is, top-down sources arising from the sentence context, modulates the use of parafoveal information where the higher the contextual predictability, the more lexical access is facilitated. The implication of the first, single and skipping results in relation to the additive-interactive debate are not conclusive.

The merits of the study are that targets were controlled for a number of variables words differ on, such as phonemes, morphemes, orthographic neighbours et cetera between the conditions. Future directions from here would be to further manipulate the data. we could look at additional fixation duration measures such as spillover effects, launch distance, landing position and parafoveal-on foveal effects. While the ANOVA is useful in highlighting significant differences among different levels of different factors, other statistical techniques such as regression could be employed here. For example, for the predictability factor because the range of Cloze completions for high and low predictable targets varied this data could be further exploited in regression analyses. The range of Cloze completions for high predictable was 0.10-1.00 for both high and low frequency targets. The range of Cloze completions for low predictable (for high and low frequency) was 0.00-0.60 and 0.00-0.45. Appendix F shows, for every target word, the Cloze completion obtained for conducting this future study.

Chapter 6

General discussion and concluding remarks

The focus of the experiments in this thesis was to explore when semantic variables affect lexical access. Semantic effects describe meaning level influences (on lexical access). It was considered that such meaning level influences included both the semantic context within which a word occurs as well as the semantic attributes of the word itself. Hence, for experimental investigation purposes, the variables contextual predictability and word imageability, respectively, were used to explore this. Contextual predictability can be seen to reflect top down processing in the language processing system. The word frequency effect (the faster response to commonly used high frequency words compared to low frequency words which occur less often) was taken to index the moment of lexical access (Balota, 1990; Pollatsek & Rayner, 1990; Sereno & Rayner, 2003). The individual effects of word frequency and contextual predictability are well established in the eye movement literature. Word imageability has remained relatively less-researched in the eye movement field and the starting point was to conduct a lexical decision study. The approach of the experiments in this thesis was to use the logic of additive factors method (Sternberg, 1969a; 1969b) to determine whether the combined effect of each respective semantic variable was additive or interactive. By utilising a variety of methodologies, behavioural and eye movement, the aim was to elucidate the time course of processing in the system, in particular whether there are semantic influences on lexical access. A further aim of the thesis was to address the issue concerning the information presented to participants in the condition of ‘invalid parafoveal preview of a target’. Several criteria were identified as being important in order to make the assumption that parafoveal processing was successfully inhibited on the pre-target fixation. Another aim was to investigate the effect of word frequency of the parafoveal word on parafoveal preview benefit as well the effect of contextual predictability of the parafoveal word on the parafoveal preview benefit. The parafoveal preview benefit was calculated by subtracting fixation durations in a condition of ‘valid preview of the target’ from ‘invalid preview of the target’. It could be that word frequency affects parafoveal preview benefits (e.g., Inhoff & Rayner, 1986) and that contextual predictability affects parafoveal preview benefit (e.g., Balota et al., 1985). In such a case, lexical access can be viewed as an interactive type process (McClelland & Rumelhart, 1989; McClelland & O’Regan, 1981). Before discussing the wider implications of the results, a brief overview of the main findings from each experimental chapter will be provided.

6.1 Overview of the main experimental results

Experiment 1 (Chapter 2) reported a lexical decision study which examined the time course of word imageability. Critical words were manipulated on their frequency (LF, HF) and imageability (low imageability, LI; high imageability, HI). The pattern of reaction time results showed a main effect of frequency, a nonsignificant main effect of imageability and a nonsignificant interaction of frequency by imageability. The pattern of accuracy data (percentage errors made in each experimental condition) also showed both the significant main effect of frequency and the nonsignificant main effect of imageability; however there was a significant interaction of frequency x imageability. Follow-ups contrasts (in items) showed an unexpected direction of results where there was a processing advantage for low imageable words compared to high imageable (for low frequency).

The next two experiments (Chapters 3 and 4 respectively) manipulated the frequency (low frequency, LF; high frequency, HF) and contextual predictability (low, medium, high; LP, MP, HP respectively) of target words in sentences while participants' eye movements were recorded. Word frequencies were collected from established norms. Contextual predictability was operationally defined via Cloze probability (the proportion of subjects who guessed the target word when presented with the text up to but not including the target). A rating task was also used to check our manipulation of the contextual predictability variable. This study was the first to use targets which had been rated as very highly predictable in their contexts. Experiment 2 showed an interaction between frequency and contextual predictability of targets in single and first fixations. Our study showed very fast single fixation durations and also very high rates of skipping the target compared to comparable eye movement frequency-predictability past studies. The clearer display used in the present study compared to these past ones could have led to the faster reading times. As such, it is possible that there were floor effect operating in some experimental conditions, which if removed, would likely show additive effects of frequency and predictability.

Experiment 3 additionally used an invalid parafoveal preview of the target word using the boundary technique (Rayner, 1975) to investigate the contribution of parafoveal processing to the frequency-predictability interaction observed in Experiment 2. Thus, the normal parafoveal preview of the target word which is obtained from the prior fixation was substituted with a nonword. The nonword was constructed in order to fulfil several criteria: word length and overall word shape with its corresponding target was maintained as well as ensuring that the nonword could be pronounced given the rules of orthography in the English language. This Experiment was the first to meet all these criteria deemed necessary in order that the nonword did not stand out and command the reader's attention but at the same time ensure that parafoveal processing of a target was inhibited on the pre-target fixation. In this way, the parafoveal access of the target word (during fixations on the pre-target) was blocked and frequency-predictability effects could be evaluated, but based only on information obtained from foveal viewing of the target.

Results in Experiment 3 showed that when targets were read with an invalid parafoveal preview, an interaction between frequency and predictability in first and single fixations was no longer observed – results showed significant main effects of word frequency and of contextual predictability and a nonsignificant interaction. This would be expected in interactive positions according to which lexical access would be slowed when parafoveal information is not available since less information is being provided to the reader (e.g., McClelland & O'Regan, 1981). However, results in the skipping measure were unclear. That is, in the skipping results, there was a nonsignificant main effect of frequency, a significant main effect of contextual predictability and a significant interaction of the two variables (in subjects only). Follow-ups to the main effect and interaction showed many effects in unexpected directions. These results are difficult to interpret when readers have been given an invalid parafoveal preview of the target and they skip the target, it is unclear what this means. There are three possibilities. It could be that targets were skipped but readers did intend to land on them, that is, they could have landed on the space or few characters past the target but leftwards attention span meant that they did pick up the target. Also possible is that when readers landed just past the target, they kept on reading and only later, came back to the target. Another possibility is that readers landed several characters past the target, so did not pick up the target from the left attention span, yet never returned to the target. The mean percent skipping across invalid preview trials was only 12% so further dividing the data would lead to very small data sets.

In latter part of Chapter 4, a cross Experiments 2 and 3 comparison examined the parafoveal preview benefit when reading low and high frequency targets embedded in low, medium and high predictable contexts. Only one study has examined the joint influences of frequency and predictability and preview (Hand, Miellet, O'Donnell, & Sereno, 2010) and used targets of moderate predictability only. In addition, Hand et al., (2010) parafoveal processing was alternatively indexed using post-hoc analyses of launch distance to the eventual target. Therefore, the cross comparison of Experiments 2 and 3 presented further data analyses with Experiment 2's data considered the condition of 'valid preview of the target' and data from Experiment 3 as the condition of 'invalid parafoveal preview of the target' (a between-groups manipulation of preview). Predictions were based on pattern of results collected in Experiments 2 and 3.

Results showed a similar pattern of significant main effects of frequency and of contextual predictability in first and single fixation durations. In the first fixation duration, results showed significant interactions of word frequency x contextual predictability as well as of contextual predictability x preview. Word x frequency x preview was nonsignificant as was the three-way of the variables. For the word frequency x contextual predictability interaction, the numerical differences in the frequency contrasts and the predictability contrasts were in the expected direction. Low frequency targets incurred longer first fixations than high frequency ones; all contrasts were significant except at the high predictable level where only the subjects analysis was significant. For the predictability contrasts, all contrasts were significant except for low frequency targets, there were no significant differences between low and medium predictable targets and for high frequency targets, there were no significant differences between medium and high predictable targets. In terms of the contextual predictability x preview interaction, we examined the simple main effects to the significant main effect of predictability in valid and invalid preview conditions. It was concluded that this interaction had arisen because of the direction of results in the invalid preview condition where there were no significant differences between medium and low predictable targets – but the same contrast in the high predictable condition was significant. Also, in the valid preview condition, there were no significant differences between the medium and high predictable conditions – but in the invalid condition, this contrast was significant. When we calculated the size of the parafoveal preview benefit, results indicated a floor effect by the fast time observed in the high predictable-valid condition.

Results in single fixation duration were the same except that word frequency x contextual predictability was now nonsignificant and contextual predictability x preview was significant in subjects analysis but marginal in items. The three-way interaction was nonsignificant in all fixation duration measures. The skipping results showed significant main effects of word frequency and of preview and a nonsignificant main effect of contextual predictability. In the skipping data interactions, only contextual predictability x preview was significant; however a closer inspection of the data showed that this could likely be attributed to effects being cancelled out. Overall, it was concluded that this study had made useful progress towards investigating the single and combined effects of the three variables and a stronger manipulation of the preview variable would be to test the combined effects frequency, predictability and preview in a within-subjects design (see Experiment 4).

Experiment 4 (Chapter 5) reported an eye movement reading study which examined the effects of word frequency, contextual predictability and parafoveal preview in a within-subjects design. The design of the experiment meant that four different groups of participants saw one of four versions of materials and a total of 80 participants took part. Results showed that in the first, single and gaze durations, all main effects were in the expected direction and were significant in subjects and items analyses. In the first and single fixations, the frequency x preview interaction was significant (in subjects only). Follow-ups to this showed that preview contrasts (valid vs. invalid) for high and low frequency targets were significant and frequency contrasts (HF vs. LF) were significant for invalid targets only. It could be that the faster reading times for targets with a valid preview led to a floor effect – the single fixation time for high frequency valid targets was 205 ms and for low frequency valid targets, 206 ms. We calculated the size of the parafoveal preview benefit and the size of the frequency effect in first and single fixation duration. In contrast to past studies which have shown that preview benefit is greater for high than low frequency targets (Hand et al., 2010; Inhoff & Rayner, 1986; Reingold et al., 2012), this experiment showed that the preview benefit was larger for low than high frequency targets. Also in the opposite direction was the frequency effect: past studies have shown that the frequency effect is greater in the valid preview than in the invalid preview condition (Hand et al., 2010; Reingold et al., 2012); in this experiment the frequency effect was larger for targets in the invalid than in the valid condition. These reversed effects could be to do with the fast reading times seen in this study. For example, in the normal preview condition, the mean single fixation duration was 209 ms – more pixelated presentations (e.g., Hand et al., 2010; Rayner et al., 2004) have observed times of 277 ms and 271 ms respectively. In the present study, a clear display could have led to targets in the high frequency-valid preview condition (already read the fastest given the preferential conditions of words being highly frequent to the reader as well being given a parafoveal preview of the target) hitting a floor, that is, it just was not possible for readers to fixate the targets any faster than they were already doing. It is possible then that if times had been slowed, given a more pixelated, dot-matrix presentation, we would have seen both parafoveal preview benefit and frequency effects in the expected direction. Therefore, it was not possible to conclude that word frequency modulates the use of parafoveal information (e.g., Inhoff & Rayner, 1986; Hand et al., 2010; Reingold et al., 2012). The predictability x preview and frequency x predictability interactions as well as the three-way was nonsignificant in first and single fixation durations. This could be because the reading times are just so fast that effects are not appearing in measures of interest.

The percent skipping data in Experiment 4 showed that all three main effects were in the expected direction and were all significant in subjects and items. Only the predictability x preview interaction was significant. Follow-ups were all, except one, in the expected direction – and these latter ones were all significant. Basically, results showed that there was a greater parafoveal preview benefit for high predictable targets than low predictable ones. However, the predictability contrasts showed that there were no significant differences between high and low predictability in the invalid preview trials. Therefore, it cannot be firmly concluded that parafoveal preview modulates word skipping. Overall, in Experiment 4 it was concluded that results from the first, single and skipping data were inconclusive regarding the issue that processing in the language processing system may be interactive and the existence of floor effects in some conditions was discussed.

6.2 Models of word recognition and the impact of semantics

Models of word recognition differ in how they envisage the impact of semantics by the language processing system. These models are rooted in theoretical perspectives and we reviewed the various understandings of the language processing system, for example the modular (Fodor, 1983) and interactive (McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989) views in Chapter 1. It could be that semantic processes, specifically the semantic context within which a word is read, can guide the early lexical access process, for example by limiting the number of candidates which are activated. This indicates a role for top down processing in the language processing system which contravenes modular or bottom-up views of processing in the system. In terms of parafoveal processing, if readers are able to use information from a high frequency versus low frequency target presented in the parafovea as well as high versus low contextual predictability, then this suggests that lexical access is influenced by a number of factors, including contextual constraint and parafoveal information because the flow of information among the component processors is multidirectional (McClelland and O’Regan, 1981). The suggestion is that lexical access is as an interactive type process (McClelland & Rumelhart, 1989; McClelland & O’Regan, 1981) in which contextual and parafoveal information are used to drive lexical access.

The question of time course of semantic variables on lexical access addresses the theoretical question of whether processing in the language system is autonomous (e.g., Fodor, 1983; Forster, 1979) or interactive (McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1981). The potential architecture of the language processing system continues to be debated (Coltheart, 1999; Lucas, 1999; Masson & Kliegl, 2013; for serial and parallel processes in eye movement control, Murray, Fischer & Tatler, 2013 and for models of word recognition, Norris, 2013). Whether semantic variables operate during the lexical access processing stage or alternatively after lexical access, for example, in the post-lexical stage are one way of addressing this theoretical issue.

6.2.1 Bottom-up and top-down processing in the language processing system

Automatic word processing that proceeds in one direction from the sensory input to higher level processing is called bottom-up processing. Whenever this bottom-up processing is altered or influenced by experience, expectation or attention from higher level brain areas, our brain exerts what is called top-down control (Rayner & Pollatsek, 1989; Rayner, 2009). Top-down reading processes refers to the uptake of information as guided by the individual's prior external knowledge and experiences leading them to create certain expectations about what the text should contain. The reader uses their external knowledge to support and enrich their understanding of the text. This external knowledge takes the form of schemas in long-term memory. Schemas consist of all that an individual knows or associates with any given concept. We can identify two types of external knowledge. The first is world knowledge including any encyclopaedic knowledge and previous knowledge of, for example, the writer of the text. This aids the reader in constructing a *content schema* for the text. Secondly, the reader uses their previous experience if they have encountered similar text before in order to determine the kind of reading that the text needs. This is via a *formal schema* and aids the reader to recognise how information is likely to be set out and how they should engage with the writer.

The task of the skilled reader is to construct an on-line, developing internal (or meaning) representation of the text. This internal representation is a non-verbal construct which construes the reader's understanding of the text. This construct is constantly updated as the reader processes more information: incoming information has to be integrated with what has gone before, so as to ensure it adds to the developing representation of the text in a ways that are meaningful, relevant and consistent. This process requires the reader to be able to identify the main ideas and to be able to relate them to previous ideas. Higher level processing involves the use of information from sentential context to build a meaning representation from the text so far by bringing in existing knowledge experiences and expectations (in the form of pre-established schemas). As such, for skilled readers, contextual information is used to enrich the meaning representation of what has been read so far. When a reader is reading a given text to do with, for example, a man replacing slates having an accident, knowing that the text is about a man working on a roof, might cause the reader to create certain expectations at a rather conscious level i.e., ‘slates appear on the roof of a house’; ‘falling from a roof is likely to cause serious injury’. Thus the wider context (what the reader already knows) influences interpretation of a given word. When the reader comes to read the word ‘roof’ (i.e., a high predictable word given the context), the top down influence of world knowledge biases expectations about upcoming words. Typically, this would be that readers would spend less time on ‘roof’ compared to if they had read a context which was less predictable. One explanation of this finding is readers use their higher order thinking skills to predict upcoming words in a sentence.

There is a substantial amount of disagreement as to how semantic constraints influence the construction of meaning. There are several possibilities: autonomous; interactive; top-down driven; bottom-up priority. Table 6.1 below offers a summary of all the possibilities of the effect of contextual information. The aim of the experiments in this thesis were to discriminate between the first two viewpoints (autonomous/modular and interactive), since these are the most prominent in the current literature.

Table 6.1 Summary of Possible Views as to how Semantic Constraints Influence the Construction of Meaning

Possible views	**Explanation**
Autonomous /modular (bottom-up driven)	Bottom-up perceptual information is most important; semantic information is only used for checking and enriching it.
Interactive	Semantic information interacts with perceptual information at all stages of processing (both sources of evidence, top-down and bottom-up, are available throughout).
Top-down driven	Context biases interpretation of text before any perceptual information processing occurs.
Bottom-up priority	A minimal amount of perceptual data is processed before semantic influences can occur.

6.2.2. Distinctions in the information processing system

As with all other cognitive processes, psychologists have understood reading processes via the information processing approach to language. Indeed the cognitive approach is synonymous with the information processing approach (Neisser, 1967). This approach views perceptual and cognitive operations as taking place in stages or steps (serially; one stage or step at a time). However, there is accumulating evidence that has led to this approach being gradually replaced by models based on production models and connectionist theory, that is on two (or more) stages or steps being active simultaneously (i.e., in parallel).

There are two relevant distinctions in the literature on human information processing. The first is the serial versus parallel organisation of processing stages. In a serial stage model, information processing occurs sequentially, one stage at a time. On the other hand, in parallel models, more than one stage can be active simultaneously (i.e., in parallel). Many parallel models are also interactive, where each level/stage of processing with those above and below it to produce evidence which supports or rules out a particular interpretation (both sources of evidence, bottom-up and top-down are available throughout, i.e., all the information is available at one time).

The second of the distinctions drawn in the literature is the discrete versus continuous transmission of information through the information processing system. In discrete models, information accumulation takes one of two possible states – no information or full information. Conversely, continuous processing refers to cases where information is gradually accumulated. Whereas traditional models (Forster, 1979) posit a series of discrete processing stages, interactive models (McClelland & Rumelhart, 1981) point instead to cascaded processing in which even high level discourse information can affect lexical processing. In contrast to serial models, parallel models of processing hold the assumption that the mind can simultaneously process large amounts of information; more than one stage of processing may occur at any one time. The most influential version of parallel models use the human brain as the metaphor: a great deal of neural activity is occurring at any one time. These are parallel distributed processing (PDP) models (McClelland & Rumelhart, 1986). Interactive approaches to word recognition, that is, the brain metaphor, challenges these serial assumptions. In interactive models, processing of a written word is subject to simultaneous bottom-up and top-down influences.

6.2.3 Stages of lexical access and post-lexical integration

Traditional models of word recognition assume that a given word must be recognised before we can access the meaning of the word. Within such models then, word-level or lexical processing precedes processing at higher levels of analysis. This means that traditionally, word processing was typically seen to comprise two stages, the early lexical access stage followed by a later post-lexical stage. Lexical access has been defined as the moment we recognise a stimulus is a word and all the information associated with the word, such as its meaning, syntactic class, sound, and spelling are accessed immediately, hence the word is recognized (Balota, 1990). Post-lexical integration refers to the processes which take place when lexical access is complete. This includes integrating the meaning of the lexical representation with the prior context. Semantic or meaning level effects on early lexical access are controversial within traditional models.

In the eye movement dependent variables, first and single fixations may indicate the initial access to representation of a word's orthography, phonology or meaning. Gaze duration is likely to reflect processes that occur between initial lexical access and later processing activities – total time is likely to indicate post-lexical processing activities. Skipping is likely to represent early processing activities.

6.2.4 Experimental results and the additive-interactive debate

In Experiment 1, we looked at the effects of frequency and imageability on reaction time and percent error. According to a modular architecture, semantics affects the output of the lexical processor (Fodor, 1981; Forster, 1979). In experimental investigations, the prediction is that there should be main effects of frequency, imageability and a nonsignificant interaction of the two variables. In an interactive view, semantic constraints can directly affect lexical access seemingly through parallel processing stages (McClelland, 1987; Morton, 1969). In an experimental study, this would be shown by significant main effects of frequency and of imageability, as well as the significant interaction, indicating that both variables are influencing the same stage of processing. Actual results showed that in the reaction time measure, there was a significant main effect of frequency, a nonsignificant main effect of imageability and a nonsignificant interaction of the two variables. In the percent error, there was a significant main effect of frequency and a significant interaction of frequency and imageability. Thus, only the percent error results lend support to an interactive position. Because there was no main effect of imageability nor a significant interaction of frequency x imageability, the reaction time results were undecided regarding additive-interactive views of lexical processing while the percent error were did suggest support for an interactive view. The reasons for a lack of main imageability effect in the reaction time and percent error were discussed.

In Experiment 2, the frequency x predictability interaction in first (marginal in items) and single fixation duration measures could be taken to lend support for an interactive architecture of the language processing system. The interactive result obtained in first and single fixations suggests that lexical access (as indexed by word frequency) is modulated by top-down sentence context (contextual predictability). However, this experiment had very fast single fixation durations as well as a high rate of skipping compared to past studies which have also examined the frequency-predictability interaction using eye movements. Results in first and single fixations appear to lend support to a modular view, which would state that the interaction in first and single fixation durations exists because floor effects are masking an actual additive relationship between frequency and predictability. The modular view would also posit that there should be an interaction of frequency and predictability in later measures, that is, gaze duration and total time, and since present results showed an additive result in these measures, it was concluded that the frequency-predictability relationship would be further investigated in the next experiment.

In Experiment 3, we examined if the frequency-predictability interaction in Experiment 2 could have been due to the parafoveal processing aspect of reading high frequency and high contextual predictability targets. According to interactive views of lexical access (McClelland & Rumelhart, 1989) when readers are denied a parafoveal preview of a target then lexical access should be slowed because the reader needs to use foveal and parafoveal information sources to drive lexical access. It is possible then that an invalid parafoveal preview could lead to a pattern of results with significant main effects of frequency and predictability and a nonsignificant interaction in first and single fixation durations. In skipping, if the decision to skip a target depends on both foveal and parafoveal aspects of text (e.g., Rayner et al., 1982) then when a parafoveal preview is denied, we would expect low levels of skipping. This is because parafoveally, readers would likely process some aspects of the nonword – we did maintain overall word shape and used pronounceable characters so that this nonword would not command attention. Thus, when the parafoveal preview is denied, an interactive view of lexical access would suggest main effects of word frequency and of predictability and a nonsignificant interaction. Overall results in the first and single fixations did show the two significant main effects and a nonsignificant interaction. However, in the skipping measure, results showed a nonsignificant main effect of frequency, a significant main effect of contextual predictability and a significant interaction of the two variables (in subjects only). Follow-ups to both the contextual predictability main effect and the interaction showed many results in unexpected directions. Overall, it was concluded that first and single fixation duration measure results could be taken to support an interactive view of lexical access.

In the cross Experiments 2 and 3 comparison, the effects of word frequency, contextual predictability and preview validated the pattern of results observed in Experiments 2 and 3 in fixation durations measures and percent skipping. The three-way was likely nonsignificant because the preview factor was manipulated between groups. In regards to the additive-interactive debate despite no firm conclusions being drawn, this experiment made good progress towards understanding the effects of word frequency, contextual predictability and preview.

In Experiment 4, the single and combined effect of frequency, predictability and preview were manipulated. Results in the first and single fixation duration measures showed that frequency x preview interaction was significant (in subjects only). Since frequency contrasts were significant for invalid targets only, it was suggested that a floor effect could be operating in the valid condition. The relative size of the frequency effect across preview conditions as well as the relative size of the parafoveal preview benefit across frequency conditions were reverse to the usual findings and were likely hitting a floor. Therefore, it was not possible to conclude that word frequency modulates the use of parafoveal information (e.g., Inhoff & Rayner, 1986; Hand et al., 2010; Reingold et al., 2012). Predictability x preview was nonsignificant in first and single fixations but was significant in skipping. The latter result showed that there was a greater parafoveal preview benefit for high predictable targets than low predictable ones which would be consistent with interactive views of lexical access. Since the predictability contrasts showed that there were no significant differences between high and low predictability in the invalid preview trials we could not firmly conclude that parafoveal preview modulates word skipping. Overall, results from the first, single and skipping data were inconclusive regarding whether processing in the language processing system is interactive.

6.2.5 Future directions

In the eye tracking experiments in this thesis, the issue of a floor effect in the data was raised several times. In the current studies, eye movements were measured using an EyeLink® 2K eye tracker. Passages of text were presented in 14-point Bitstream Vera Sana Mono font which used black characters on a white background and is a non-proportional font. Past eye tracking studies of either frequency-predictability or frequency-predictability-preview used a Dual-Purkinje eye tracker and materials were presented in a dot-matrix font (cyan letters on a black background) usually in a dimly lit room (e.g., Hand, Miellet, Sereno & O'Donnell, 2010; Rayner, Ashby, Pollatsek & Reichle, 2004). These factors could make the text difficult to read because characters might appear pixelated or blurred. When comparing single fixation durations in for example Experiment 2 (see Table 3.17 in Chapter 3, p. 152) it was shown that the mean single fixation duration in Experiment 2 was much faster at 205 ms compared to 277 ms and 271 ms in Hand et al. (2010) and Rayner et al. (2004). Similarly, the skipping rate was an average of 25% in Experiment 2, compared to 19% and 18% in the other two studies respectively (see Table 3.18 in Chapter 3, p. 153).

If it is the case that a good quality presentation display led to faster fixation durations as well as more skipping (a clear font makes the accessibility of parafoveal information better so targets are more likely to be skipped) compared to presentations which have used point-plotting displays, then the question arises how to remove this floor. One solution to compensate for the clear display is that we introduce a font that is lower in contrast – thus trying to 'blur' the text so that it slows down reading. However, this could be problematic because introducing poor stimulus quality could interact with the variables of interest (for studies on word frequency and stimulus quality, see Bangert, Abrams & Balota, 2012; Scaltritti, Balota & Peressotti, 2013; Yap, Balota & Tan, 2013). Perhaps it is not feasible to circumvent the clear display presentations which exist in current eye tracking machines. Thus, another possibility is that the frequency-predictability and/or preview relationship should be investigated using a different technique altogether. Event-related potentials (ERPs) offer one avenue. ERPs offer continuous recording and can help elucidate the time course of frequency and contextual predictability.

6.2.5.1 Event-related potential studies of frequency x predictability

Some early ERP studies examined the combined effects of word frequency and contextual predictability to see whether the effect was additive or interactive (Van Petten & Kutas, 1990; 1991, Van Petten, 1993). Van Petten and Kutas (1990) examined word frequency, predictability and sentence position of the target word. They measured the N400 amplitude in a sentence reading task. The materials were 338 single line sentences with an average of 9.3 words per sentence. Targets were open-class (content) words which were nouns, verbs, adjectives, and adverbs ending with 'ly'. These words appeared in a highly predictable context (mean Cloze was 0.83, SD=0.17) and were sentence final. In addition, all the words in the sentence were sorted into three groups according to the position they appeared in the sentence as an indication of the degree of contextual constraint. Words were then sorted into a further six groups according to word frequency.

Results in Van Petten and Kutas’ (1990) study showed that low frequency words elicited a larger N400 amplitude than did the high frequency words. The results also showed that low frequency words which occurred near the beginning of the sentence elicited larger N400s than did high frequency words. However this frequency effect did not exist for words which occurred near the end of the sentence. This suggests that the more context the participant had prior to a target, the less the effect of frequency. The authors concluded that this interaction between sentence context and word frequency is at an early stage in word recognition specifically in which word frequency can be replaced by contextual constraint provided by a sentence during the early stages of word recognition. However, this argument rests on the view that the N400 reflects early lexical processes.

More recently, Dambacher, Kliegl, Hofmann and Jacobs (2006) examined the combined effect of frequency and predictability (along with word position) using a multiple linear regression. They examined the effects of these variables in two time windows: an early time window, the P200 component, ranging from 140-200 ms post-stimulus and secondly, in a late time window, the N400, ranging from 200 and 500 ms post-stimulus. Dambacher et al., (2006) reported a significant main effect of word frequency on the P200, peaking at 170 ms after post-stimulus. There was also a significant main effect of predictability on the P200. However, there was no significant interaction between the two variables on this early component. Results from the regression analyses showed that low frequency words elicited larger P200 amplitudes than high frequency words. Predictability and the interaction of frequency and predictability were significant. In terms of the N400 component, frequency was marginally significant, there was a main effect of predictability and there was a significant interaction between frequency and predictability. Results from the regression analyses showed that the strongest predictor for the N400 was predictability. The interaction of frequency and predictability was significant. Low frequency words elicited a larger effect than did high frequency words.

Penolazzi, Hauk and Pulvermüller (2007) examined the effects of frequency (high vs. low) and predictability (high vs. low cloze probability) and word length (short vs. long) using an orthogonal design in a sentence reading task. All target words were nouns and appeared in a single line sentence and always appeared in sixth word position of the sentence, that is, targets were not sentence finals. Target words were preceded by a highly predictable context. By changing one word in this preceding text, the sentence context was made low predictable. A participant saw one version of the materials. Predictability was determined by a Cloze rating task in which a separate group of participants were presented with sentence fragments which consisted of the sentence up to and not including the target word. Participants had to write in their three most probable words which they thought fitted in with the context of the preceding text. A sentence was considered to have a high Cloze probability if the target word had been written by at least 50% of the participants but also had not been written in the low predictability sentence fragment.

Penolazzi et al., (2007) examined five time intervals: two early intervals before 200 ms and three later time intervals. They reported main effects of frequency at around 120 ms and predictability appeared 180 ms after stimulus onset, however, both these variables interacted with word length. Predictability interacted with word length up to around 300 ms. Predictability also interacted with topographical variables. Similarly, in the second time window (170-190 ms), there were main effects of frequency but this interacted with word length. There was also a main effect of predictability and this interacted with topographical variables. In a time window corresponding to the N400, they reported additive effects: main effects of word frequency and predictability on the N400 component. Thus the results from this experiment did not show early frequency effects or early predictability effects independently of word length.

Sereno et al., (2003) investigated the issue of the locus of sentence context by examining the effect of frequency, predictability and ambiguity in a sentence reading task. The authors investigated effects in an early time window from 132-192 ms after stimulus onset which corresponded to the N1 component. Participants were presented with a single line sentence. Target words were nouns and were the sentence final. Targets were either ambiguous words, or low frequency words or high frequency words and were presented in both neutral (low predictable context) and in highly predictable contexts. Target words appeared twice: in neutral and in highly predictable contexts and participants read one version so that they did not read a given target more than once. There were three main results on the early N1 component. First, Sereno et al. replicated previous findings by showing a frequency effect in this early component. That is, HF words were read faster than LF words and this was in both low and high predictable contexts. Second, context effects were also found in the N1, however only for LF words appearing in the highly predictable context. Moreover, this effect was marginally significant. The third main result was to do with responses to ambiguous words. Overall, the conclusion was that contextual predictability has an early locus at the same time as word frequency and thus there was strong support for interactive accounts of language processing. The implication is that higher level conceptual processes, as well as lower level perceptual (bottom-up), influence early lexical processing.

One future study could address the design issue of the majority of ERP studies which have typically presented sentence word-by-word at a rate of around two words per second or around 500 ms per word. This presentation rate of a target is relatively slow compared to normal reading, particularly from the eye movement literature which has shown that the average fixation duration is 250 ms for normal reading (Sereno, Rayner & Posner, 1998). Materials used in the eye tracking Experiment 5 would be augmented by carrying out an EEG study with them. That is, we could divide 240 passages of text so that there are 60 items per condition (collapsed across preview since preview is not a factor in EEG studies). The interest would be in examining the pattern of effects in the early components since evidence exists to suggest that component such as N1 and P2 are sensitive to early lexical processes (e.g., Scott et al., 2010; Sereno et al., 2003). Results should shed light on the temporal contingencies of word recognition processes and help to inform theories of word recognition.

List of references

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behaviour Research Methods, Instruments, and Computers, 31*, 531-552.

Altarriba, J., Bauer, L.M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers,* 31, 578-602

Altarriba, J., Kambe, G., Pollatsek, A., and Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations: Evidence from fluent Spanish-English bilinguals. *Perception and Psychophysics, 63*, 875-891.

Altaribba, J., Kroll, J.F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed language sentences: Evidence from eye fixation and naming times. *Memory & Cognition, 24*, 477-492.

Ashby, J., Rayner, K., & Clifton, C., Jr. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology, 58A*, 1065-1086.

Balota, D.A. (1990). The role of meaning in word recognition. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading (pp. 143-164)*. Hillsdale, NJ: Erlbaum.

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

Balota, D.A., Cortest, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap., M.J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283-316.

Balota, D. A., Ferraro, R. F., & Connor, L. T. (1991). On the early influence of meaning in word recognition:Areview of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings (pp. 187–222)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*, 364-390.

Balota, D.A. & Rayner, K. (1983). Parafoveal visual information and semantic contextual constraints. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 726–738.

Balota, D.A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables. In D.Besner & G. W. Humphreys (Ed.'s), *Basic processes in reading visual word recognition (pp. 198-232).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bangert, A.S., Abrams, R.A., & Balota, D.A. (2012). Reaching for words and nonwords: Interactive effect of word frequency and stimulus quality on the characteristics of reaching movements. . *Psychonomic Bulletin & Review, 19*, 513-520.

Barca, L., Burani, C., & Arduino, L. S. (2002).Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, and Computers, 34*, 424-434.

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

Becker, C.A (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8*, 493-512.

Binder, K. S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1162-1172.

Binder, J.R., Westbury, C.F., McKiernan, K.A., Possing, E.T. and Medler, D.A. (2005). Distinct brain systems for processing abstract and concrete words. *Journal of Cognitive Neuroscience, 17,* 905-917.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments and Computers, 33*, 73–79.

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 582–594.

Bloom, P.A. & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition, 8*, 631-642.

Boles, D. B. (1983). Dissociated imageability, concreteness, and familiarity in lateralized word recognition. Memory & Cognition, 15, 511–519.

Bransford, J.D., & Johnson, J.D. (1972). Contextual prerequisites for understanding: some investigations of comprehension an recall. *Journal of Verbal Learning & Verbal Behavior, 11*, 717-726.

Braun, M., Hutzler, F., Ziegler, C., Dambacher, M., & Jacobs, A.M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping, 30,* 1977-1989.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125–147). Oxford: Elsevier.

Carroll, P., & Slowiaczek, M.L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14*, 509-522.

Carroll, J.B., & White, M.N. (1973). Word frequency and age-of-acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*, 85-95.

Chiarello, C., Senehi, J., & Nuding, S. (1987). Semantic priming with abstract and concrete words: Differential asymmetry may be postlexical. *Brain and Language, 31,* 43–60.

Clark, H. (1973). The language-as-fixed-effect: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior*, *12*, 335-339

Clark, J.M. & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers, 36*, 371-383.

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257-271.

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

Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences, 3,* 115-120.

Coltheart, M. (2004). Are there lexicons? *Quarterly Journal of Experimental Psychology, 57A,* 1153-1171.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dualroute cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Cortese, M.J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers, 36*, 384-387.

Dambacher, M., Kliegl, R., Hofmann, M. and Jacobs, A.M. (2006). Frequency and predictability effect on event-related potentials during reading. *Brain Research,* 89-103.

Day, J. (1977). Right-hemisphere language processing in normal right-handers. *Journal of Experimental Psychology: Human Perception & Performance, 3*, 518-528.

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

Donders, F. C. (1868) Twee werktuigen, tot bepaling van den tijd, voor psychische processen benoodigd' appeared in: '*Onderzoekingen gedaan in het Physiologisch Laboratorium Utrecht'rs 1868'*, *2*, 21–25. Translated in: Donders, F.C. (1969). On the speed of mental processes. In W. G. Koster (Ed.), *Attention and Performance II. Acta Psychologica, 30*, 412-431.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: on the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16*, 79-103.

Ehrlich, S.F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641-655.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*, 621–636.

Engbert, R., Nuthmann, A., Richter, E.M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777-813.

Feldman, L. B., Basnight-Brown, D. & Pastizzo, M. J., (2006). Semantic influences on morphological facilitation: Concreteness and family size. *The Mental Lexicon*, *1*, 59–84.

Ferlazzo F, Conte S, Gentilomo A. (1993). Event-related potential and recognition memory within the 'levels of processing' framework. *Neuroreport, 4*, 667-670.

Fiebach, C.J. & Friederici, A.D. (2003). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia 42*, 62–70.

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

Francis and Kučera (1967). *Computational analysis of present- day American English*. Providence, RI: Brown University Press.

Fodor, J.A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. W. Walker (Eds.), *New approaches to language mechanisms (pp. 257–287)*. Amsterdam: North Holland.

Forster, K.I. (1979). Levels of processing and the structure of the language processor. In W.E. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett (pp. 27-85)*. Hillsdale, NJ: Erlbaum.

Forster, K.I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465-495.

Forster, K.I. (1989). Basic issues in lexical processing. In W.Marslen-Wilson (Ed.), *Lexical Representation and Process* (pp. 75-107). Cambridge, Ma.: M.I.T. Press.

Forster, K.I. & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The *turple* effect. *Memory & Cognition*, *30*, 1106-1116.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 word. *Behavior Research Methods & Instrumentation, 12*, 395-427.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 18*, 381-398.

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518-565.

Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114*, 357–375.

Hand, C.J., Miellet, S., O'Donnell, P., & Sereno, S.C. (2010). Word frequency and contextual predictability effects in reading: It depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1294-1313.

Hand, C.J., O'Donnell, P.J. & Sereno, S.C. (2012). Word-initial letters influence fixation durations during fluent reading. *Frontiers in Psychology, 3 (85)*, 1-19.

Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F. & Marslen-Wilson, W.D. (2006). The time course of visual word recognition as revealed as revealed by linear regression analysis of ERP data. *Neuroimage, 30*, 1383-1400.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller, F. & Rogers, T. (2006). When would you prefer a SOSSAGE to a SAUSAGE? At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience, 18*, 818-832.

Hauk, O. & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115*, 1090-1103.

Henderson, J.M., Dixon, P., Peterson, A., Twilley, L.C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 82-97.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit during eye movements in reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 1150-1170.

Holcomb, P.J., Kounios, J., Anderson, J.E., & West, W. (1999). Dual-Coding, Context-Availability, and Concreteness Effects in Sentence Comprehension: An Electrophysiological Investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*,721-742.

Howell, J. R., & Bryden, M. P. (1987). The effects of word orientation and imageability on visual half-field presentations with a lexical decision task. *Neuropsychologia, 25*, 527–538.

Holmes, V. M., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behavior, 15*, 559–566.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology*, *34*, 293-304.

Hyönä, J. & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology, 46*, 475–483.

Inhoff, A.W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior*, *23*, 612-624.

Inhoff, A. W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 544-555.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–54). Oxford: Elsevier Inc.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431-439.

Izura, C., Pérez, M.A., Agallou, E., Wright, V.C., Marin, J., Stadthagen-González, H., & Ellis, A.W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory & Language, 64*, 32-58.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *104*, 130–136.

Johnston, R.A. & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition*, *13*, 789-845.

Juhasz, B.J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*, 684–712.

Juhasz, B.J. & Pollatsek, A. (2011). Lexical influences on eye movements in reading. In S. Liversedge, I. Gilchrist, and S. Everling (Eds.)., *The Oxford Handbook of Eye Movements (pp. 873-894).* New York: Oxford University Press.

Juhasz, B.J. & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1312-1318.

Juhasz, B.J. & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations in reading. *Visual Cognition, 13*, 846-863.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Kennedy, A. (2000). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology, 53A*, 429-455.

Kennedy, A., Murray, W.S., & Boissiere, C. (2004). Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology, 16*, 128-153.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kennedy, A., Pynte, J., Murray, W.S., Paul, S-A. (2012). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. *The Quarterly Journal of Experimental Psychology, 66*, 601-618.

Kennison, S.M. & Clifton, C. (1995). Determinants of Parafoveal Preview Benefit in High and Low Working Memory Capacity Readers: Implications for Eye Movement Control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 68-81.

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J. Hyönä, Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 103-117). Oxford: Elsevier.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 262-284.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*, 12-35.

Kliegl, R., Olson, R.K., & Davidson, B.J. (1982). Regression analyses as a tool for studying reading processes: Comments on Just and Carpenter's fixation theory. *Memory and Cognition*, *10*, 287-296.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting Dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 804–823.

Kousta, S.T., Vigliocco, G., Vinson, D.P., Andrews, M. and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, *140(1)*, 14-34.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 92–107.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial landing sites in words*. Acta Psychologica, 104*, 191-214.

Lee, C. & Federmeier, K.D. (2008). To watch, to see, and to differ: An event-related potential study of concreteness effects as a function of word class and lexical ambiguity. *Brain and Language 104*, 145–158.

Lima, S.D. & Inhoff, A. (1985). Lexical access during eye fixations in reading: effects of word-initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 272–285.

Liversedge, S.P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, 4, 6-14.

Lucas, M. (1999). Context effects in lexical access: A meta-analysis. *Memory and Cognition, 27,* 375-398.

Marschark, M., & Surian, L. (1992). Concreteness effects in free recall: The roles of imaginal and relational processing. *Memory & Cognition, 20,* 612-620.

Masson, M.E.J., & Kliegl, R. (2013). Modulation of additive and interactive effects in lexical decision by trial history. *. Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 898-914.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 8*6, 287-330.

McClelland, J.L. (1987). The case for interactionism in language processing. In M.Coltheart (Ed.), *Attention and performance: Vol.12. The psychology of reading (pp. 363-383).* Hillsdale, NJ: Erlbaum.

McClelland, J.L. & O'Regan, J.K. (1981). Expectations increase the benefit derived from parafoveal visual information in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 634-644.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–407.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. Cambridge, MA: MIT Press.

McConkie, G. W., Hogaboam, T. W., Wolverton, G. S., Zola, D., & Lucas, P. A. (1979). Toward the use of eye movements in the study of language processing. *Discourse Processes, 2,* 157-177.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophy ics, 17*, 578-586.

McConkie, G.W. Wolverton, G.S., & Zola, D. (1984). Instrumentation consideration in research involving eye-movement contingent stimulus control. In A.G. Gale and F. Johnson (Eds.), *Theoretical and Applied Aspects of Eye Movement Research (pp.* 39-47). Amsterdam: Elsevier.

McDonald, S.A., & Shillcock, R.C. (2003a). Eye movements reveal the on-line computations of lexical probabilities during reading. *Psychological Science, 14*, 648-652.

McDonald, S.A., & Shillcock, R.C. (2003b). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, *43*, 1735-1751.

Miellet, S., O'Donnell, P.J., & Sereno, S.C. (2009). Parafoveal magnification visual acuity does not modulate the perceptual span in reading. *Psychological Science*, *20*, 721-728.

Miellet, S., Sparrow, L., & Sereno, S.C. (2007). Word frequency and predictability effects in reading French: An evaluation of the E-Z Reader model. *Psychonomic Bulletin & Review*, *14*, 762-769.

Monaghan, J., & Ellis, A.W. What exactly interacts with spelling-sound consistency in word naming? *Journal of experimental psychology: Learning, Memory & Cognition*, *28*, 183-206.

Morris, R.K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 92-103.

Morrison, R.E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance, 10*, 667-682.

Morrison, C. M., Chappell, T.D., & Ellis, A.W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A*, 528-559.

Morrison, R.E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Perception & Psychophysics, 30,* 395–396.

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

Morton, J. (1979). Facilitation in word recognition: experiments causing change in the logogenmodel. In: Kolers, P.A., Wrolstad, M.E. & Bouma, H. (Eds.), *Processing of Visible Language (pp.* 259-268). New York: Plenum.

Murray, W.S., Fischer, M.H., Tatler, M.W. (2013). Serial and parallel processes in eye movement control: Current controversies and future directions. *The Quarterly Journal of Experimental Psychology, 66*, 417-428.

Murray, W.S. & Forster, K.I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review*, 111, 721-756.

Neisser, U. (1967). *Cognitive psychology*. Appleton-Century-Crofts: New York.

Nelson, D. L., & Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language, 31*, 237-260.

Nittono, H.,Suehiro, M., & Hori, T., (2002). Word imageability and N400 in an incidental memory paradigm. *International Journal of Psychophysiology, 44*, 219–229.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition, 22*, 93-136.

Norris, D. (2013). Models of visual word recognition. *Trends in Cognitive Sciences*, *17*, 517- 524.

O'Regan, J. K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25*, 501-509.

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp 395-453). (Vol. 4 of Reviews of Oculomotor Research). Elsevier, Amsterdam.

O'Regan, J. K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In: K. Rayner (Ed.) *Eye movements and visual cognition: Scene perception and reading (pp.333-354).* Springer Verlag, New York.

O'Regan, J.K., Lévy-Schoen, A., & Jacobs, A. (1983). The effect of visibility on eye movement parameters in reading. *Perception and Psychophysics, 34*, 457-464.

Paap, K.R., Newsome, S.I., McDonald, J.E., & Schvaneveldt, R.W. (1982). An interactive-verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89,* 573-594.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

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

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum.

Paivio, A., & O'Neill, B.J. (1970). Visual recognition thresholds and dimensions of word meaning. *Perception and Psychophysics, 8*, 273-275.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1196–1204.

Paivio, A., Yuille, J.C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*, 1-25.

Pachella, R.G. (1974). The interpretation of reaction time in information-processing research. In B.H. Kantowitz (Ed.). *Human Information Processing: Tutorials in performance and cognition* (pp. 41-82). Hillsdale, NJ: Lawrence Erlbaum.

Paller, K.A., Kutas, M., Shimamura, A., & Squire, L.R. (1987). Brain responses to concrete and abstract words reflect processes that correlate with later performance on a test of stem-completion priming. *Electroencephalography and Clinical Neurophysiology (Supplement 40),* 360-365.

Papagno, C., Fogliata, A., Catricala, E., & Miniussi, C. (2009). The lexical processing of abstract and concrete nouns. *Brain Research, 1263*, 78-86.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological Psychology, 74*, 374-388.

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

Pollatsek, A., Lesch, M., Morris, R.K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 148-162.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading (pp. 143-163)*. Hillsdale, NJ: Erlbaum.

Pollatsek, A., Rayner, K, & Balota, D.A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics, 40*, 123-130.

Posner, M.I. & McCandliss, B.D. (1993). Converging methods for investigating lexical access. *Psychological Science, 4*, 305-309.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (eds.), *Reading as a perceptual process (pp. 165-191).* Amsterdam: Elsevier Science.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology, 66*, 429-452.

Ransdell, S.E. & Fischler, I..(1987). Memory in a monolingual mode: When are bilinguals at a dis-advantage? *Journal of Memory and Language, 26*, 392-405.

Ransdell, S.E. & Fischler, I. (1989). Effects of Concreteness and Task Context on Recall of Prose among Bilingual and Monolingual Speakers. *Journal of Memory & Language, 28*, 278-291.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65-81.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory and Cognition, 4*, 443-448.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*, 21-30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457-1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E.D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 720-730.

Rayner, K. & Bertera, J.H. (1979). *Reading without a fovea. Science*, *206*, 468-469.

Rayner, K. & Balota, D.A. (1989). Parafoveal preview effects and lexical access during eye fixations in reading. In W. Marslen-Wilson (Ed.), *Lexical representation and processes (pp. 261-290).* Cambridge, Mass.: MIT Press.

Rayner, K., Balota, D.A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40*, 473-483.

Rayner, K., Binder, K.S, Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little effect on initial landing positions in words. *Vision Research*, *41*, 943-954.

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191-201.

Rayner, K., & Fischer, M.H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysics, 58*, 734-747.

Rayner, K., Fischer, M.H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*, 1129-1144.

Rayner, K. & Juhasz, B.J. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology, 16*, 340-352.

Rayner, K. & Juhasz, B.J. & Pollatsek, A. (2007). Eye movements during reading. In M.J. Snowling and C. Hulme (Ed.'s), *The Science of Reading: A Handbook (pp. 79-98).* Wiley-Blackwell.

Rayner, K. & Pollatsek, A. (1989). *The psychology of reading.* Englewood Cliffs, NJ: Prentice-Hall.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31*, 537-550.

Rayner, K., & Pollatsek, A. (2006). Eye movement control in reading. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of Psycholinguistics (pp. 613-658).* London: Elsevier.

Rayner, K., Pollatsek, A., & Reichle, E.D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences, 26*, 507-518.

Rayner, K, & Raney, G.E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3*, 238-244.

Rayner, K., & Sereno, S.C. (1994). Eye movements and reading: Psycholiniguistic studies. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics (pp. 57–81).* New York: Academic Press.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1188-1200.

Rayner, K., Sereno, S.C., Morris, R.K., Schmauder, A.R. & Clifton, C., Jr. (1989). Eye movements and on-line comprehension processes. *Language and Cognitive Processes, 4*, SI 21-50.

Rayner, K., & Well, A.D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*, 504-509.

Reichle, E.D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445-476.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125-157.

Reilly, J., Chrysikou., E. G., & Ramey, C. H. (2007). A hybrid semantic-phonological model of the age of acquisition of English nouns. *Psychonomic Bulletin & Review, 14*, 1164-1170.

Reilly, R. G., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J. Hyönä, Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements* (pp. 429–455). Amsterdam: Elsevier.

Reingold, E.M., Reichle,E.D., Glaholt, M.G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65*, 177-206.

Richards, L. G. (1976). Concreteness as a variable in word recognition. *American Journal of Psychology, 89*, 707-718.

Richardson, J. T. E. (1975). Imagery, concreteness, and lexical complexity. *Quarterly Journal of Experimental Psychology, 27*, 211–223.

Richardson, J.T.E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology, 67*, 315-325.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*, 487–494.

Rubenstein, H., Lewis, S.S., & Rubenstein, M. (1971). Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. *Journal of Verbal Learning and Verbal Behavior, 10*, 57-62.

Rubin, D. C. (1980). Fifty-one properties of 125 words: A unit analy- sis of verbal behavior. *Journal of Verbal Learning & Verbal Behav-ior, 19*, 736-755.

Rumelhart, D. E., & McClelland, J. L. (1986). PDP models and general issues in cognitive science. In D. E. Rumelhart, J. L. McClelland, and the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. *Volume 1: Foundations*.Cambridge, MA: Bradford Books/MIT Press.

Samson, D., & Pillon, A. (2004). Orthographic neighborhood and concreteness effects in the lexical decision task. *Brain and Language, 91*, 252–264.

Scaltritti, M., Balota, D.A. & Peressotti, F. (2013). Exploring the additive effects of stimulus quality and word frequency: The influence of local and list-wide prime relatedness. *The Quarterly Journal of Experimental Psychology, 66*, 91-107.

Schilling, H.E.H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*, 1270-1281.

Schroyens, W., Vitu, F., Brysbaert, M. & Y'dewalle, G. (1999). Eye Movement Control During Reading: Foveal Load and Parafoveal Processing. *The Quarterly Journal of Experimental Psychology, 52A*, 1021-1046.

Schuberth, R.E., & Eimas, P.D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 27-36.

Schustack, M., Ehrlich, S., & Rayner, K. (1987). The complexity of contextual facilitation in reading: Local and global influences. *Journal of Memory and Language, 26*, 322-340.

Schwanenflugel, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instruments, & Computers, 18*, 363-371.

Schwanenflugel, P. J. (1991). Why are abstract words hard to under- stand?. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings (pp.223-250).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Schwanenflugel, P. J., Akin, C., & Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition, 20*, 96–104.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory & Language, 27*, 499-520.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 82–102.

Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly, 24*, 114–126.

Scott, G.G., O'Donnell, P.J., Leuthold, H., & Sereno, S.C. (2010). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology 80*, 95–104.

Seidenberg, M. S. (1985). Constraining models of word recognition. *Cognition, 19*, 1-30.

Seidenberg, M.S. & McClelland, J.L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Sereno, S.C. (1992). Fast priming during eye fixations in reading. *Journal of experimental psychology: Human perception and performance, 18,* 173-84.

Sereno, S.C. (1995). Resolution of lexical ambiguity: evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 582-595.

Sereno, S.C., Brewer, C.C., & O'Donnell, P.J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science, 14*, 328-333.

Sereno, S.C., O'Donnell, P.J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate bias effect. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 335-350.

Sereno, S.C., Pacht, J.M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science, 3*, 296-300.

Sereno, S. C. & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 173-184.

Sereno, S.C., & Rayner, K. (2000). Spelling-sound regularity effects on eye fixations in reading. *Perception & Psychophysics, 62*, 402-409.

Sereno, S.C. & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*, 489-493.

Sereno, S.C., Rayner, K., & Posner, M.I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *NeuroReport, 9*, 2195-2200.

Slattery, T.J., Pollatsek, A., & Rayner, K. (2007). The effect of the frequencies of three consecutive content words on eye movements during reading. *Memory & Cognition, 35*, 1283-1292.

Slattery, T.J, Staub, A., & Rayner, K. (2011). Saccade launch site as a predictor of fixation duration in reading: comments on Hand, Miellet, O'Donnell and Sereno (2010). *Journal of Experimental Psychology: Human Perception and Performance, 38*, 251-261.

Stadthagen-Gonzalez, & Davis, C.J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38*, 598-605.

Stanovich, K. E. & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition, 7*, 77-85.

Stanovich, K.E., & West, R.F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 658-672.

Stanovich, K.E., & West, R.F. (1983). On priming by sentence context. *Journal of Experimental Psychology: General, 112*, 1-36.

Starr, M.S. & Rayner, K. (2001). Eye movements during reading: some current controversies. *Trends in Cognitive Sciences, 5*, 156-163.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics (pp. 327-342).* Oxford, UK: Oxford University Press.

Sternberg, S. (1969a). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.

Sternberg, S. (1969b). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-315.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1140–1154.

Taylor, W. L. (1953). Cloze Procedure: a new tool for measuring readability. *Journalism Quarterly, 30*, 415-433.

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes, 22*, 727–779.

Tolentino, L.C., & Tokowicz, N. (2009). Are pumpkins better than heaven? An ERP investigation of order effects in the concrete-word advantage. *Brain & Language ,110*, 12–22.

Tsai, P-S.,Yu, B., Lee, C-J., Tzeng, C-Y., Hung, D.L., & Wu, D. (2009). An event-related potential study of the concreteness effect between Chinese nouns and verbs. *Brain Research, 1253*, 149-160.

Van Hell, J. G., & De Groot, A. M. B. (1998). Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology, 51A*, 41–63.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8*, 485-531.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*, 380-393.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open-and closed-class words. *Memory and Cognition, 19*, 95-112.

Wattenmaker, W. D., & Shoben, E. J. (1987). Context and the recallability of concrete and abstract sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 140–150.

West,W.C., Holcomb, P.J., (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*, 1024–103.

West, R.F, & Stanovich K.E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development, 49*, 712-727.

West, R.F, & Stanovich K.E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 385-399.

Whaley, C.P (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17*, 143-154.

Whatmough, C.,Verret, L., Fung, D., Cherktow, H. (2004). Common and contrasting areas of activation for abstract and concrete concepts: An H2 15O PET study. *Journal of Cognitive Neuroscience, 16*, 1211-1226.

Williams, R.S**.** & Morris, R.K. (2004). An eye movement analysis of word familiarity and vocabulary acquisition in skilled reading. *European Journal of Cognitive Psychology, 16*, 312-339

Wilson, M.D. (1988). The MRC Psycholinguistic Database: Machine readable dictionary, Version 2. *Behavioral Research Methods, Instruments, & Computers, 20*, 6-11.

Winnick, W.A. & Kressel, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning, and free recall as a function of abstractness-concreteness and word frequency. *Journal of experimental psychology, 70*, 163-168.

Yap, M.J., Balota, D.A. & Tan, S.E. (2013). Additive and interactive effects in semantic priming: Isolating lexical and decision processes in the lexical decision task. . *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 141-158.

Zhang, Q., Guo, C., Ding, J., & Wang, Z. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language 96*, 59–68.

Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics, 36*, 277-284.

Appendices

Appendix A

Experimental materials – Chapter 2 (Experiment 1)

Experimental materials consisted of 200 words (100 critical words and 100 nonwords) presented one at a time on the computer screen. Words are numbered here only for the purpose of identifying them. Conditions 1 to 4 were LF-LI; LF-HI; HF-LI and HF-HI respectively with 25 words (25 critical words and 25 pseudowords) per condition. There were 10 practice trials (not shown). Experimental trials were presented in a different random order for each participant.

<u>Critical words</u>

	Low Frequency-Low Imageablity (LF-LI)	Low Frequency-High Imageablity (LF-HI)	High Frequency-Low Imageablity (HF-LI)	High Frequency-High Imageablity (HF-HI)
1	rust	kilt	loan	diet
2	dent	vase	type	trip
3	cube	cage	area	heat
4	liar	lamp	year	edge
5	oath	stork	error	note
6	brawl	gravy	guide	park
7	pouch	siren	scene	army
8	hobby	jewel	title	love
9	tweed	stain	force	head
10	lobby	rifle	range	metal
11	grief	hedge	staff	brain
12	starch	kennel	voice	trial
13	muzzle	napkin	place	fight
14	puzzle	ribbon	height	radio
15	armour	cellar	search	space
16	poison	saddle	length	music
17	genius	statue	speech	letter
18	temper	pepper	member	gallery
19	exhaust	bagpipe	health	weather
20	apology	mermaid	article	holiday
21	hygiene	steeple	history	student
22	portion	pyramid	problem	teacher
23	mechanic	costume	company	evening
24	monument	ornament	question	meeting
25	composer	fountain	business	magazine

Nonwords

	Low Frequency-Low Imageablity (LF-LI)	Low Frequency-High Imageablity (LF-HI)	High Frequency-Low Imageablity (HF-LI)	High Frequency-High Imageableity (HF-HI)
1	lask	bast	timp	hity
2	blen	hink	trab	felp
3	bame	tace	abia	leem
4	noil	plew	yalk	quib
5	nisk	bruck	drile	wope
6	malve	hanty	larse	prad
7	fough	pober	phyde	enty
8	habby	grall	antle	velm
9	prall	stime	ploin	virp
10	lunny	foble	bloaf	roble
11	clife	terge	cheen	pring
12	thince	baffen	vathe	greel
13	mollen	emblew	pluce	pight
14	pummit	repaim	quance	renia
15	impate	crider	sounge	swoth
16	torgon	wumble	stogue	sanic
17	orkler	chanus	fleach	retter
18	famper	fentry	yander	lindery
19	extrink	blimble	thring	hartmen
20	enogaly	sawtrin	tisibel	numbial
21	curgeon	phottle	jundary	shemper
22	shenvel	ornable	grister	rimbern
23	mantheon	cromdel	sponika	attrage
24	indoment	naprator	blaption	primful
25	andesker	hunction	timbling	quallegy

Appendix B

Specifications of critical words – Chapter 2 (Experiment 1)

Condition 1 LF-LI

Critical word	Number of Letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Orthographic neighbours	Imageability	Age-of-Acquisition
rust	4	3.36	1	4	1	13	538.50	380.00
dent	4	3.60	1	4	1	16	499.67	369.00
cube	4	3.82	1	4	1	6	507.67	468.00
liar	4	4.07	2	3	2	1	425.00	308.00
oath	4	5.87	1	2	1	6	371.50	408.00
brawl	5	1.68	1	4	1	4	447.50	494.00
pouch	5	3.56	1	3	1	5	541.50	367.00
hobby	5	6.33	2	4	1	4	494.00	361.00
tweed	5	8.24	1	4	1	2	450.00	419.00
lobby	5	12.36	2	4	1	3	462.00	406.00
grief	5	14.62	1	4	1	1	433.33	452.00
starch	6	1.42	1	5	1	3	453.50	489.00
muzzle	6	2.14	2	4	1	4	513.00	461.00
puzzle	6	6.13	2	4	1	3	510.00	320.00
armour	6	9.69	2	4	1	2	536.00	400.00
poison	6	10.31	2	4	1	1	511.75	368.50
genius	6	12.22	2	6	2	1	441.50	475.50
temper	6	13.69	2	5	1	3	489.00	333.00
exhaust	7	5.72	2	6	1	0	518.50	454.00
apology	7	7.16	4	7	2	0	400.00	406.00
hygiene	7	8.28	2	5	1	0	459.00	481.00
portion	7	11.87	2	5	1	0	399.00	411.00
mechanic	8	3.08	3	7	2	0	530.00	433.00
monument	8	7.92	3	9	1	0	543.00	444.00
composer	8	10.33	3	7	3	2	467.00	461.00
mean	5.76	7.10	1.84	4.72	1.24	3.20	477.68	414.76
SD	1.30	3.94	0.80	1.54	0.52	3.91	48.99	53.70

Condition 2 LF-HI

Critical word	Number of Letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Orthographic neighbours	Imageability	Age-of-Acquisition
kilt	4	0.91	1	4	1	11	561.00	300.00
vase	4	5.09	1	3	1	9	612.67	321.00
cage	4	10.49	1	3	1	13	612.50	300.00
lamp	4	13.92	1	4	1	12	611.67	284.00
stork	5	0.99	1	5	1	5	570.00	356.00
gravy	5	1.83	2	5	1	2	594.00	269.00
siren	5	2.68	2	5	1	2	578.00	347.00
jewel	5	3.96	1	4	1	1	621.00	292.00
stain	5	5.26	1	4	1	6	558.67	343.00
rifle	5	7.61	2	4	1	0	581.00	322.00
hedge	5	9.84	1	3	1	3	581.50	306.00
kennel	6	1.79	2	4	1	3	580.00	322.00
napkin	6	2.13	2	6	1	0	582.00	342.00
ribbon	6	6.94	2	5	1	1	563.00	286.00
cellar	6	6.96	2	4	1	1	599.50	387.00
saddle	6	7.80	2	4	1	2	578.00	344.00
statue	6	10.08	2	5	1	1	562.00	406.00
pepper	6	10.40	2	4	1	2	586.33	320.67
bagpipe	7	0.27	2	6	2	0	618.50	405.00
mermaid	7	0.94	2	5	1	0	604.00	313.50
steeple	7	1.07	2	5	1	1	559.00	361.00
pyramid	7	5.76	3	7	1	0	613.00	397.00
costume	7	6.93	2	6	1	0	559.00	331.50
ornament	8	2.90	3	8	1	0	594.00	303.00
fountain	8	7.27	2	5	1	1	602.00	389.00
mean	5.76	5.35	1.76	4.72	1.04	3.04	587.29	333.91
SD	1.20	3.78	0.60	1.21	0.20	4.00	20.76	39.56

Condition 3 HF-LI

Critical word	Number of Letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Orthographic neighbours	Imageability	Age-of-Acquisition
loan	4	41.13	1	3	1	14	390.80	460.75
type	4	178.03	1	3	1	5	393.50	414.75
area	4	349.41	3	4	1	3	394.00	392.00
year	4	727.18	1	3	1	14	409.00	269.00
error	5	41.06	2	3	2	1	408.00	425.00
guide	5	63.93	1	3	1	3	440.00	367.67
scene	5	73.26	1	3	1	2	406.00	425.00
title	5	106.64	2	4	1	3	413.00	375.00
force	5	168.38	1	4	1	3	382.75	403.00
range	5	219.92	1	4	1	5	368.67	456.00
staff	5	236.47	1	4	1	2	455.67	424.00
voice	5	275.40	1	3	1	1	429.50	275.00
place	5	497.73	1	4	1	5	373.25	330.67
height	6	40.12	1	3	2	1	436.00	294.00
search	6	67.89	1	3	1	1	366.00	344.00
length	6	75.51	1	4	1	0	379.33	330.00
speech	6	83.47	1	4	1	0	430.00	338.50
member	6	184.58	2	5	1	0	399.00	392.00
health	6	256.07	1	4	1	2	410.67	410.50
article	7	71.83	3	6	1	0	424.50	405.00
history	7	206.49	2	7	2	0	376.00	391.00
problem	7	282.09	2	7	1	0	411.00	367.00
company	7	427.52	3	7	2	0	426.00	436.00
question	8	254.54	2	7	2	0	398.00	314.00
business	8	371.33	2	6	1	1	441.00	436.00
mean	5.64	212.00	1.52	4.32	1.20	2.64	406.47	379.03
SD	1.19	167.18	0.71	1.46	0.41	3.80	24.65	54.63

Condition 4 HF-HI

Critical word	Number of Letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Orthographic neighbours	Imageability	Age-of-Acquisition
diet	4	44.34	2	4	1	9	486.00	442.00
trip	4	48.06	1	4	1	8	475.00	314.00
heat	4	62.30	1	3	1	11	476.75	299.67
edge	4	80.89	1	2	1	1	502.50	308.00
note	4	109.78	1	3	1	12	504.50	303.67
park	4	116.67	1	4	1	15	591.40	283.75
army	4	121.80	2	4	1	2	615.50	337.50
love	4	232.27	1	3	1	16	508.00	277.25
head	4	399.91	1	3	1	13	537.67	312.00
metal	5	47.76	2	4	1	2	564.00	327.00
brain	5	49.99	1	4	1	7	601.67	332.67
trial	5	68.93	2	5	2	3	516.00	433.00
fight	5	71.44	1	3	1	9	511.50	272.50
radio	5	89.97	3	5	1	3	613.00	317.00
space	5	131.33	1	4	1	6	482.67	353.00
music	5	158.07	2	6	1	0	549.00	272.00
letter	6	135.90	2	4	1	7	616.00	324.00
gallery	7	47.07	3	6	1	0	581.50	454.50
weather	7	58.02	2	4	1	4	488.33	358.33
holiday	7	75.16	3	6	2	0	587.67	312.00
student	7	82.20	2	6	2	0	615.00	399.00
teacher	7	89.47	2	4	2	1	576.00	298.50
evening	7	141.17	2	5	1	0	559.00	303.00
meeting	7	206.42	2	5	2	1	485.50	390.50
magazine	8	49.51	3	7	1	0	614.00	391.00
mean	5.36	108.74	1.76	4.32	1.20	5.20	546.33	336.63
SD	1.35	78.50	0.72	1.22	0.41	5.18	51.97	53.08

Appendix C

Experimental Materials – Chapter 3 (Experiment 2) and Chapter 4 (Experiment 3)

Experimental materials consisted of 150 sets of short passages over two lines on the computer screen in the format shown below. Sentences are also numbered here only. For the purpose of identifying the target word, it is shown in **bold** here. The word in parentheses is the invalid preview orthographically legal letter string used in Experiment 3. The word in bold is the target word used in Experiment 2; in Experiment 3, the invalid preview changed to the respective targets when the participants eyes crossed an invisible pre-defined boundary (placed after the last character of the pre-target word). Conditions 1 to 6 were LF-LP, LF-MP, LF-HP, HF-LP, HF-MP and HF-HP respectively with 25 passages in each condition. There were 48 comprehension questions in total, half with yes-answers and half with no-answers to make sure participants were paying attention. There were 5 practice trials which were the same for every participant. Experimental trials were presented in a different random order for every participant.

Condition 1 LF-LP

1 Arthur was at home, preparing vegetables to accompany his dinner.
He steamed some (gron)**peas** and spooned them onto the side of his plate.

2 Overnight, vandals had ruined my prized cherry blossom tree.
They had ripped off the (dest)**bark** and scattered it across the lawn.

3 Anthony and Harriet were allowed to go out and play after school.
Their parents wanted them home before (bant)**dusk** because it was a weeknight.

4 The careless trainee caused havoc on the petrol station forecourt.
He knocked over an entire (fure)**drum** of oil and it made a dreadful mess.

5 Patsy strongly believed that the UK judicial system needed reform.
It was clear that (yeat)**jail** was not enough of a deterrent to criminals.

6 City officials planned to improve levels of public welfare.
They were to upgrade the (acrem)**sewer** system with improved filters.

7 Jessica was keen to try and improve her diet by eating more fruit.
She went to the shop to buy some (roter)**melon** chunks to eat at her desk.

8 The house hunters had a very good idea of what they were looking for.
They wanted a house with a large (eddus)**attic** to convert into a home office.

9 It was very difficult finding a new leather jacket that fitted.
I eventually found one at a (aboff)**stall** at the local outdoor market.

10 Justin was an enthusiastic baker and was eager to try new things.
He was excited about trying his new recipe for (reary)**icing** at the weekend.

11 The old pirate sat in the ale house thinking back to his younger days.
He fondly remembered his **(**gessal)**parrot** who had been a great companion at sea.

12 Zak wasn't enjoying working as the mechanic in the farming village.
He hated it when the (peasen)**grease** from the engines got onto his overalls.

13 Sara always remembers to collect her morning paper on the way to work.
She enjoys the (goncho)**puzzle** pages and eagerly tries to finish the crossword.

14 Heather had spent the last few nights tossing and turning in bed.
She thought she should change her (yetter)**pillow** for a new feather one.

15 Tracy got a glass and began making herself a Bloody Mary.
She added some (joggan)**pepper** to give it a little bit of a kick.

16 The bird watchers were silent as they waited for the rare swallows.
They were startled when a (noddel**)rabbit** suddenly dashed out from the bushes.

17 Paul had to go into town to pick up some groceries.
He got on the (creedom**)scooter** and revved up the engine.

18 The electrical goods shop was having a huge sale this weekend.
There was a good deal on a (thortom)**blender** that had multiple settings.

19 The old friends were heading away for a golfing weekend in Portugal.
It would be nice to have a (vacuune**)reunion** every year but it was not possible.

20 The distraught child was in floods of tears and inconsolable.
He'd lost hi (littarm)**balloon** and the man selling them had none left.

21 The evening news carried a special report about conflict in the Sudan.
The story was told by a (vilipon**)refugee** who had managed to escape the country.

22 The local community centre was offering a variety of evening classes.
Denise enrolled in the (jelling)**pottery** class since she wanted to make vases.

23 My Mum has many delicate china trinkets arranged on her mantle piece.
She carefully dusts each (rimonoul)**ornament** every week to keep them pristine.

24 Tina often had a hard time at friends' parties when she was a child.
Growing up with (bendulum)**diabetes** meant that she had to watch what she ate.

25 German U-boats patrolling the Atlantic fired torpedoes at the convoy.
One of them struck a (sountane**)civilian** ship and several dozen souls were lost.

Condition 2 LF-MP

26 Colin couldn't resist the advances of the sexy new secretary.
He was overcome with (tark)**lust** and embraced her passionately.

27 The critically acclaimed restaurant was fully booked once again.
They had hired a talented (skib)**chef** who had transformed their menu.

28 Elspeth ran away from her parents' home when she was seventeen.
She joined a (noth)**cult** that promised to take her away on a space ship.

29 The little boy enjoyed dressing up and pretending to be Superman.
He would put on a (zoya)**cape** and zoom around the house as if he were flying.

30 Lucy returned home after another hard day at the office.
She slumped onto the (ribe**)sofa** and turned on the television.

31 When preparing a turkey, you do not have to throw away the giblets.
These can be used to make (quing**)gravy** to be served with the roasted bird.

32 Rory was going to dig all day in the potato fields.
He picked up his (egiln)**spade** and headed off to work.

33 The school children were impressed after their trip to the aquarium.
They admired the (clend)**shark** as it slipped through the water like a knife.

34 The bar brawl ended with Frank being hit hard square in the mouth.
The punch broke his (brell**)tooth** and would need to be capped.

35 The Emperor celebrated his victory by arranging a lavish banquet.
It was a sumptuous (barch)**feast** which was heartily devoured by his Generals.

36 The magazine had a special feature dedicated to Britney Spears.
It included a full size (gealon)**poster** of her from her latest concert tour.

37 Dan set the table in preparation for his romantic meal that night.
In the middle of the table, he placed a (murtha**)candle** which he later lit.

38 Dave wanted to build a new bookcase but couldn't find his toolbox.
Eventually, he had to borrow a (berrow)**hammer** and nails from his neighbour.

39 Lola admired the grace of the Bolshoi dancers performing "Swan Lake".
She loved going to the (liddal)**ballet** whenever this company came to town.

40 A recent biography revealed Cuthbert's passion for going on safari.
He was a ruthless (fealon)**hunter** with a reputation for killing big game.

41 The Master insisted that his butler pressed his shirts immaculately.
Hobbs would carefully starch the (vatton**)collar** of each shirt every morning.

42 Holidaymakers on Corfu were unprepared for the intense heat wave.
The next day, many of them had (vorlone)**sunburn** that needed medical attention.

43 Will's hair was a mess after his friends poured green paint over him.
He used nearly a whole bottle of (clunger)**shampoo** to try and get it out.

44 Violet and Quentin were having a heated argument in their lounge.
Swearing loudly, she picked up a (recline)**cushion** and threw it at his head.

45 Matt had a habit of continually hunching over his keyboard.
He knew that his bad (quedric)**posture** could lead to future back problems.

46 In many parts of Africa it is common for no rain to fall for months.
It is known that (teeryth)**drought** often causes crop failure in these regions.

47 The school cafeteria was in a ghastly state after the pupils had left.
It was unfair to expect the (stamous)**cleaner** to come in and sort out this mess.

48 Kieran planned to get his wife an expensive birthday present.
He knew she really wanted a (smoltern)**necklace** which was made from pearls.

49 Thomas Edison was one of the great pioneers of the Industrial Era.
He is still highly respected as an (crosslen)**inventor** centuries after his death.

50 Returning from work, Holly saw that her house had been broken into.
She called to report the (deapting)**burglary** and waited for the police to arrive.

Condition 3 LF-HP

51 Martin lost his temper and broke his wife's favourite ornament in two.
He would need to find some (phem)**glue** to repair it before she got home.

52 Jimmy ran down the hill, gripping the string in the strong winds.
He loved to play with his (delm)**kite** but rarely got the right conditions.

53 Gordon bought another replacement fuse for his unreliable kettle.
He removed the screws from the (girp)**plug** and noticed burn marks inside.

54 Protestors picketed the zoo about the conditions the lion was kept in.
They wanted it removed from its (noya)**cage** and allowed to roam in a paddock.

55 The timer buzzed and Mum knew that her apple pie was finally ready.
She lifted it out of the (vire)**oven** and placed it on the counter to cool.

56 Joey excitedly told his parents he saw a striped horse at the zoo.
His parents explained that the animal was a (maken)**zebra** from Africa.

57 Ponies and horses are not suited to travelling across deserts.
The best animal for this is the (swart)**camel** as it rarely needs water.

58 Alison gasped as she dropped a glass of red wine onto the carpet.
Hurriedly, she tried to remove the (whone)**stain** with some soap and water.

59 Old Mrs. Greeble was warty, haggard, and had a fearsome black cat.
The older kids said that she was a (anlet)**witch** to scare the younger ones.

60 After his morning jog, Gregor was happy to take a long, hot shower.
When he stepped out, he reached for his (darch)**towel** but it wasn't there.

61 Mick left his car headlights on overnight and his car wouldn't start.
He lifted up the (dessol)**bonnet** and attached jump leads to the battery.

62 Harry went to Saville Row in London to purchase his new suit.
He had an appointment to be measured by a (fortus)**tailor** at great expense.

63 At work, the boiler had broken and we were freezing at our desks.
We arranged for a portable (fankon)**heater** to be brought into the office.

64 At the party, Ryan discovered that all the beer cans were warm.
No one had put any in the (dastyr)**fridge** and they tasted unpleasant.

65 Brigit sat in admiration, examining the structure of the delicate web.
She was amazed that one (egelon)**spider** could produce something so intricate.

66 The rebels decided to assassinate the President by tainting his food.
One of them infiltrated the kitchen and added the (graver)**poison** to his soup.

67 Nick gazed hungrily into the tank at the expensive seafood restaurant.
He selected the biggest (tefelom)**lobster** to be taken away and cooked.

68 The danger posed by mosquitoes in hot climates is well known.
The main threat is the spread of (notious)**malaria** which can cost lives.

69 Diana’s supply of clean clothes was diminishing rapidly.
She really needed to do her (funelop)**laundry** before she completely ran out.

70 Emma was overcome by a strong scent at the store’s beauty counter.
A bottle of expensive (junction)**perfume** had fallen from the shelf and smashed.

71 Little Kirsty spent hours dressing up as a witch on Halloween.
Everyone at school thought that her (seaform)**costume** was the best this year.

72 Most retailers allow customers to return goods within thirty days.
However, they have to provide their (senough)**receipt** as proof of purchase.

73 The newlyweds left the church to loud cheers from their guests.
Everyone threw (vertilla)**confetti** as they made their way to the wedding car.

74 Julie was getting ready for a date and nervously put on her make-up.
She wore red (taychoal)**lipstick** as it flattered her and matched her outfit.

75 Often, women are paid less than men for doing similar jobs.
The fight for (spartulp)**equality** is still ongoing in most professions.

Condition 4 HF-LP

76 The zookeepers were busy preparing for their latest arrival.
They were getting a baby (doss)**bear** that had been born in America.

77 The sales team were pushing hard as the end of the month loomed.
They had been set a certain (pait)**goal** by the director of the company.

78 Chloe loved going outside to have fun with her assorted toys.
She loved to play with her (teth)**ball** in the yard when the sun shone.

79 After his mother’s death, Leo was left with large medical bills.
He wanted to sell her (fook)**land** as quickly as he could to raise cash.

80 Bertie arrived in York without having arranged accommodation.
He checked into the first (wace)**room** he found that he could afford.

81 Juliet kissed her husband on the cheek as he was leaving for work.
She noticed that he had left his (gleas)**phone** and ran outside after him.

82 James had scratchy tonsils and suspected he was getting a bad cold.
The next morning, his smarm **| voice** was hoarse and it was painful to speak.

83 The student in the flat was unhappy with their useless flatmate.
He was having his (grase)**power** cut off because of unpaid utility bills.

84 The men looked very presentable in their white shirts and black ties.
They were going to a (grelp)**party** dressed as characters from Reservoir Dogs.

85 The burglar was quiet and efficient as he stole the valuables.
He quickly ran to the (bream)**house** across the street and robbed it too.

86 Meeting Winston, you would never have guessed he was as old as he was.
Age had not affected his (saring)**memory** the way it affected many other people.

87 Paddy was unsure about what to do next as he neared his thirties.
He didn't have a (vesume)**career** which was either enjoyable or challenging.

88 Local businesses donated to a regeneration fund for the town centre.
There were plans for a (puelor)**garden** to be built with colourful flowers.

89 The House of Commons was full to the rafters for the important debate.
There wasn't a single (rondaw)**member** of any of the parties who didn't attend.

90 Carlos spent thousands on drinking binges and gluttonous meals out.
He didn't think that (banilt)**health** was something he needed to worry about.

91 The politician was greeted with boos when he visited the school.
It was obvious that the (geddle)**public** were not happy with his policies.

92 There was uproar amongst the audience at the town planning meeting.
There were plans to build a (leshing)**factory** on the local playing fields.

93 The new manager was finding it difficult to exert any real authority.
He realised that (virgoul)**respect** had to be earned, and wouldn't come easily.

94 Elaine received bad news from home and needed to get time off work.
She asked to swap her (sarbuck)**weekend** shift so that she could visit her Mum.

95 Dmitri was enjoying his work in the Immunology lab at the hospital.
He had to put each (rabbain)**culture** into refrigeration at the end of the day.

96 It was difficult for the young soldier to be posted so far from home.
He received a (gerlain)**picture** that his daughter had drawn and he shed a tear.

97 The General in charge of security in Afghanistan had to be tactful.
He had to take into account the (teching)**history** between the various tribes.

98 Tanya had to draw a picture of something she had done on holiday.
She drew a picture of a (nesslure)**mountain** that she he had climbed near Oban.

99 Helen thought that going to the new romantic comedy would be fun.
She phoned her (terythan)**daughter** to see if she would like to go with her.

100 The retired couple holidayed in Spain at least six months of the year.
It made sense when they bought a (jingrelp)**property** and moved there for good.

Condition 5 HF-MP

101 My favourite hobby is going to see musicals at the theatre.
I normally pay extra so that my (nand)**seat** is near to the stage.

102 I struggled to read the badly printed manual for my new computer.
It had little space between the lines of (dord)**text** and strained my eyes.

103 Sue spent hours preparing a variety of dishes for her dinner party.
Her guests agreed that the (tunk)**food** was wonderful so it was all worth it.

104 Dianne was fed up dealing with sullen and un-cooperative colleagues.
It had been a difficult (mant)**week** but fortunately it would soon be over.

105 Overall, Rose was very satisfied both personally and professionally.
She had a good (tolo)**life** and hoped things would stay settled for a while.

106 Ken had forgotten to water Isla's geranium while she was on holiday.
When she got home, she saw that the (ghock)**plant** was completely dried out.

107 Inflation commonly rises by a small percentage every year.
The result is an increase in the (gorer)**price** of goods that we purchase.

108 Edgar was worried about getting burgled when he went out at night.
He usually left a (fryth)**light** on to make it look as if someone was home.

109 Mood around the office was glum and the boss needed to take action.
Organising a party for the (wholt)**staff** would hopefully boost morale.

110 Trying to sleep on Christmas Eve was never easy when we were kids.
It was the most exciting (cryth)**night** of the year without a doubt.

111 Doug was annoyed when the commercials started blaring from his TV.
He quickly turned the (antion)**volume** down and went to make a cup of tea.

112 This was Maximillion's third appearance in court in five months.
He was sure he would be sent to (jarve)**prison** this time for his crimes.

113 Many animals must hibernate in order to survive harsh climates.
At the end of the (sorkan)**winter** they will wake up and forage for food.

114 The group thought that Larry was the best decision maker amongst them.
They chose him to be the (fanton)**leader** and he graciously accepted.

115 Mary felt bad about pretending to be sick to avoid dinner with Tim.
She hadn't been much of a (heanal)**friend** and apologised the next day.

116 My neighbours and I wanted to commemorate the Queen's Golden jubilee.
We organised a big party in the (chound)**street** which went on all night.

117 Dr. Fox visited the ward to answer questions about the operation.
When he met his (yodench)**patient** he assured him it was a routine procedure.

118 Hepatitis affects the liver and can be transmitted by transfusions.
It is a serious (breanen)**disease** which requires immediate hospitalisation.

119 Angela loved to knit and saw a great idea for a nice spring jumper.
She cut out the (geddace)**pattern** from the magazine and went to buy the wool.

120 The owner of the large estate built an ostrich farm on his land.
Everyone in the (nethepo)**village** thought that he was a bit of an eccentric.

121 The supervisor thought she should speak to the student personally.
She planned to have a (corbery)**meeting** to discuss the student's progress.

122 The UK is experiencing a vast influx of foreign migrant workers.
They come to our (ninship)**country** to try and make better lives for themselves.

123 Buddhism is growing in popularity and has many famous followers.
The principles of this (mitapane)**religion** emphasise finding inner contentment.

124 Geoff headed to the pub to watch the final of the F.A. Cup.
They would always watch (benddeth)**football** no matter who was playing.

125 Airports have recently taken numerous steps to prevent terror attacks.
There are now more (cronnelp)**security** checks in order to protect our safety.

Condition 6 HF-HP

126 Gary had just started a new job helping tidy up the barber's shop.
He had to sweep up the (torm)**hair** from the floor at the end of each day.

127 When Alex arrived at his friend's house, he rang the bell.
He heard footsteps behind the (bire)**door** as his friend came to let him in.

128 Lisa had moved to London to start a new job at a large legal firm.
It was the first time she had lived in a big (selp)**city** and she was excited.

129 Detective Mills arrived at the murder scene early in the morning.
A neighbour discovered the (holp)**body** after hearing screams in the night.

130 Ron and Jen had especially requested a room on the hotel's top floor.
They knew that it would give them a fabulous (weam)**view** of the ocean.

131 The night after her day at the zoo, Natalie fell into a deep sleep.
However, she had a very unusual (hower)**dream** about being chased by a chimp.

132 At the school play, Bill waited nervously for the curtain to rise.
When he went onto the (adepi)**stage** he gave a breathtaking performance.

133 Adam's behaviour at school was getting out of control.
He kept disrupting the (stean)**class** and would have to be sent to the Head.

134 Joyce was responsible for arranging her tennis match with Molly.
She had booked the (snock)**court** at noon so they could lunch afterwards.

135 Despite keeping spending to a limit, Liz struggled to budget properly.
She always ran out of (wrimp)**money** before the end of the month.

136 In the morning, most people brush their teeth and use mouthwash.
This freshens their (derilt)**breath** before they go to work or school.

137 Businesses have simple models in which they try to maximise income.
They try to make as much (jontal)**profit** as possible to please investors.

138 The young boy recklessly kicked his ball in front of the house.
One day, he broke a (anster)**window** and blamed it on his little brother.

139 Two men in the pub started fighting very viciously.
The barman phoned the (gation)**police** as more people began to get involved.

140 The traders arrived at 5 AM to get their stalls ready for business.
Saturday was a busy day at the (ristol)**market** and setting up early was vital.

141 I planned a big celebration for my parents' 50th wedding anniversary.
I invited the whole of my (lentig)**family** to a reception in a fancy restaurant.

142 After nearly 25 years with a mortgage, Jack only had one month left.
He only had to make one more (jogreel)**payment** and he'd finally own his house.

143 Ben slept through his alarm and was going to be late for his train.
He hurriedly drove to the (cholane)**station** only to see it pulling away.

144 I still receive letters for the previous tenants of my flat.
I wish that they would change their (uttrion)**address** as it is very annoying.

145 Nicola was revising frantically for her end of year degree exams.
She would spend hours in the (belving)**library** with her head buried in books.

146 Ian prepared his sandwiches for work before he went to bed.
This meant that in the (neamery)**morning** he only had to put them in his bag.

147 The binmen had not removed Stanley's garbage for nearly three weeks.
He decided to phone the (nonsect)**council** so that he could make a complaint.

148 Interpol knew the thief of the Mona Lisa planned to keep it to himself.
It would be too difficult to sell the (quesdary)**painting** as it is too well-known.

149 The latest Cosmopolitan had a picture of George Clooney on the cover.
Sales of the (ampearic)**magazine** would receive a boost because of his popularity.

150 Terry had just found out that his wife had gone into labour.
He rushed to the (tangulet)**hospital** to be present at the delivery.

Appendix D

Target word characteristics – Chapter 3 (Experiment 2) and Chapter 4 (Experiment 3)

LF=low frequency; HF=high frequency; LP=low predictability; MP=moderate predictability; HP =high predictability; Freq. per mill.=frequency of occurrence per million words; Imageability=score from 100 (low imageability) to 700 (high imageability); Rating=rating task score (1-6, where 1 is low predictable and 6 is highly predictable); Cloze=Cloze task rating score. All figures are displayed to 2 decimal places.

Condition 1 LF-LP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
peas	4	5.77	1	671.43	4.67	0.05
bark	4	6.20	1	520.00	4.00	0.05
dusk	4	6.71	1	559.00	4.33	0.00
drum	4	10.29	1	627.00	4.00	0.00
jail	4	13.24	1	629.00	4.72	0.00
sewer	5	2.13	2	538.00	3.83	0.05
melon	5	2.19	2	685.71	3.94	0.05
attic	5	6.89	2	557.14	3.89	0.00
stall	5	7.83	1	461.00	4.11	0.00
icing	5	8.84	2	604.76	4.00	0.00
parrot	6	4.00	2	670.00	4.44	0.00
grease	6	4.12	1	517.00	5.28	0.00
puzzle	6	6.13	2	513.00	3.94	0.00
pillow	6	7.53	2	622.00	5.56	0.05
pepper	6	10.40	2	575.00	3.78	0.00
rabbit	6	14.58	2	637.00	4.17	0.00
scooter	7	0.62	2	569.00	3.17	0.00
blender	7	0.73	2	623.81	3.89	0.05
reunion	7	6.01	4	357.14	5.00	0.05
balloon	7	6.73	2	619.00	3.39	0.00
refugee	7	9.16	3	500.00	5.11	0.00
pottery	7	9.56	3	585.71	4.61	0.00
ornament	8	2.90	3	594.00	5.61	0.00
diabetes	8	7.04	4	469.00	3.33	0.00
civilian	8	14.90	4	357.14	4.17	0.00
Mean	5.88	6.98	2.08	562.47	4.28	0.01
SD	1.33	3.88	0.95	86.84	0.66	0.02

Condition 2 LF-MP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
lust	4	5.22	1	395.00	5.83	0.60
chef	4	7.20	1	620.00	6.06	0.70
cult	4	10.03	1	419.00	4.28	0.35
cape	4	10.84	1	569.00	6.44	0.70
sofa	4	10.97	2	608.00	5.72	0.65
gravy	5	1.83	1	597.00	5.11	0.30
spade	5	3.03	1	549.00	5.56	0.70
shark	5	3.38	1	642.00	5.28	0.35
tooth	5	6.10	1	635.00	5.44	0.50
feast	5	9.54	1	592.00	6.28	0.65
poster	6	7.40	2	600.00	5.50	0.40
candle	6	8.59	1	600.00	6.11	0.70
hammer	6	11.76	2	629.00	5.39	0.40
ballet	6	13.51	2	600.00	6.33	0.75
hunter	6	14.19	2	567.00	4.78	0.45
collar	6	14.62	2	582.00	5.22	0.60
sunburn	7	0.67	2	629.00	6.28	0.40
shampoo	7	2.87	2	623.81	5.89	0.65
cushion	7	5.30	2	613.00	3.72	0.20
posture	7	5.82	2	433.00	6.00	0.55
drought	7	6.94	1	367.00	6.17	0.70
cleaner	7	10.03	2	532.21	4.83	0.40
necklace	8	2.78	2	618.00	5.06	0.20
inventor	8	3.04	3	534.00	5.94	0.55
burglary	8	5.93	3	476.00	6.33	0.60
Mean	5.88	7.26	1.64	561.20	5.58	0.52
SD	1.33	4.01	0.64	80.34	0.69	0.17

Condition 3 LF-HP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
glue	4	6.99	1	640.00	5.61	0.95
kite	4	7.76	1	646.00	5.83	0.90
plug	4	8.13	1	558.00	5.56	0.90
cage	4	10.49	1	596.00	5.89	0.90
oven	4	12.98	2	609.00	6.67	1.00
zebra	5	2.21	1	650.00	6.72	1.00
camel	5	4.22	2	616.00	6.50	1.00
stain	5	5.26	1	529.00	6.22	0.95
witch	5	6.59	1	633.00	6.22	1.00
towel	5	8.84	2	578.00	6.11	1.00
bonnet	6	4.13	2	590.48	5.50	1.00
tailor	6	4.48	2	500.00	6.44	0.95
heater	6	4.82	2	566.67	6.67	1.00
fridge	6	5.91	1	620.00	6.06	1.00
spider	6	6.28	2	620.00	6.28	1.00
poison	6	10.31	2	538.00	6.28	0.95
lobster	7	2.79	2	613.00	5.89	0.95
malaria	7	2.99	4	484.00	6.44	0.95
laundry	7	5.80	2	562.00	6.28	1.00
perfume	7	5.96	2	590.48	6.50	1.00
costume	7	6.93	2	530.00	6.11	0.95
receipt	7	11.73	2	435.00	6.39	0.95
confetti	8	0.71	3	647.62	6.17	0.85
lipstick	8	4.16	2	680.95	6.50	0.85
equality	8	16.22	4	346.00	6.39	1.00
Mean	5.88	6.67	1.88	575.17	6.21	0.96
SD	1.33	3.58	0.83	75.28	0.34	0.05

Condition 4 HF-LP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
bear	4	62.19	1	625.00	3.83	0.00
goal	4	65.70	1	506.00	5.67	0.00
ball	4	82.34	1	634.00	4.50	0.00
land	4	233.26	1	580.00	3.72	0.00
room	4	320.96	1	555.00	5.11	0.05
phone	5	84.98	1	630.00	3.89	0.00
voice	5	275.40	1	451.00	5.94	0.05
power	5	351.02	2	450.00	4.00	0.00
party	5	441.56	2	596.00	4.89	0.00
house	5	547.72	1	636.00	5.44	0.05
memory	6	82.34	3	391.00	4.17	0.00
career	6	84.58	2	421.00	5.00	0.05
garden	6	120.57	2	643.00	3.83	0.00
member	6	184.58	2	402.00	5.00	0.00
health	6	256.07	1	427.00	4.11	0.00
public	6	428.80	2	451.00	4.50	0.00
factory	7	47.30	3	608.00	3.89	0.00
respect	7	60.84	2	346.00	5.06	0.00
weekend	7	72.79	2	271.43	3.78	0.00
culture	7	93.44	2	313.00	4.22	0.00
picture	7	110.24	2	580.00	4.44	0.00
history	7	206.49	3	422.00	3.83	0.00
mountain	8	43.00	2	644.00	3.56	0.00
daughter	8	98.71	2	490.00	3.50	0.00
property	8	134.32	3	455.00	5.83	0.00
Mean	5.88	179.57	1.80	501.10	4.47	0.01
SD	1.33	142.00	0.71	112.98	0.74	0.02

Condition 5 HF-MP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
seat	4	65.41	1	612.00	5.33	0.55
text	4	80.59	1	470.00	5.83	0.75
food	4	207.57	1	495.00	6.39	0.45
week	4	351.94	1	430.00	4.78	0.30
life	4	611.76	1	440.00	5.00	0.20
plant	5	90.26	1	617.00	6.11	0.75
price	5	203.60	1	360.00	6.06	0.65
light	5	250.49	1	536.00	5.56	0.75
staff	5	249.91	1	469.00	6.00	0.65
night	5	388.10	1	606.00	6.56	0.50
volume	6	61.39	2	469.00	6.28	0.50
prison	6	68.46	2	574.00	6.39	0.60
winter	6	79.04	2	627.00	6.11	0.60
leader	6	101.98	2	503.00	5.89	0.70
friend	6	181.93	1	559.00	4.61	0.40
street	6	209.89	1	604.00	5.17	0.45
patient	7	90.97	2	529.00	6.50	0.70
disease	7	98.47	2	463.00	6.39	0.60
pattern	7	99.04	2	456.00	5.28	0.70
village	7	123.31	2	584.00	3.78	0.25
meeting	7	233.02	2	449.00	5.89	0.70
country	7	348.90	2	555.00	6.17	0.75
religion	8	48.76	3	434.00	6.39	0.60
football	8	74.58	2	611.00	6.56	0.40
security	8	152.12	4	394.00	6.33	0.55
Mean	5.88	178.86	1.64	513.84	5.81	0.56
SD	1.33	133.85	0.76	77.88	0.71	0.16

Condition 6: HF-HP

Target	Number of letters	Frequency per million	Number of syllables	Imageability	Rating	Cloze
hair	4	150.79	1	595.00	6.72	0.90
door	4	252.46	1	597.00	5.89	0.95
city	4	254.44	1	605.00	6.11	0.95
body	4	273.97	1	629.00	6.50	0.95
view	4	277.23	1	375.00	6.06	1.00
dream	5	49.64	1	479.00	5.83	0.95
stage	5	169.40	1	610.00	6.17	1.00
class	5	199.30	1	560.00	6.39	1.00
court	5	316.46	1	591.00	6.22	1.00
money	5	341.77	2	589.00	6.61	0.95
breath	6	55.54	1	480.00	6.22	1.00
profit	6	61.72	2	240.00	6.33	0.90
window	6	107.10	2	602.00	6.33	0.95
police	6	288.06	2	630.00	6.50	1.00
market	6	318.07	2	585.00	6.00	0.90
family	6	363.53	3	589.00	6.39	1.00
payment	7	57.44	2	465.00	6.50	1.00
station	7	69.49	2	565.00	6.00	1.00
address	7	70.22	2	399.00	6.44	1.00
library	7	86.72	3	587.00	6.11	0.95
morning	7	182.03	2	579.00	6.28	0.95
council	7	299.39	2	408.00	5.94	1.00
painting	8	45.59	2	605.00	6.50	0.90
magazine	8	49.51	3	590.00	6.17	0.95
hospital	8	156.67	3	601.00	6.39	1.00
Mean	5.88	179.86	1.76	542.20	6.26	0.97
SD	1.33	109.79	0.72	97.31	0.23	0.04

Appendix E

Experimental materials – Chapter 5 (Experiment 4)

Experimental materials consisted of a total of 240 target words, each of which was presented in the second line of a two line passage of text, in the format shown below on the computer screen. Sentences are numbered here only. For the purpose of identifying the target in the passage of text, it is shown in **bold**. The word shown in parentheses is the invalid preview letter string used in place of the target. This invalid letter string changed to the target itself when the reader's eyes crossed the invisible boundary, which was placed immediately after the last letter of the pre-target word. The 8 conditions respectively were: HF-HP Valid; HF-HP Invalid; HF-LP Valid; HF-LP Invalid; LF-HP Valid; LF-HP Invalid; LF-LP Valid and LF-LP Invalid (with a total of 30 passages of text in each of the 8 conditions). This lead to a total of 240 passages of text; of which half were HF targets and half were LF targets. There were also a total of 77 questions, with either a yes/no answer required. The questions were based on content in either line 1 or 2 of the passage of text. This was to ensure that readers were paying attention during the course of the experiment. There were eight practice trials, which were the same for every participant. The experimental trials were presented in a different random order for every participant.

HF = high frequency; LF = low frequency; HP = high predictability; LP = low predictability

1	HF-HP	Simon was very careful when he shaved before his job interview.
	HF-LP	Simon used his arms to brace his fall on the pavement.
		An unsightly cut on his (toin)**face** would look bad to his employers.
2	HF-HP	Jessica's cat was moulting over all the furniture in her flat.
	HF-LP	Jessica started cleaning her flat by dusting the surfaces.
		The sofa was covered in (tine)**hair** and needed to be vacuumed.
3	HF-HP	The fisherman began to row his little boat on his fishing trip.
	HF-LP	He was determined to return with a prize worthy of a top champion.
		When he got the middle of the (bula)**lake** he cast his line.
4	HF-HP	The referee's watch said time was up and he raised his whistle.
	HF-LP	The amateur referee ran to the traffic cone located near the pitch.
		Placing it to his (tayn)**lips** he gave a loud blast when the game ended.
5	HF-HP	I often make mistakes when giving people their birthday gifts.
	HF-LP	People always say that I'm absent-minded but well-intentioned.
		Forgetting to sign the (mesk)**card** is normally my biggest mistake.
6	HF-HP	Trips on the train can be very boring without something to do.
	HF-LP	Some people keep entertained on long trips by watching DVDs.
		Remember to take a good (fand)**book** and dull journeys will seem shorter.
7	HF-HP	The consensus was that the serial strangler had struck again.
	HF-LP	The examination was carried out as a matter of normal procedure.
		The marks around the (crat)**neck** confirmed the detective's conclusion.
8	HF-HP	Everyone knew that "EastEnders" was just beginning.
	HF-LP	We were so busy baking cakes in the kitchen we had forgotten the time.
		We recognised the familiar (blorn)**theme** tune and sat down to watch.

9 HF-HP Dave and Gordon watched the boxers in the match exchange blows.
HF-LP Dave and Gordon were going to watch the match on the HDTV in the pub.
Afterwards, they agreed that the (dryth)**fight** was very exciting.

10 HF-HP The analysts found a tendency for increased spending in the recession.
HF-LP The Gulf air would bring warmer temperatures to the UK.
It was forecast that this (leask)**trend** would continue for a while.

11 HF-HP One advantage of Britain being an island is the fabulous beaches.
HF-LP There are many popular tourist destinations in France.
It is very common for people to visit the (wrick)**coast** for a holiday.

12 HF-HP More bets are placed on the Grand National than any other race.
HF-LP Sporting events are often wagered upon to win money.
People choose the (frour)**horse** they think will win and place their bets.

13 HF-HP I was dozing on the sofa as I waited for a call from the plumber.
HF-LP I was dozing on the sofa and nearly fell asleep.
All of a sudden, the (gleas)**phone** rang and completely startled me.

14 HF-HP There was a height restriction to get on the rollercoaster.
HF-LP At the funfair, the group of friends raced towards the Ghost Train.
Some of the kids were too (cleck)**short** to go on the ride.

15 HF-HP Sarah had saved money to have veneers fitted at the dentist.
HF-LP Sarah paid for a hospital where she would have private clinicians.
When they were finished, her (frilt)**teeth** looked fabulous.

16 HF-HP The couple finally got pregnant after trying for months.
HF-LP The teenagers tried for weeks to get into the boarded up building.
They were extremely (degip)**happy** when they eventually succeeded.

17 HF-HP The vet examined the critically ill dog in her surgery.
HF-LP Meg approached the bird that was caught in the net.
She tried to prevent it from (agorp)**dying** but it was too late.

18 HF-HP The farmer worked hard all day in his fields.
HF-LP Ian’s first day at work had been just as eventful as he’d hoped.
He was extremely (brean)**tired** when he came home.

19 HF-HP The toddler held onto the furniture to keep himself upright.
HF-LP The old man’s health had begun to deteriorate at an increasing pace.
On his own, he was unable to (choul)**stand** without falling down.

20 HF-HP Marvin had to go to the shops to buy a new ink cartridge.
HF-LP Marvin couldn’t meet up with his friends to brag about his trip.
At present, he was unable to (gowel)**print** of his colour photos.

21 HF-HP Dan was traumatised by seeing the mutilated body as a child.
HF-LP Dan’s phobia had been caused by a snake biting him as a child.
He could never get rid of the (serpo)**image** from his mind's eye.

22 HF-HP The famous soprano received a standing ovation from the audience.
HF-LP The memorial ceremony for Michael Jackson was beautiful.
People threw flowers at the (chups)**stage** to show their admiration.

23 HF-HP Sheena had to shop for many things in many different stores.
HF-LP The teacher prepared for the classes she would have next week.
She made up several (frade)**lists** so that she remembered everything.

24 HF-HP The painters were told not to damage any of the furniture.
HF-LP The artists were commissioned to design a mural for the lobby.
Before they began, they had to (acrom)**cover** everything with sheets.

25 HF-HP Parents must ensure appropriate supervision in their absence.
HF-LP The cooker is one appliance where safety is a priority.
It is important never to leave a (stoft)**child** unattended in the kitchen.

26 HF-HP The night after her day at the zoo, Natalie fell into a deep sleep.
HF-LP Natalie's parents told her they would be living in Africa next year.
She had a very unusual (hower)**dream** about being chased by a chimp.

27 HF-HP Susan was bored in the lecture and time passed slowly.
HF-LP Susan's guidance counsellor had many curious objects on his shelves.
She kept looking at the (shest)**clock** and counted down the minutes.

28 HF-HP The park keepers took good care of the lawns.
HF-LP The hotel owners were expecting external evaluators to arrive any day.
They made sure that the (pance)**grass** was cut every day.

29 HF-HP It was a cold day and Barbara had forgotten her gloves.
HF-LP Last night, Barbara went for a walk to think about moving house.
She decided to keep her (trule)**hands** in her pockets for warmth.

30 HF-HP Seth could easily carry six plastic chairs at a time.
HF-LP The minister invested in durable banquet tables for the church hall.
They were incredibly (fryth)**light** and could be stacked together.

31 HF-HP The class of 300 students had arrived to sit their final exam.
HF-LP They had been revising hard for the British Citizenship exam.
Everybody in the (loth)**hall** was determined to do well in the test.

32 HF-HP Robbie and his dad were getting ready to play catch.
HF-LP Edward and Gillian had been searching for over an hour.
They finally found the (teth)**ball** in the back corner of the front closet.

33 HF-HP The librarian was disgraced at the damage to the returned book.
HF-LP The student unwillingly handed in his jotter to be graded.
There wasn't a single (gopa)**page** without a tear or smudge.

34 HF-HP Crashing the car into the living room left Jim with a large bill.
HF-LP The work was scheduled to take place over the winter months.
Rebuilding the (muth)**wall** would be the most expensive part of the job.

35 HF-HP Many young boys dream of becoming soldiers when they grow up.
HF-LP Kids are often asked what they would like to be when they grow up.
A career with the (narp)**army** seems an exciting adventure to children.

36 HF-HP The man we had hired to replace our slates had a fatal accident.
HF-LP The man's family were consoled that he had not suffered in agony.
The fall from the (mant)**roof** had killed him instantly on impact.

37 HF-HP We decided to take the children to play on the swings on Sunday.
HF-LP The children waited eagerly in the car for their parents.
A trip to the (yeat)**park** is a lovely treat when the sun is shining.

38 HF-HP Many people are opting to leave cities for a quieter life.
HF-LP Not everybody lives near their place of work.
They move to (creef)**rural** areas and commute to work instead

39 HF-HP George found a marquee to host his son's wedding reception.
HF-LP George was told the mountain top cabin provided spectacular outlooks.
It was ideal for the (nouch)**even**t so he hired it immediately.

40 HF-HP The Ministry of Defence discovered a spy in their operation.
HF-LP The tour guide arranged and led art history trips to Eastern Europe.
He was a Russian (nysol)**agent** who was relaying details to Moscow.

41 HF-HP The knife was blunt and Nigel was struggling to cut the turkey.
HF-LP Nigel began to shape the marble sculpture with a small chisel.
He asked his wife for a (clurg)**shar**p one and continued to carve.

42 HF-HP Terry went to the new gardening centre.
HF-LP Terry had never been to Stirling's out-of-town shopping centre.
He bought a rare (ghock)**plant** for his garden.

43 HF-HP The young couple were wanting to furnish their new dining room.
HF-LP The older couple visited the store on the high street.
They selected a (futir)**table** that was exactly what they wanted.

44 HF-HP People sometimes tried to drown themselves by jumping off the bridge.
HF-LP The dreadful car accident happened at 6am when the roads were quiet.
A boy was pulled from the (cusic)**river** by a passer-by walking their dog.

45 HF-HP The DVD is now the most common form of movie entertainment.
HF-LP Many things are being replaced by more modern innovations.
It seems that the (antom)**video** will soon be a thing of the past.

46 HF-HP Mark's car was damaged by the side-on crash at the junction.
HF-LP Mark was annoyed when he realised how much he would have to spend.
He would need new (bream)**doors** before his car was roadworthy.

47 HF-HP The thugs were arrested and brought to the police station.
HF-LP he owners of the castle knew what to do with unruly peasants.
They put them in the (ratir)**cells** overnight as punishment.

48 HF-HP Doctors warn against excess cholesterol and promote exercise.
HF-LP School age children are taught about the importance of their health.
Looking after the (brock)**heart** is an important task for all age groups.

49 HF-HP Inflation commonly rises by a small percentage every year.
HF-LP The MP asked for additional anti-terrorism measures to be imposed.
The result is an increase in the (quain)**price** of goods that we purchase.

50 HF-HP The lawyers were behind schedule in selecting the jurors.
HF-LP Kyle and Simone were waiting for a call from the company.
They were hoping to begin the (favid)**trial** as quickly as possible.

51 HF-HP The yacht crew were pleased with the favourable strong wind.
HF-LP The new route would help them arrive earlier than planned.
They used it to gain (nysol)**speed** and were able to win the race.

52 HF-HP David increased his vocabulary by reading lots of books.
HF-LP David was glad he had studied moral philosophy at university.
His knowledge of difficult (miate)**word**s was far better than others.

53 HF-HP The joiner hadn't smoothed the edges of the cabinets yet.
HF-LP The carpenter had just finished staining the wooden doors.
They were still quite (sumpt)**rough** and not ready to be varnished.

54 HF-HP Johnny liked his father to read to him before bedtime.
HF-LP Johnny enjoyed his first day at primary school.
There was one particular (chung)**story** he liked about a tiger.

55 HF-HP Only when they had no choice, the cannibals would eat monkey flesh.
HF-LP The aliens had strange customs and eating habits.
They preferred the taste of (frove)**human** flesh over animals.

56 HF-HP I love the feeling of sand under my feet and the sound of waves.
HF-LP I often take my children out on the weekends.
Going to the (trest)**beach** to collect shells is an enjoyable activity.

57 HF-HP Ted was diabetic and had to monitor what he ate.
HF-LP Ted had to monitor his diet carefully for his heart condition.
If he ate too much (rapon)**sugar** he could become unwell.

58 HF-HP Stuart did not want to travel to London by bus or plane.
HF-LP Stuart wanted to see his favourite band in London.
He bought tickets for the (deece)**train** to Waterloo on the Internet.

59 HF-HP Matthew’s younger sister was born several years after him.
HF-LP Her brother, Matthew, was working towards his black belt in karate.
Because he is (uthos)**older** than her, he is protective.

60 HF-HP The patient had been cared for in the hospital for weeks.
HF-LP His long stay was not as tiresome as he had thought it would be.
He had a favourite (searn)**nurse** who looked after him.

61 HF-HP Keith was known to be a terrible but well-meaning local gossip.
HF-LP Aunt Lillian calls all her relatives on a regular basis.
If something happens, the entire (bine)**town** will know within hours.

62 HF-HP Tour Pleasure Cruises were ideal for people who enjoyed diving.
HF-LP The children liked to visit their rich uncle’s coastal villa.
They would jump off the (krid)**boat** into the cool blue water.

63 HF-HP The emergency services were attending to a terrible car crash.
HF-LP The vehicle had been set on fire in the multi-storey car park.
The police closed the (sart)**road** until the wreckage was removed.

64 HF-HP Tom had accidentally walked on a piece of glass at the beach.
HF-LP Tom had an injury but he had to remain in uniform all day long.
He finally examined the cut on his (dech)**foot** and it looked infected.

65 HF-HP Farms are often very valuable to potential property developers.
HF-LP Their new investment would mean they could retire early.
They will use the (trab)**land** to build new homes and make lots of money.

66 HF-HP Our holiday in the Canadian wilderness came to a terrifying end.
HF-LP Our militia took cover in the hills but it did not go as planned.
Our camp was attacked by a (dace)**bear** that had smelled our food.

67 HF-HP Men who are proposing often spend many hours in jeweller's shops.
HF-LP The father gave his son some timely advice.
Choosing the right (semp)**ring** may make the difference in the outcome.

68 HF-HP Much care went into watering the field before the football match.
HF-LP The rival teams were due to play at the refurbished sports complex.
On the day of the game, the (yolet)**pitch** looked better than ever.

69 HF-HP Jean couldn't stay long at her father's because she was running late.
HF-LP Jean had to wait a few minutes before Greg came to the door.
She was only paying him a (pinel)**quick** visit to see if he was ok.

70 HF-HP Their spacious lounge could easily accommodate thirty people.
HF-LP Their lounge required major refurbishment.
It was very (femps)**large** with high ceilings and a fireplace.

71 HF-HP The secretary sliced the tip of her finger on the letter.
HF-LP Harriet's hands smarted and stung as she was doing the recycling.
She hated getting these (joyac)**paper** cuts and swore loudly.

72 HF-HP Daphne's computer wasn't letting her open the application.
HF-LP Daphne followed the instructions down to the very last detail.
She kept getting (amone)**error** messages and called the support line.

73 HF-HP The shepherd had spotted a wolf prowling around his fields.
HF-LP The man heard that a neighbour's goats had been savaged by a predator.
He kept a close eye on his (clury)**sheep** to protect his flock.

74 HF-HP Last year, Frank narrowly missed the bronze and came in fourth.
HF-LP Frank worked furiously to submit his art portfolio on time.
This year, he hoped to finish at least (kloat)**third** in the competition.

75 HF-HP In cities, there are often special bus and taxi lanes.
HF-LP To make our roads less busy, it is advised to walk short journeys.
Sometimes there are also (squim)**cycle** lanes to ease traffic.

76 HF-HP Ms. Hart had the flu and needed her classes to be covered**.**
HF-LP Linda's whole face had become swollen from the infection.
She would be unable to (frint)**teach** for at least a week.

77 HF-HP There had been a terrible crash at the weekend's Grand Prix.
HF-LP An unstable canister had fallen off one of the maintenance lorries.
Oil had leaked onto the (foust)**track** and caused a massive pile-up.

78 HF-HP The Queen has never voted in a General Election.
HF-LP Several politicians will join the anti-war demonstration.
Members of the (miped)**royal** family are not allowed to.

79 HF-HP Mood around the office was glum and the boss needed to take action.
HF-LP The manager of the private nursery had extra funds to spend.
He decided to organise an event for the (chelt)**staff** to boost morale.

80 HF-HP Craig knew the law about carrying illegal weapons in public.
HF-LP A while ago, Craig’s parents had grounded him for a whole month.
He still carried a (timba)**knife** despite the risk of being caught.

81 HF-HP Melanie and Danielle shared the eighty jelly beans evenly.
HF-LP Melanie and Danielle shared the big bag of sweets between them.
Each girl received (brelp)**forty** sweets and ate them greedily.

82 HF-HP During apartheid in South Africa, most races could not vote.
HF-LP In the past, politics was controlled by the elite in society.
Only people who were (stods)**white** could take part in the elections.

83 HF-HP Tiger Woods was angry when he was distracted playing a shot.
HF-LP Fred felt he would have an advantage playing in the home ground.
Apparently, someone in the (sneck)**crowd** cheered as he hit the ball.

84 HF-HP Little Peter was still afraid to cross the road by himself.
HF-LP Tom was rebuilding the motor on his car but was having some problems.
He needed the aid of a competent (elath)**adult** to help him with the task.

85 HF-HP Henry had been injured in a scrum at school.
HF-LP Henry had injured himself during a routine warm-up at the gym.
He was unable to play (siple)**rugby** for several weeks.

86 HF-HP Rob liked the pub’s drink specials but disliked their hygiene.
HF-LP Rob ordered a baguette and a pint from the pub for lunch.
He noticed that his (phrew)**glass** was cracked and told the waitress.

87 HF-HP The pirates located the spot where the treasure was buried.
HF-LP Stan and his gang had taken what wasn’t rightfully theirs.
They opened up the (sload)**chest** and marvelled at the booty inside.

88 HF-HP Jennifer tried a cigarette for the first time and loved it.
HF-LP Jennifer had discovered a new way of talking to strangers.
She started to regularly (varts)**smoke** when she went out.

89 HF-HP Marcus almost hurt himself badly at the gym lifting weights.
HF-LP The delivery of birdbaths for Marcus’ garden shop arrived.
He had picked ones that were too (drisp)**heavy** for him to lift.

90 HF-HP Jared had forgotten to bring CDs on his long car journey.
HF-LP Jared knew he should revise for his test which was just next week.
Instead, he turned the (anter)**radio** on and tuned in a rock station.

91 HF-HP The teenager had broken out in terrible acne.
HF-LP His back had become a bit swollen and ached a little.
There was one particular (wyal)**spot** that needed to be squeezed.

92 HF-HP New parents often find it impossible to get a full night's sleep.
HF-LP We hadn’t realised how thin the walls were at the holiday resort.
The noise of a (felp)**baby** crying in the next room makes sleeping hard.

93	HF-HP	The cyclist was overcome by the steepness of the incline.
	HF-LP	The disabled man was having trouble coping with everyday life.
		He couldn’t get up the (keft)**hill** as his as legs were too painful.
94	HF-HP	Our house cat loves to pounce on unsuspecting sparrows in the garden.
	HF-LP	The farmer’s cat was getting old but was still a great hunter.
		Going out to catch a (foat)**bird** was a treat the cat looked forward to.
95	HF-HP	Gran had become very senile and couldn't look after herself.
	HF-LP	Our new budgie was in danger of being eaten by our cat.
		We decided to put her in a (leam)**home** where she would be less vulnerable.
96	HF-HP	The drunken man was taken to hospital after being head-butted.
	HF-LP	Darren’s parents took him to A&E and he was finally admitted.
		The doctor repaired his broken (vern)**nose** and prescribed painkillers.
97	HF-HP	Our day of kite flying was cut short when it became stuck.
	HF-LP	The wet blanket from the boat was on the line drying in the breeze.
		It got caught in a (hain)**tree** and we had to climb up and retrieve it.
98	HF-HP	The burglar wore soft shoes to avoid being heard.
	HF-LP	Paul often helped himself to sweets once his mum left the kitchen.
		He was always very (praik)**quiet** and had never been caught.
99	HF-HP	Harry hated missing the beginning of films at the cinema.
	HF-LP	Harry was always running late for every occasion.
		Once again, he had missed the (whurl)**start** of the movie and was annoyed.
100	HF-HP	At the end of season sale, prices were much reduced.
	HF-LP	She was surprised at the quality of the items in the vintage store.
		The clothes were (slony)cheap but still of very high quality.
101	HF-HP	Shoppers were excited about the clothes shop being built in town.
	HF-LP	The odd design of the exterior of the business drew much attention.
		When it was opened, the new (cleam)**store** attracted many customers.
102	HF-HP	Every morning, Jeff would walk past the baker's shop.
	HF-LP	Jeff liked to visit the summer market in the town centre.
		He enjoyed the smell of (faunt)**bread** and frequently bought a loaf.
103	HF-HP	During the War, German submarines targeted supply convoys.
	HF-LP	During the war, the enemy forces caused distress and destruction.
		They would attack the (aboge)**ships** that carried weapons and food.
104	HF-HP	We approached the prison camp where our friend was being held captive.
	HF-LP	It was getting dark and we made our final approach with great care.
		In the distance we saw a (peast)**guard** pacing back and forth with a rifle.
105	HF-HP	Tony wanted to win in this year's maths competition.
	HF-LP	Tony hoped that he would soon be able to call himself Mr. Universe.
		He wanted the (garem)**prize** money more than anything else.
106	HF-HP	The young couple were eager to get onto the property market.
	HF-LP	The average teenager reaches many milestones after leaving school.
		Deciding to buy a (drern)**house** is one of life's big commitments.

107 HF-HP The gang leader had been gunned down as he left his house.
HF-LP The police spokesman read out a statement about the kidnapping.
It was done by members of a (wrach)**rival** gang in a revenge attack.

108 HF-HP I could feel something in my shoe which dug into my heel.
HF-LP The child ran home with something he had taken from the garden.
It was a small (chran)**stone** which had come from the gravel path.

109 HF-HP I didn't realise how hot my curry was as I ate a big spoonful.
HF-LP After my public talk, my nerves were completely frazzled.
The peculiar sensation in my (smoll)**mouth** only went away with a few beers.

110 HF-HP The assistant at the bank spilled ink all down his front.
HF-LP Tom gorged himself on the freshly picked blueberries.
This left a stain on his (cload)**shirt** and he was angry at himself.

111 HF-HP Jack's aunt was supposed to pick him up after school.
HF-LP After riding the rollercoaster, Jack looked for his brother.
Instead, it was his (vands)**uncle** who was waiting for him.

112 HF-HP The plumber couldn't mend the boiler until next week.
HF-LP Paul decided to build a cabin to house his hunting equipment.
He had to order the (quafe)**parts** he needed from a specialist shop.

113 HF-HP The Archbishop of Canterbury presided over the King's coronation.
HF-LP At the end of the cartoon, the wizard turned the frog back into a man.
When he placed the (unrew)**crown** on his head, the ceremony was complete.

114 HF-HP Liz and her friends polished off all the food in her flat.
HF-LP Liz was friends with the workmen who were remodelling her kitchen.
The refrigerator was (rejip)**empty** after they left.

115 HF-HP The army designed new camouflage to be used in forests.
HF-LP The novelty birthday cake was shaped like a mountain range.
It was mostly dark (porve)**green** but had patches of black and brown.

116 HF-HP Maria's only son was graduating today from Oxford.
HF-LP Maria was meeting with the Headmaster when he summoned in her son.
As she watched him, she felt so (queck)**proud** of his achievements.

117 HF-HP Mary's young son gave her a kick as she washed the dishes.
HF-LP Mary suddenly noticed a fox staring at her from her back window.
She was so surprised, she dropped a (ghuba)**plate** and it smashed.

118 HF-HP At school, Miss Jones told only the boys to leave early.
HF-LP The Mother Superior noticed that all the communion wine was gone.
She wanted to talk to the (quate)**girls** about the incident.

119 HF-HP Keith liked to listen to Mozart, the Beatles, and techno.
HF-LP The Mother Superior noticed that all the communion wine was gone.
He liked all kinds of (siver)**music** with no particular preference.

120 HF-HP Adam’s behaviour at school was getting out of control.
HF-LP Keith was undecided which genre he should choose for his essay.
He kept disrupting the (stoin)**class** and got into fights during break times.

121 LF-HP I thought that a nice set of antlers would look good in my study.
LF-LP I had finally thought of the perfect gift for my father-in-law.
I went out to hunt for a (choy)**stag** that would make a good trophy buck.

122 LF-HP The doctor had prescribed me medicine for my angina.
LF-LP The doctor said I was run down and should take some time off.
He told me to take a (yoth)**pill** with water after my evening meal.

123 LF-HP The safari hunter startled his prey when he fired a shot at it.
LF-LP The vet went to get medicine for the sick animal at the park.
He hurried to his (pory)**jeep** and sped off after the lion.

124 LF-HP The hot sun shone on the pool bar as I ordered a gin and tonic.
LF-LP The bowl of punch has lots of fruit in it and I ordered a glass.
It came with a wedge of (bown)**lime** and loads of ice.

125 LF-HP I have a friend who is always stuffing his face with chewing gum.
LF-LP I have a friend who is always stuffing his face with sweets.
There is a strong smell of (neek)**mint** whenever John opens his mouth.

126 LF-HP Planners are building a new public transport system in Edinburgh.
LF-LP With congestion, councils must redesign towns to aid commuters.
Constructing a (bewe)**tram** system is a useful way of easing traffic.

127 LF-HP Ian's wound was almost completely healed so he decided to pick it.
LF-LP Ian couldn't wait to see how it looked underneath.
He carefully pulled the (rext)**scab** and felt satisfied when it came off.

128 LF-HP Andrea constantly suffered from severe eczema.
LF-LP Andrea had an appointment to see a doctor at the clinic.
She was always (shalp)**itchy** and constantly scratched her arms.

129 LF-HP Maude added two brown sugars to her cappuccino.
LF-LP Maude ordered ice cubes for her milkshake at McDonald's.
She put her spoon through the (drelk)**froth** and stirred them in.

130 LF-HP Swimmers in the sea spotted the distinctive fin and called out.
LF-LP The lifeguards raised the alarm at the tourist resort.
Everyone hurried from the (clouf**)shark** as soon as the warning went out.

131 LF-HP Brad and Phoebe bought a large box of popcorn at the movies.
LF-LP Brad and Phoebe both ordered the fish curry at the restaurant.
However, it was too (metig)**salty** so they didn't eat much of it.

132 LF-HP Nadia had been practising her tennis stroke for six hours.
LF-LP Nadia had been practising most of the afternoon.
She now had a pain in her (ottun)**elbow** and went to get an ice pack.

133 LF-HP David's toilet bag was searched for sharp objects at the airport.
LF-LP David was travelling through Europe and decided to go to a festival.
Security confiscated his (minan)**razor** and he'd have to buy a new one.

134 LF-HP Jamie loved basketball but he was very short for his age.
LF-LP Ian enrolled in the Dutch school when his parents moved to Amsterdam.
In gym class, he felt like a (frout)**dwarf** next to his classmates.

135 LF-HP My parents met in the Seventies and loved to go out dancing.
LF-LP The couple spent a lot of their free time together.
Every weekend they would go to a (bowen)**disco** and dance the night away.

136 LF-HP The shopkeeper suspiciously eyed the girl in the hooded top.
LF-LP George decided that the next time she came in, he would do something.
He knew she was a (klead)**thief** and hoped to catch her red-handed.

137 LF-HP After his morning jog, Gregor was happy to take a long, hot shower.
LF-LP In the morning, Gregor forgot where he had last put his contacts.
When he stepped out, he reached for his (frest)**towel** but it wasn't there.

138 LF-HP If Cinderella was to go to the ball, she would need a miracle.
LF-LP Nina joked that she required a makeover to look good for the party.
She needed the help of her (toing)**fairy** godmother to make her over.

139 LF-HP Maria carried a donor card in case she was in an accident.
LF-LP Maria correctly filled in her medical details on the hospital forms.
Doctors could use any (amper)**organ** in the event of her death.

140 LF-HP The sun's heat can be used as a renewable source of energy.
LF-LP Rural communities are often encouraged to upgrade their houses.
People can use (raken)**solar** panels on their roofs for power.

141 LF-HP Zoe had a habit of forgetting to check food in the oven.
LF-LP Zoe found life difficult when she moved out of her parents' house.
Most of the time she (frask)**burnt** her meals and had to start over.

142 LF-HP Pierre had entertained kids at the circus for fifty years.
LF-LP Pierre loved to make people laugh and had made a career of it.
He had enjoyed being a (skane)**clown** but it was time to retire.

143 LF-HP The amateur cyclists found it difficult to ascend the hill.
LF-LP The boy scouts made their way through the forest in the downpour.
The path was (whang)**steep** so they took several breaks.

144 LF-HP When Geoffrey got a nosebleed, Dawn nearly keeled over.
LF-LP When her boyfriend got into a fight, Dawn became quite upset.
We thought she was going to (broud)**faint** after seeing all the blood.

145 LF-HP At the ceilidh, Steven vigorously spun Emma round and round.
LF-LP Steven dared Emma to run across the bridge stretching above the river.
This made her very (brimp)**dizzy** but she still had a good time.

146 LF-HP Farmer Joe dusted his fields to combat the insect infestation.
LF-LP Steve was not impressed by the antics of the local hooligans.
They had started to attack his (snage)**crops** and he had to act quickly.

147 LF-HP Maria's boyfriend had been drinking and he slapped her in the face.
LF-LP Maria saw the football coming towards her and tried to dodge it.
She could feel her left (stoud)**cheek** redden and started to cry.

148 LF-HP The window cleaner always carried a supply of hot tea.
LF-LP Luke was always drinking green tea at work.
He kept it in a (thund)**flask** that he filled up every morning.

149 LF-HP Ponies and horses are not suited to travelling across deserts.
LF-LP Some animals have always been used for transportation.
The best animal for this is the (seach)**camel** as it rarely needs water.

150 LF-HP Emily had never seen such an enormous bowl of ice cream.
LF-LP Emily was famished and didn't care who was watching.
She excitedly grabbed a (equim)**spoon** and began to stuff herself.

151 LF-HP Egyptian pharaohs were often given extravagant burial chambers.
LF-LP The history lecturer talked about leaders from ancient civilisations.
Building an ornate (fank)**tomb** was customary to celebrate each ruler.

152 LF-HP The aid worker visited the poorest and dirtiest part of the city.
LF-LP The private sector built more housing for the expanding population.
Many people lived in the (edor)**slum** and their conditions were terrible.

153 LF-HP I stood at the bottom of the sandy slope and prepared to run.
LF-LP When I returned home, I felt like going for an outdoor sprint.
I ran to the top of the (larn)**dune** and felt a burning in my calves.

154 LF-HP Experts say using live bait will improve your angling success.
LF-LP On our camping trip, we avidly followed advice from specialist books.
Placing a fresh (rece)**worm** on the end of the line should attract a fish.

155 LF-HP Panic struck when I ran out of bread making the kids' sandwiches.
LF-LP I was making Lea's lunch when I remembered her friend was coming too.
I hurried out to get a (bant)**loaf** so I could finish making the lunches.

156 LF-HP The witch cast a spell to transform the prince into another form.
LF-LP In the fantasy novel, the android Kyra attacked commander Devlin.
Turning him into a (tump)**frog** left her free to take over the kingdom.

157 LF-HP Everyone knows what happens at the end of Little Red Riding Hood.
LF-LP The thriller's turning point is when some animals are found dead.
The woodsman kills the (vutt)**wolf** with an axe and the story ends well.

158 LF-HP Waste had overflowed onto the road and the smell was awful.
LF-LP The minor tremor caused much damage to our neighbourhood.
Luckily, workers repaired the broken (ranom)**sewer** within two hours.

159 LF-HP The young couple were shopping for new living room furniture.
LF-LP The young couple had extra money from their wedding to spend.
They purchased a new (meast)**couch** that had very soft upholstery.

160 LF-HP Sebastien's holiday to Cuba had been the trip of a lifetime.
LF-LP The old gentleman had become accustomed to fine wine and dining.
He often enjoyed a (zepia)**cigar** after dinner.

161 LF-HP Hounds used for hunting are trained in special kennels.
LF-LP Some farmers train their dogs to do special tasks.
They are taught to chase (bence)**foxes** out of their burrows.

162 LF-HP Fiona always had two cups of strong coffee to wake her up.
LF-LP Fiona used her fruit juicer relentlessly every morning.
This made her feel more (skack)**alert** and ready to tackle the day.

163 LF-HP Will heard that a large mist was forecasted to move inland that night.
LF-LP Will was driving fast to meet his friends for that night’s game.
It suddenly became (boppa)**foggy** as Will drove and he had to slow down.

164 LF-HP The couple took the cruise ship across the Pacific to Hawaii.
LF-LP The day was hot and the couple made their way upstairs.
The breeze from the (varew)**ocean** kept them cool on deck.

165 LF-HP After paying for her groceries, the cashier gave Lisa her change.
LF-LP Lisa checked that she had her key before leaving her flat.
She put it in her (yamor)**purse** and carried the bags to her car.

166 LF-HP The old professor dressed as a stereotypical academic.
LF-LP When she met her blind date, she chuckled at his fashion sense.
His jacket was (drack)**tweed** and had patches on the elbows.

167 LF-HP Ryan's friends influenced him to drink at the school disco.
LF-LP Ryan parents wanted to know why he quit his new job.
It was because of pressure from his (grize)**peers** that he did it.

168 LF-HP At their local pub, the office workers all ordered gin and tonics.
LF-LP The restaurant staff were extremely busy with two birthday parties.
The manager started to cut up a (haver)**lemon** for the drinks.

169 LF-HP Lorna had gone on a five mile run in the midday sun.
LF-LP Lorna had managed to build up her flat-pack wardrobe by herself.
You could see the (canch)**sweat** running down her face by the end.

170 LF-HP The toddler ran up to his parents with his face covered in snot.
LF-LP The toddler was distraught about his broken toy.
His mum leaned over and (majit)**wiped** his nose with a soft tissue.

171 LF-HP The new store carried the latest range of denim clothing.
LF-LP Kate had been chatting online to Jim for many weeks on a dating site.
Kate treated herself to expensive (gruce)**jeans** for her big date.

172 LF-HP Emma prayed her parents would get her a cute pet this Christmas.
LF-LP Emma asked for just one present when she wrote her letter to Santa.
Her heart leapt when she saw a beautiful (jogyr)**puppy** sitting patiently.

173 LF-HP Sidney had tried a new shampoo for his terrible dandruff.
LF-LP Sidney tried the oil the pharmacist had recommended to him.
He massaged it into his (noady)**scalp** before rinsing it out well.

174 LF-HP Gavin was diving for oysters and got lucky when he opened one.
LF-LP Gavin was happy to return from his snorkelling holiday.
He had found a large (queck)**pearl** and gave it to his wife.

175 LF-HP Heroin addicts often tie a belt tightly around their arms.
LF-LP Doctors often apply pressure to certain regions of the body.
This allows them to locate some (naron)**veins** that they inject into.

176 LF-HP The priest smiled as the bride and groom exchanged their vows.
LF-LP Tim couldn’t believe that Jean was on time for once.
They stood at the (ethem)**altar** and looked deeply into each other's eyes.

177 LF-HP The neighbour's Alsatian kept coming into Valerie's garden.
LF-LP Valerie recently took in a stray dog from the shelter.
She got her son to build a (krean)**fence** to keep it away from her roses.

178 LF-HP Peter liked extra cheese and mushrooms as toppings.
LF-LP Flying first class to America, Peter was asked if he wanted anything.
He ordered a large (grice)**pizza** with a side of wedges.

179 LF-HP Old Mrs. Greeble was warty, haggard, and had a fearsome black cat.
LF-LP Miss Dearborn lived at number 31 in Alder Crescent.
The older kids said that she was a (redal)**witch** to scare the younger ones.

180 LF-HP The teacher scrawled the sentences onto the blackboard.
LF-LP The teacher turned back to the lesson after scolding the kids.
The noise of the (stutt)**chalk** sent shivers up everyone's spine.

181 LF-HP Theseus used string to guide his way through the minotaur's lair.
LF-LP Brian loved to explore but he always took precautions.
Getting back out the (weam)**maze** would be easier if he left a trail.

182 LF-HP Our Pacific dive gave us the opportunity to see natural coral.
LF-LP The old shipwreck was a favourite setting for several reasons.
Swimming around the (surk)**reef** was a wonderful way to see exotic fish.

183 LF-HP My sister used to play with girly toys and had a clear favourite.
LF-LP My grandmother fondly recounted stories about when she was young.
As a child, there was one (neft)**doll** she refused to be without.

184 LF-HP On holiday in America, we spent Saturday shopping for clothes.
LF-LP My friends and I decided to spend Monday together.
We drove to the (soth)**mall** on the outskirts of town and had a good day.

185 LF-HP The good witch used magic to make children's wishes come true.
LF-LP The children loved when the teacher enacted the stories.
She would take out a (mesk)**wand** and wave it about.

186 LF-HP Traditional wine producers do not like screw caps on wine bottles.
LF-LP Traditionalists prefer natural products for wine production.
They think the best substance to use is (nand)**cork** instead of plastic.

187 LF-HP I needed to roll a double six with my last throw to win the game.
LF-LP I am rather competitive and always want to beat my opponent!
I picked up the (torm)**dice** and said a silent prayer as I threw them.

188 LF-HP The Eskimo family hunted for weeks to prepare for the Arctic winter.
LF-LP The family hunted for weeks to prepare for the approaching winter.
They stocked their (spizz)**igloo** with enough food to last months.

189 LF-HP The nuclear plant had contaminated the area with noxious waste.
LF-LP The surveyors spent several weeks in an area outside of town.
The land was very (bisco)**toxic** and could not be used for decades.

190 LF-HP The band included trumpets, trombones, French horns and tubas.
LF-LP The band included several members of the same family.
It was the most famous (hiver)**brass** ensemble in Canada.

191 LF-HP The boys got into a fist fight in the playground.
LF-LP The two squabbling boys were finally left unsupervised on the bus.
They began to furiously (yexel)**punch** each other in the face.

192 LF-HP After rhumba and tango classes, the pair wanted to try something new.
LF-LP The friends were deciding which evening class they should take.
They thought that the (antis)**salsa** class would be the most fun.

193 LF-HP In music class, Ricky discovered that he had natural rhythm.
LF-LP Tim was new to the class and his teacher needed to assess his level.
His teacher sat him at the (hover)**drums** and told him to play away.

194 LF-HP After many washes, Karl's shirt had lost most of its colour.
LF-LP The dog's blanket in the back of Karl's truck had seen better days.
It was so badly (tefal)**faded** that he needed to buy a replacement.

195 LF-HP Dr. Adams was still drunk when he was due to start work.
LF-LP Dr. Adams made his way into work, despite feeling under the weather.
He would need to (anlom)**sober** up quickly or he would be sacked.

196 LF-HP Robert was polishing his shoes before his big job interview.
LF-LP Robert was dismayed to see that the ornaments had gathered dust.
He wanted them to be (whemp)**shiny** enough to see his face in them.

197 LF-HP Before the new school year, all the furniture was replaced.
LF-LP The council awarded a small grant to the high school's library.
Pupils would have new (boafe)**desks** that were free from graffiti.

198 LF-HP Alison's eyes were watering as she chopped the vegetables.
LF-LP Alison normally steamed her food but today she was in a hurry.
She added the (vazir)**onion** and peppers into the oil in the pan.

199 LF-HP The pregnant girl's family had a history of multiple births.
LF-LP The girl had expected her appointment to be straightforward.
The nurse told her she had (bowe)**twins** when she went for her scan.

200 LF-HP The witness did not get a good look at the mugger.
LF-LP The police wanted to use her account to narrow down the suspects.
Her description was (niper**)vague** and not very helpful.

201 LF-HP Jill's friends were drinking red wine all night in her flat.
LF-LP Jill shuddered as the rain battered against her doors and windows.
In the morning, she noticed an enormous (whone**)stain** on the carpet.

202 LF-HP The triumphant King arranged a sumptuous and lavish banquet.
LF-LP Sir Blakewell smiled as he recalled the events from that evening.
It was a delightful (druch)**feast** which was heartily devoured by all.

203 LF-HP The anthropologist studied the ways of different African peoples.
LF-LP Last year, Ray travelled around Africa with his camcorder.
Each month he filmed a different (durko)**tribe** to record their customs.

204 LF-HP Karen had jumped and landed awkwardly while ice skating.
LF-LP Someone spiked the volleyball directly at Karen in gym class.
She badly hurt her (velts)**ankle** and would need an x-ray.

205 LF-HP I couldn't stop sneezing as I cleaned out the storage room.
LF-LP I had offered to tidy up the sitting room in my student flat.
Everything was (boulp)**dusty** and got up my nose as I worked.

206 LF-HP Poachers still illegally hunt elephants for their tusks.
LF-LP It is illegal to hunt endangered species but some disobey this.
It is possible to buy (crung)**ivory** items on the black market.

207 LF-HP The bottle of coke had been opened a few days ago.
LF-LP Liam's friend handed him a drink at the party.
Liam drank some, but it was not (dromp)**fizzy** and tasted bad.

208 LF-HP The letter Lucas had posted was returned to him.
LF-LP Lucas had hastily written a cheque to his window cleaner.
He had forgotten to put a (ching)**stamp** on it before posting it.

209 LF-HP Maintaining a healthy digestive system requires roughage.
LF-LP Some breakfast cereals are actually not very healthy.
Foods that are high in (tober)**fibre** are recommended by experts.

210 LF-HP In the Disney film, Belle falls in love with the castle's monster.
LF-LP The fairytale ends happily as would be expected.
The village beauty and the (tarch)**beast** live happily ever after.

211 LF-HP An arena was built in London to mark the new millennium.
LF-LP The citizens were not pleased to see what their taxes had paid for.
Everyone agreed that the (brin)**dome** was a terrible waste of money.

212 LF-HP The elephant family prepared to travel to their destination.
LF-LP The tourists watched the animals departing through their binoculars.
The members of the (lunk)**herd** set off, trunk to tail, on their journey.

213 LF-HP Al's head was itching terribly and one look revealed the problem.
LF-LP Al begged his mum to let him stay home from school.

He had caught (fren)**lice** from one of the other children at school.

214 LF-HP Pouring salt around the creature that eats my lettuces is pleasing.
LF-LP The bait is placed in the container and set on the ground at night.
This traps the (whep)**slug** and I take perverse pleasure in its demise.

215 LF-HP We tearfully said our goodbyes to Tommy as he boarded the ferry.
LF-LP We said our goodbyes to Tommy as he was leaving the neighbourhood.
We stood and waved at the end of the (gron)**pier** as the ship left port.

216 LF-HP Betty only needed an egg white to make her meringue nest.
LF-LP Betty disliked wasting food when she was baking.
Later, she used the (geth)**yolk** to make a separate dish.

217 LF-HP It is important to protect clothes from being eaten by insects.
LF-LP You have to be careful not to leave clothes outside overnight.
A determined and hungry (zelt)**moth** could ruin an entire wardrobe.

218 LF-HP The school football team persevered despite the wet and dirty pitch.
LF-LP The school football team had travelled a long way to play the match.
The children returned home (selig)**muddy** but happy to have won.

219 LF-HP The gypsies travelled along the canal in the middle of the night.
LF-LP Gypsies were illegally trying to get into the neighbouring country.
They hid in the cargo of a slow moving (demys)**barge** afraid of discovery.

220 LF-HP Flo couldn't eat the sticky toffee because of her dentures.
LF-LP When her young grandson gave her a sweet, Flo eagerly ate it.
It was far too (stimp)c**hewy** and got stuck to her false teeth.

221 LF-HP Sean suffered from the symptoms of Parkinson's Disease.
LF-LP Sean had been lifting weights all morning and he could see the effects.
His arms were (chelp)**shaky** and his family was worried.

222 LF-HP The grey squirrel was foraging at the foot of the oak tree.
LF-LP Sammy often hid things in the back garden for safekeeping.
He recovered the (narem)**acorn** that he had buried last winter.

223 LF-HP The record company wanted Tara to record some new songs.
LF-LP Tara’s company were putting pressure on her to work harder.
They wanted her next (otlar)**album** to come out before Christmas.

224 LF-HP Rory was going to dig all day in the potato fields.
LF-LP Rory’s wife made him sandwiches for his long day ahead.
He picked up his (equir)**spade** and headed off to work.

225 LF-HP Luke's first job was working at the supermarket.
LF-LP Luke’s first job wasn’t exciting but at least he would earn money.
His responsibility was to (clomb)**stack** the shelves.

226 LF-HP The driver hadn’t seen the trench that had been dug to drain water.
LF-LP The driver sped past the countryside house carelessly.
The car crashed into the (fedul)**ditch** and had to be written off.

227 LF-HP Leon was unhappy with the tough bread he got with his soup.
LF-LP Leon was unhappy with the coffee he was served in a UK cafe.
He complained that it was (cluts)**stale** and the waitress apologised.

228 LF-HP Tina's mother was baking in the kitchen.
LF-LP Tina was preparing a picnic for their family outing.
She made lots of (rebra)**cakes** for the whole family to enjoy.

229 LF-HP The cause of death was a hammer blow to the head.
LF-LP The results from the investigation were recorded on the certificate.
The damage to the victim's (cleth)**skull** was quite sickening.

230 LF-HP Everyone was excited about going to see the big cats at the zoo.
LF-LP The teachers had organised a trip to Blair Drummond’s adventure park.
The children wanted to see (doric)**lions** and tigers most of all.

231 LF-HP Albert thought he looked good with his new facial hair.
LF-LP Albert thought his new look made him look just like David Beckham.
His friends disagreed and thought his (trunt)**beard** looked awful.

232 LF-HP Eve's cat had begun to scratch her new furniture.
LF-LP Eve was in charge of the injured leopard at the zoo.
She would need to get its (stoun)**claws** cut to prevent further damage.

233 LF-HP The music teacher hired removal men when he moved house.
LF-LP Simon hired removal men when he moved house.
He couldn't move his (gream)**piano** alone because it was too heavy.

234 LF-HP When I visited Paris, I tried a well-known French delicacy.
LF-LP I went to a restaurant and ordered something I had never tried.
I scooped out the inside of a (crink)**snail** and swallowed it whole.

235 LF-HP The Big Ranch restaurant's speciality was high quality beef.
LF-LP He was known to have a healthy appetite.
Bill ordered a huge (choul)**steak** and a pitcher of beer.

236 LF-HP Tania first prepared the tomatoes, cucumber and lettuce.
LF-LP Tania was surprised at how quickly her guests devoured the appetisers.
She finished making the (nitch)**salad** with oil and vinegar dressing.

237 LF-HP The child couldn't sleep after watching the monster movie.
LF-LP The young child had been to the fun fair for the first time.
It had been really (wrimp)**scary** and she was afraid to be alone.

238 LF-HP The children were confused by a horse with stripes on its side.
LF-LP The children were adamant that they should ride on the horse.
I explained that it was actually a (miden)**zebra** and they seemed content.

239 LF-HP It was a lovely summer's day until the sun went away.
LF-LP The children were adamant that they should ride on the horse.
It disappeared behind a (stest)**cloud** and it became colder.

240 LF-HP Frank’s wife died in giving birth to their son.
LF-LP Frank still missed not having his mother around.
He would visit her (pream)**grave** on Sundays.

Appendix F

Target word characteristics – Chapter 5 (Experiment 4)

HF targets (items 1-30)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
1	face	4	372.51	1	3	1	12	28	231.00	570.75	599.00	0.05	0.40
2	hair	4	147.74	1	3	1	5	38	165.00	595.00	583.00	0.18	0.55
3	lake	4	42.89	1	3	1	22	40	280.00	639.75	585.00	0.00	0.60
4	lips	4	55.21	1	4	2	14	23	190.00	620.00	676.19	0.00	0.70
5	card	4	52.83	1	4	1	14	29	245.00	598.00	565.00	0.00	0.85
6	book	4	249.04	1	3	1	15	25	193.00	624.80	609.00	0.35	0.90
7	neck	4	58.94	1	3	1	5	22	223.00	626.67	587.00	0.15	0.95
8	theme	5	41.50	1	3	1	3	13	459.00	286.33	336.00	0.00	0.10
9	fight	5	71.44	1	3	1	9	37	272.50	511.50	455.00	0.00	0.15
10	trend	5	28.04	1	5	1	2	3	511.00	314.33	328.00	0.00	0.20
11	coast	5	49.72	1	4	1	3	20	339.00	586.50	562.00	0.00	0.35
12	horse	5	79.14	1	4	1	10	16	207.00	658.00	613.00	0.05	0.40
13	phone	5	67.53	1	3	1	3	37	248.00	630.33	624.00	0.20	0.50
14	short	5	192.09	1	4	1	7	10	295.00	440.50	351.00	0.05	0.50
15	teeth	5	49.02	1	3	1	1	18	198.00	639.00	618.00	0.00	0.55
16	happy	5	110.08	2	4	1	6	9	258.00	511.00	355.00	0.00	0.60
17	dying	5	29.53	2	4	2	5	9	328.57	423.81	230.00	0.25	0.65
18	tired	5	40.24	1	4	2	9	14	223.00	419.00	304.76	0.45	0.65
19	stand	5	107.46	1	5	1	3	8	247.62	485.50	450.00	0.05	0.65
20	print	5	31.67	1	5	1	3	2	387.00	514.00	404.76	0.00	0.70
21	image	5	80.39	2	4	1	0	2	400.00	309.52	342.86	0.15	0.75
22	stage	5	169.40	1	4	1	7	12	271.43	610.00	561.90	0.05	0.75
23	lists	5	30.82	1	5	2	8	9	293.00	492.00	380.95	0.00	0.80
24	cover	5	112.28	2	4	1	13	8	289.00	443.00	502.00	0.00	0.80
25	child	5	256.06	1	4	1	3	9	217.00	636.33	581.00	0.00	0.85
26	dream	5	48.54	1	4	1	4	5	292.50	481.00	386.00	0.00	0.88
27	clock	5	29.47	1	4	1	8	17	228.00	641.80	591.00	0.00	0.95
28	grass	5	42.41	1	4	1	9	14	189.00	645.50	599.00	0.00	0.95
29	hands	5	197.59	1	5	2	11	7	130.00	622.67	657.14	0.00	0.95
30	light	5	224.99	1	3	1	9	39	295.00	536.00	550.00	0.00	0.95
	Mean	4.77	102.29	1.13	3.83	1.17	7.43	17.43	270.19	537.09	499.59	0.07	0.65
	SD	0.43	86.35	0.35	0.70	0.38	4.96	11.84	84.34	108.54	125.45	0.12	0.25

HF targets (31-60)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
31	hall	4	125.80	1	3	1	19	42	331.00	595.33	565.00	0.15	0.45
32	ball	4	73.60	1	3	1	20	47	137.50	640.67	615.00	0.60	0.60
33	page	4	105.04	1	3	1	12	20	262.50	555.67	571.00	0.05	0.60
34	wall	4	124.28	1	3	1	15	33	176.19	598.00	589.00	0.05	0.70
35	army	4	121.80	1	4	1	2	8	337.50	615.50	543.00	0.20	0.90
36	roof	4	44.78	1	3	1	9	27	519.00	552.00	586.00	0.25	0.90
37	park	4	116.67	1	4	1	15	16	283.75	591.40	579.00	0.10	0.95
38	rural	5	66.24	1	5	1	2	1	513.00	437.50	358.00	0.00	0.10
39	event	5	111.93	1	5	1	1	0	391.00	295.00	379.00	0.00	0.10
40	agent	5	46.16	1	5	2	1	1	466.00	405.00	357.14	0.00	0.20
41	sharp	5	52.78	1	4	1	5	5	230.00	489.67	476.19	0.00	0.35
42	plant	5	86.87	1	5	1	5	7	258.00	600.83	594.00	0.00	0.45
43	table	5	207.33	1	4	1	5	8	219.67	629.00	604.00	0.05	0.50
44	river	5	100.59	1	4	1	13	12	264.33	634.67	585.00	0.00	0.55
45	video	5	64.79	1	5	1	2	0	223.81	600.00	609.52	0.00	0.65
46	doors	5	252.46	1	4	2	4	29	192.67	618.00	585.71	0.00	0.65
47	cells	5	59.13	1	4	2	12	29	483.50	571.50	357.14	0.05	0.65
48	heart	5	146.09	1	4	1	4	17	265.50	652.33	605.00	0.25	0.65
49	price	5	190.71	1	4	1	4	6	330.00	360.00	320.00	0.00	0.65
50	trial	5	68.93	1	5	2	3	1	433.00	516.00	446.00	0.00	0.70
51	speed	5	80.13	1	4	1	5	15	372.00	416.50	247.62	0.05	0.70
52	words	5	190.53	1	4	2	9	18	267.00	422.50	285.71	0.00	0.75
53	rough	5	35.40	1	3	1	10	24	299.00	384.67	452.00	0.00	0.80
54	story	5	141.00	1	5	1	4	5	193.00	491.00	427.00	0.00	0.80
55	human	5	210.97	1	6	1	0	3	369.00	543.00	583.00	0.10	0.80
56	beach	5	40.92	1	3	1	7	26	232.00	659.00	612.00	0.00	0.90
57	sugar	5	37.40	1	4	1	0	1	263.00	616.00	620.00	0.05	0.95
58	train	5	81.71	1	4	1	7	12	321.50	563.50	592.00	0.05	0.95
59	older	5	92.07	1	4	2	5	4	295.00	479.00	348.00	0.00	0.95
60	nurse	5	35.22	1	3	1	3	13	280.60	617.00	588.00	0.10	1.00
	Mean	4.77	103.71	1.00	4.03	1.20	6.77	14.33	307.00	538.34	502.67	0.07	0.66
	SD	0.43	58.63	0.00	0.81	0.41	5.49	12.87	99.65	98.10	118.15	0.12	0.25

HF targets (items 61-90)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
61	town	4	185.08	1	3	1	7	20	310.00	571.50	556.00	0.00	0.25
62	boat	4	53.08	1	3	1	11	39	313.50	586.00	637.00	0.00	0.45
63	road	4	265.81	1	3	1	9	48	206.00	634.50	583.00	0.20	0.65
64	foot	4	73.92	1	3	1	12	12	156.50	619.00	558.00	0.00	0.75
65	land	4	216.39	1	4	1	15	21	304.50	567.00	604.00	0.00	0.80
66	bear	4	58.46	1	3	1	20	44	202.50	620.75	585.00	0.05	0.90
67	ring	4	60.31	1	3	1	13	28	263.00	605.33	593.00	0.00	1.00
68	pitch	5	29.74	1	3	1	9	30	413.00	449.00	423.81	0.00	0.15
69	quick	5	56.04	1	4	1	3	13	272.50	329.33	343.00	0.05	0.20
70	large	5	361.88	1	4	1	5	9	284.00	429.50	280.95	0.00	0.35
71	paper	5	166.04	2	4	1	8	14	257.00	571.00	599.00	0.00	0.40
72	error	5	41.06	2	3	2	1	11	425.00	408.00	373.00	0.00	0.50
73	sheep	5	29.77	1	3	1	7	32	192.33	639.25	622.00	0.05	0.55
74	third	5	220.28	1	3	1	0	18	257.14	430.00	204.76	0.20	0.55
75	cycle	5	34.11	2	4	1	1	8	273.00	611.00	285.71	0.05	0.60
76	teach	5	28.41	1	3	1	5	20	295.00	403.67	353.00	0.00	0.65
77	track	5	64.50	1	4	1	6	12	383.00	504.50	547.00	0.00	0.65
78	royal	5	160.10	1	4	2	1	4	364.00	531.00	295.24	0.00	0.65
79	staff	5	236.47	1	4	1	2	8	410.00	469.00	470.00	0.00	0.65
80	knife	5	27.89	1	3	1	1	12	230.00	648.50	612.00	0.25	0.70
81	forty	5	28.96	2	5	2	4	4	320.00	370.00	214.29	0.00	0.75
82	white	5	255.27	1	4	1	4	11	241.00	608.00	472.00	0.00	0.75
83	crowd	5	47.24	1	4	1	2	11	334.00	551.00	546.00	0.60	0.80
84	adult	5	54.67	2	5	1	0	1	290.00	590.50	492.00	0.00	0.80
85	rugby	5	35.99	2	5	2	0	0	371.43	590.48	480.95	0.05	0.85
86	glass	5	104.12	1	4	1	3	5	239.50	610.00	635.00	0.10	0.90
87	chest	5	39.26	1	4	1	4	18	302.00	572.33	580.00	0.05	0.95
88	smoke	5	39.47	1	4	1	4	6	288.00	627.50	541.00	0.00	0.95
89	heavy	5	98.88	2	4	1	2	9	228.57	495.00	413.00	0.15	0.95
90	radio	5	89.97	3	5	1	3	2	317.00	613.00	615.00	0.05	1.00
	Mean	4.77	105.44	1.30	3.73	1.13	5.40	15.67	291.45	541.85	483.82	0.06	0.67
	SD	0.43	90.97	0.53	0.69	0.35	4.92	12.52	67.28	90.28	133.21	0.12	0.24

HF targets (items 91-120)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
91	spot	4	50.61	1	4	1	8	15	261.00	513.50	512.00	0.15	0.50
92	baby	4	90.06	2	4	1	2	13	164.00	639.00	589.00	0.00	0.55
93	hill	4	77.10	1	3	1	16	49	256.00	607.00	588.00	0.00	0.80
94	bird	4	41.18	1	3	1	5	48	219.50	640.33	602.00	0.15	0.70
95	home	4	524.78	1	3	1	13	25	239.25	602.40	539.00	0.00	0.80
96	nose	4	44.06	1	3	1	14	45	173.00	608.33	628.00	0.05	0.85
97	tree	4	65.52	1	3	1	5	16	174.25	660.60	604.00	0.10	0.90
98	quiet	5	65.21	1	5	1	3	0	221.00	426.00	389.00	0.05	0.15
99	start	5	213.12	1	5	1	6	10	321.50	299.00	166.67	0.00	0.30
100	cheap	5	36.73	1	3	1	2	20	322.00	310.50	180.95	0.05	0.35
101	store	5	46.56	1	4	1	14	11	394.00	498.00	548.00	0.10	0.40
102	bread	5	34.91	1	4	1	6	19	189.00	629.75	622.00	0.15	0.45
103	ships	5	49.07	1	4	2	10	20	228.33	646.00	628.57	0.00	0.50
104	guard	5	32.52	1	4	1	0	18	356.75	549.75	517.00	0.55	0.55
105	prize	5	33.43	1	4	1	3	17	252.00	524.67	474.00	0.05	0.55
106	house	5	512.32	1	3	1	7	20	166.00	646.00	608.00	0.20	0.60
107	rival	5	28.43	2	4	1	0	5	426.19	271.43	252.38	0.00	0.65
108	stone	5	86.52	1	4	1	9	16	199.00	620.75	614.00	0.00	0.65
109	mouth	5	98.52	1	3	1	5	6	152.38	613.00	568.00	0.00	0.65
110	shirt	5	28.72	1	3	1	7	24	257.00	630.67	616.00	0.35	0.70
111	uncle	5	36.74	2	4	1	0	1	192.00	574.00	580.00	0.00	0.75
112	parts	5	127.74	1	5	2	17	12	338.75	333.60	261.90	0.00	0.80
113	crown	5	56.50	1	4	1	8	11	242.50	626.75	586.00	0.15	0.78
114	empty	5	59.77	2	5	1	0	0	267.00	446.50	374.00	0.00	0.85
115	green	5	148.59	1	4	1	5	19	241.50	629.67	460.00	0.00	0.90
116	proud	5	32.87	1	4	1	0	12	296.00	339.00	327.00	0.10	0.95
117	plate	5	42.08	1	4	1	7	14	227.00	587.67	595.00	0.35	0.95
118	girls	5	101.06	1	4	2	2	19	183.00	650.33	619.05	0.00	0.95
119	music	5	158.07	2	6	1	0	3	272.00	549.00	512.00	0.00	1.00
120	class	5	189.09	1	4	1	8	11	240.00	560.00	370.00	0.00	1.00
	Mean	4.77	103.73	1.17	3.90	1.10	6.07	16.63	249.06	541.11	497.72	0.09	0.68
	SD	0.43	122.86	0.38	0.76	0.31	5.00	12.39	68.82	120.36	141.04	0.13	0.22

LF targets (items 121-150)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
121	stag	4	3.28	1	4	1	9	9	460.00	526.50	585.71	0.00	0.20
122	pill	4	6.10	1	3	1	16	53	341.00	628.00	610.00	0.00	0.40
123	jeep	4	4.87	1	3	1	7	19	413.00	635.00	622.00	0.05	0.45
124	lime	4	6.87	1	3	1	17	26	401.50	570.00	590.00	0.25	0.50
125	mint	4	7.20	1	4	1	16	17	298.00	512.50	604.76	0.05	0.75
126	tram	4	4.46	1	4	1	9	12	442.86	290.00	600.00	0.05	0.80
127	scab	4	0.58	1	4	1	6	4	252.38	571.50	559.00	0.00	0.85
128	itchy	5	0.84	2	3	2	0	3	219.05	328.57	357.14	0.00	0.10
129	froth	5	1.49	1	4	1	4	3	428.57	310.00	471.43	0.00	0.20
130	shark	5	3.34	1	4	1	9	13	252.00	642.33	611.00	0.00	0.20
131	salty	5	1.86	2	5	2	2	5	238.10	512.00	490.00	0.00	0.25
132	elbow	5	11.47	2	4	1	0	1	222.50	616.00	607.00	0.00	0.30
133	razor	5	4.52	2	4	2	1	12	400.00	575.00	638.10	0.00	0.40
134	dwarf	5	6.62	1	5	1	1	2	276.00	624.00	514.29	0.00	0.40
135	disco	5	7.49	2	5	1	1	2	361.00	579.00	504.76	0.00	0.50
136	thief	5	8.07	1	3	1	1	10	330.00	543.00	519.00	0.05	0.55
137	towel	5	8.84	1	4	1	6	6	215.00	577.50	671.43	0.00	0.55
138	fairy	5	8.51	2	4	1	4	30	228.50	590.50	433.00	0.00	0.65
139	organ	5	13.89	2	5	1	0	0	356.00	576.00	596.00	0.05	0.70
140	solar	5	14.32	2	4	2	4	17	502.00	435.00	242.86	0.00	0.70
141	burnt	5	12.24	1	4	2	3	15	310.00	463.00	490.48	0.00	0.80
142	clown	5	3.86	1	4	1	3	11	210.00	649.67	627.00	0.25	0.80
143	steep	5	17.22	1	4	1	7	13	341.00	440.50	361.90	0.00	0.85
144	faint	5	18.43	1	4	1	6	10	400.50	450.33	462.00	0.00	0.85
145	dizzy	5	4.06	2	4	1	2	13	268.00	376.00	319.05	0.05	0.90
146	crops	5	17.50	1	5	2	8	12	314.00	517.00	495.24	0.00	0.90
147	cheek	5	20.14	1	3	1	4	33	267.00	600.50	565.00	0.05	0.90
148	flask	5	3.16	1	5	1	2	1	457.00	576.00	595.00	0.20	0.95
149	camel	5	4.22	2	4	1	1	8	312.67	616.33	597.00	0.15	1.00
150	spoon	5	7.84	1	4	1	5	12	165.50	614.67	614.00	0.00	1.00
	Mean	4.77	7.78	1.33	4.00	1.20	5.13	12.40	322.77	531.55	531.80	0.04	0.61
	SD	0.43	5.51	0.48	0.64	0.41	4.73	11.26	88.93	101.77	104.33	0.07	0.27

LF targets (items 151-180)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
151	tomb	4	6.84	1	3	1	5	28	428.50	624.33	573.00	0.10	0.25
152	slum	4	3.24	1	4	1	11	14	507.00	414.00	419.05	0.00	0.30
153	dune	4	1.89	1	3	1	15	44	471.43	450.00	530.00	0.00	0.50
154	worm	4	6.04	1	3	1	10	12	251.00	619.00	611.00	0.05	0.65
155	loaf	4	3.98	1	3	1	4	19	296.00	537.50	557.14	0.15	0.80
156	frog	4	5.08	1	4	1	3	5	229.67	648.75	619.00	0.20	0.90
157	wolf	4	9.07	1	4	1	3	2	316.00	632.00	595.00	0.25	1.00
158	sewer	5	2.13	1	3	2	6	13	447.62	538.00	552.38	0.00	0.15
159	couch	5	5.59	1	3	1	7	12	310.00	572.33	578.00	0.00	0.15
160	cigar	5	5.21	2	5	1	0	1	367.00	634.50	580.00	0.05	0.20
161	foxes	5	4.54	2	6	2	5	4	252.38	671.43	657.14	0.00	0.25
162	alert	5	17.32	2	4	1	1	6	488.00	440.50	400.00	0.00	0.30
163	foggy	5	1.36	2	4	2	6	8	314.29	576.19	390.48	0.00	0.35
164	ocean	5	21.39	2	3	1	0	1	300.00	650.00	593.00	0.05	0.35
165	purse	5	7.13	1	3	1	8	24	256.00	604.25	572.00	0.15	0.40
166	tweed	5	8.24	1	4	1	2	4	419.00	450.00	570.00	0.05	0.45
167	peers	5	9.33	1	4	2	10	40	539.00	359.00	276.19	0.00	0.50
168	lemon	5	12.30	2	5	1	2	9	260.67	645.00	608.00	0.00	0.55
169	sweat	5	14.33	1	4	1	3	8	368.50	540.33	569.00	0.00	0.60
170	wiped	5	12.29	1	4	2	6	6	257.14	295.24	290.48	0.00	0.65
171	jeans	5	12.81	1	4	1	5	23	300.00	660.00	666.67	0.00	0.65
172	puppy	5	4.79	2	4	2	4	8	203.00	635.00	623.00	0.00	0.75
173	scalp	5	3.96	1	5	1	6	1	414.29	590.00	614.29	0.10	0.80
174	pearl	5	7.68	1	3	1	1	32	356.00	610.67	597.00	0.00	0.85
175	veins	5	8.46	1	4	2	2	30	435.00	545.67	504.76	0.05	0.90
176	altar	5	9.86	2	4	1	2	1	407.00	575.00	547.62	0.00	0.90
177	fence	5	16.72	1	4	1	2	10	346.00	592.50	597.00	0.00	0.90
178	pizza	5	3.68	2	5	1	1	0	271.43	690.48	671.43	0.00	0.95
179	witch	5	6.59	1	3	1	6	18	280.00	633.25	640.00	0.00	1.00
180	chalk	5	9.80	1	3	1	0	18	260.00	628.00	634.00	0.00	1.00
	Mean	4.77	8.06	1.30	3.83	1.23	4.53	13.37	345.06	568.76	554.55	0.04	0.60
	SD	0.43	4.89	0.47	0.79	0.43	3.62	11.98	91.19	97.08	101.26	0.07	0.28

LF targets (items 181-210)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
181	maze	4	5.08	1	3	1	15	52	414.00	572.50	528.00	0.00	0.30
182	reef	4	5.48	1	3	1	5	30	520.00	457.50	452.38	0.00	0.50
183	doll	4	7.28	1	3	1	12	26	180.50	607.33	588.00	0.00	0.65
184	mall	4	3.11	1	3	1	18	43	589.00	516.50	459.00	0.00	0.75
185	wand	4	1.53	1	4	1	11	13	272.00	536.50	570.00	0.05	0.80
186	cork	4	9.52	1	4	1	13	18	386.00	623.33	608.00	0.05	0.85
187	dice	4	3.54	1	3	1	13	31	415.50	552.00	623.81	0.05	0.95
188	igloo	5	0.20	2	4	1	0	0	328.57	671.43	614.29	0.00	0.10
189	toxic	5	13.10	2	6	2	3	1	480.00	511.00	257.14	0.00	0.15
190	brass	5	15.97	1	4	1	8	14	414.00	552.00	577.00	0.00	0.20
191	punch	5	15.54	1	4	1	5	8	299.00	550.75	548.00	0.10	0.25
192	salsa	5	0.17	2					547.62	528.57	557.14	0.10	0.30
193	drums	5	7.32	1	5	2	4	10	257.00	596.50	685.71	0.00	0.40
194	faded	5	15.34	2	5	2	10	19	300.00	376.19	271.43	0.00	0.45
195	sober	5	6.68	2	4	1	3	5	561.00	290.00	242.86	0.00	0.50
196	shiny	5	7.67	2	4	2	3	8	241.00	542.00	411.00	0.10	0.55
197	desks	5	3.81	1	5	2	2	2	272.00	610.75	609.52	0.00	0.55
198	onion	5	6.57	2	5	1	2	0	270.50	618.50	632.00	0.00	0.60
199	twins	5	12.53	1	5	2	4	7	307.00	531.50	542.86	0.00	0.60
200	vague	5	15.89	1	3	1	3	9	505.50	275.33	272.00	0.30	0.60
201	stain	5	5.26	1	4	1	6	20	343.00	558.67	535.00	0.00	0.65
202	feast	5	9.54	1	4	1	3	11	360.00	592.00	610.00	0.10	0.65
203	tribe	5	8.06	1	4	1	5	8	401.00	526.67	504.00	0.10	0.70
204	ankle	5	10.41	2	4	1	1	3	251.00	645.00	608.00	0.00	0.75
205	dusty	5	8.28	2	5	2	7	8	339.00	502.00	452.00	0.00	0.85
206	ivory	5	9.77	3	5	1	0	2	523.81	529.00	571.00	0.25	0.85
207	fizzy	5	1.06	2	4	2	3	11	242.86	509.52	457.14	0.00	0.90
208	stamp	5	14.23	1	5	1	4	4	269.00	552.00	609.52	0.00	0.90
209	fibre	5	16.64	2	4	1	0	8	511.00	433.00	300.00	0.20	0.95
210	beast	5	9.87	1	4	1	6	18	353.50	569.33	564.00	0.15	1.00
	Mean	4.77	8.31	1.43	4.14	1.28	5.83	13.41	371.81	531.25	508.69	0.05	0.61
	SD	0.43	4.95	0.57	0.79	0.45	4.77	12.67	113.29	90.83	125.94	0.08	0.25

LF targets (items 211-240)

Item number	Target	Number of letters	Frequency per million	Number of syllables	Number of phonemes	Number of morphemes	Number of orthographic neighbours	Number of phonological neighbours	Age-of-Acquisition score	Imageability rating	Concreteness rating	Cloze Neutral	Cloze Biasing
211	dome	4	7.10	1	3	1	15	23	384.00	522.00	517.00	0.00	0.10
212	herd	4	7.43	1	3	1	11	34	325.00	554.00	565.00	0.15	0.35
213	lice	4	1.83	1	3	1	14	32	413.00	563.00	543.00	0.00	0.55
214	slug	4	2.24	1	4	1	8	11	209.52	570.00	671.43	0.00	0.60
215	pier	4	5.78	1	3	1	5	49	420.00	571.67	588.00	0.00	0.75
216	yolk	4	1.11	1	3	1	2	11	280.95	525.00	593.00	0.00	0.80
217	moth	4	2.99	1	3	1	6	14	290.48	588.50	550.00	0.25	0.80
218	muddy	5	6.70	2	4	2	3	17	204.76	522.00	476.19	0.00	0.20
219	barge	5	3.50	1	4	1	5	14	457.14	570.00	561.90	0.00	0.25
220	chewy	5	0.32	2					276.19	323.81	328.57	0.10	0.25
221	shaky	5	5.08	2	4	2	4	3	321.43	376.19	333.33	0.00	0.30
222	acorn	5	2.59	2	5	1	2	1	270.00	545.00	638.10	0.00	0.40
223	album	5	23.01	2	5	1	0	0	352.38	523.81	533.33	0.00	0.40
224	spade	5	3.03	1	4	1	7	16	204.76	549.00	565.00	0.00	0.45
225	stack	5	6.98	1	4	1	11	17	409.00	443.50	366.67	0.00	0.45
226	ditch	5	6.78	1	3	1	8	26	292.00	574.00	555.00	0.05	0.50
227	stale	5	4.33	1	4	1	13	22	368.00	365.00	393.00	0.00	0.55
228	cakes	5	10.61	1	4	2	19	25	197.00	648.00	657.14	0.05	0.60
229	skull	5	12.34	1	4	1	3	9	361.50	636.50	570.00	0.05	0.60
230	lions	5	9.96	1	5	2	3	4	232.67	640.50	623.81	0.23	0.70
231	beard	5	9.74	1	4	1	4	25	260.00	630.33	580.00	0.00	0.70
232	claws	5	3.16	1	4	2	9	11	350.00	611.00	614.29	0.05	0.80
233	piano	5	20.21	3	5	1	0	0	249.00	645.00	615.00	0.05	0.80
234	snail	5	2.99	1	4	1	1	3	265.00	617.33	579.00	0.20	0.85
235	steak	5	4.32	1	4	1	5	18	329.00	648.50	646.00	0.10	0.85
236	salad	5	10.86	2	5	1	0	4	358.50	638.00	595.00	0.00	0.85
237	scary	5	1.94	2	5	2	5	5	195.00	338.10	276.19	0.00	0.90
238	zebra	5	2.21	2	5	1	0	1	352.50	649.00	564.00	0.00	0.90
239	cloud	5	23.07	1	4	1	2	15	206.50	637.25	554.00	0.45	0.95
240	grave	5	23.11	1	4	1	9	19	339.00	634.50	535.00	0.40	1.00
	Mean	4.77	7.51	1.33	4.00	1.21	6.00	14.79	305.81	555.35	539.60	0.07	0.61
	SD	0.43	6.71	0.55	0.71	0.41	4.99	11.70	74.67	96.21	101.43	0.12	0.25