UNIVERSITY OF SOUTHAMPTON

FACULTY OF HUMANITIES

Modern Languages

Semantic Neighbourhood Density Effects in Word Identification during Normal Reading: Evidence from Eye Movements

by

Badriya Al Farsi

Thesis for the degree of Doctor of Philosophy

August 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF HUMANITIES

Applied Linguistics

Doctor of Philosophy

SEMANTIC NEIGHBOURHOOD DENSITY EFFECTS IN WORD IDENTIFICATION DURING NORMAL READING: EVIDENCE FROM EYE MOVEMENTS

by Badriya Humaid Al Farsi

Eye movement studies (e.g., lexical ambiguity and semantic plausibility studies) suggesting that word meaning can influence lexical processing relied on contextual information. Therefore, these studies provide only a limited insight into whether the semantic characteristics of a fixated word can be accessed before the completion of its unique word identification. The present thesis investigated the effect of the semantic characteristics of a word in its lexical processing during normal reading. In particular, four experiments were carried out to examine the effects of semantic neighbourhood density (SND, defined by mean distance between a given word and all its co-occurrence neighbours falling within a specific threshold in semantic space, Shaoul & Westbury, 2010a) in normal reading. The findings indicated that the SND characteristics of the fixated word influenced the lexical processing of the fixated word itself and the subsequent words, as evident in early reading time measures associated with lexical processing. These results suggest that a word's semantic representation can be activated and can influence lexical processing before the completion of unique word identification during normal reading. The findings were discussed in terms of Stolz & Besner's (1996) embellished interactive-activation model (McClelland & Rumelhart, 1981) and the models of eye movement control during reading.

Table of Contents

List of tables

List of figures

DECLARATION OF AUTHORSHIP

I, BADRIYA HUMAID AL FARSI

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Semantic Neighbourhood Density Effects in Word Identification during Normal Reading: Evidence from Eye Movements

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Experiment 1, 3, and 4 are in preparation for submission to the following manuscript:
   Al Farsi, B. H., Rule, S., & Liversedge, S. P. (2014). Semantic neighborhood density effect in lexical identification during normal reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Signed:...

Date: ....

Acknowledgements

I would like to express my appreciation to my supervisor, Prof Simon Liversedge, for helping me to immerse myself in the field of eye movements during reading, and challenging me to think deeply about the topics of this thesis. Many thanks also to my second supervisor Dr Sarah Rule for providing some intriguing questions about some aspects of the thesis and for her help in recruiting some participants for my experiments.

My deep gratitude goes to Dr Denis Drieghe for his assistance with the *R* script that I wrote to analyse the eye movement data and for his help in understanding the results obtained from running linear-mixed effect models. I would like also to thank Dr Hazel Blyth, for her advice on the design of some experiments in this thesis and on dealing with eye movement data. And I also owe thanks to Dr Anne McBride, for her assistance in conducting the norming studies for Experiment 3 and 4. My thanks also go to Dr Chris Westbury, Dr Cyrus Shaoul and Dr Lori Buchanan for clarifying (via email correspondence) some aspects related to semantic distributional modelling.

I would like also to thank all the participants who devoted some of their time to take part in my experiments. Without their help, the research questions of this PhD would not have been answered.

Finally, I would like to thank my father, who would be proud of me, and mother for their assurance and support, my friend Amal for her emotional support, and my son for bearing with me throughout the PhD process.

Definitions and Abbreviations

ANOVA: analysis of variance

ARC: average radius of co-occurrence

GAG: guidance by attentional gradient

HF: high frequency

HON: high number of orthographic neighbours

HSND: high semantic neighbourhood density

IA: interactive-activation

LF: low frequency

LME: linear-mixed effects

LON: low number of orthographic neighbours

LSND: low semantic neighbourhood density

ms: milliseconds

SAS: serial attention shift

SND: semantic neighbourhood density

Chapter 1: General Introduction

Reading is a sophisticated task that involves a series of efficient and highly automated processes, some of which are contingent on others so they may take place (Rayner, 1998, 2009). When we read a written text in English, our eyes move from left to right. The eyes make jerky eye movements with a series of pauses (called *fixations*) during which the eyes remain stationary. Between these fixations, the eyes make fast movements (called *saccades*) in which the eyes move forward (or sometimes backward) to bring the next word (or portion of text) into the centre of vision (called *fovea*). At the beginning of a fixation, visual information of the currently fixated word is extracted automatically. The extraction of visual information involves detecting individual letters of the fixated word through the analysis of their visual features (e.g., horizontal lines, corners, etc.). This process of letter detection is known as orthographic encoding. Upon orthographic encoding, words are identified (the lexical identification process).

Identifying a printed word in text entails first activating its representations stored in the long-term memory, including the spelling of the word or how the word looks (i.e., the orthographic information), its sounds or pronunciation (i.e., phonological information) and its meaning (i.e., semantic information), and then accessing its syntactic category (e.g., verb, noun, or adjective, etc.). Once the syntactic category of the word is available, syntactic processing takes place whereby a reader computes the structural relationships between the individual words in a sentence. This syntactic processing allows the reader to understand who or what did what to whom. Upon extracting the meaning of individual words and the structural relations between the words in the sentence, the meaning of the whole sentence is constructed in a word-by-word basis as each word in the sentence is read (Pickering, 1999; Pickering & Traxler, 1998).

Identifying individual words is an essential part of the reading process that should occur first so that understanding the structural relations between words and comprehending the sentence as a whole may take place. Therefore, lexical processing (or word identification) has received much attention in the literature, with many researchers exploring the effects of different aspects of word representation (orthographic, phonological and semantic information of words) in word identification. While the influence of many orthographic and phonological factors in normal reading are well researched and understood (Rayner, 1998; 2009), the effect of semantic characteristics of a word in its identification has received relatively less attention. A complete account of lexical processing during normal reading requires understanding the role of word meaning in lexical processing. Currently, little is understood about whether and how meaning of an individual word is extracted during early stages of word identification in reading. This thesis primarily focuses on how word meaning contributes to the process of word identification during normal reading using eye movement recording methodology.

In this thesis, eye movement recording during reading was used as a methodology that does not disturb the cognitive processing that occurs during normal reading. In the first place, eye movements are inherent to the reading process as our eyes move across text during normal reading. Secondly, participants read a given text without being asked to make an overt decision about the presented stimuli to indicate that they have identified the presented stimuli. In this way, the eye movement data reflect only the cognitive processes taking place in normal reading without requiring a secondary task of making a decision about the stimuli. It should be noted that how long the eyes remain fixated on a given word is largely associated with the ease or difficulty with which a word is identified; words that are difficult to identify are fixated for a longer time than words that are relatively easy to process (Rayner, 1998; 2009). The difficulty with which a word is lexically identified pertains to the characteristics (orthographic, phonological and possibly semantic information) associated with the word itself, an issue that will be elaborated upon later in this chapter. Because of this link between eye movements and linguistic processing, eye movement recording during normal reading has been used as a nonintrusive methodology.

This chapter will provide a general introduction to the characteristics of eye movements and the models of eye movement control during reading, then the Introduction will turn to discussing visual and lexical processing as part of the reading process. Understanding the characteristics of eye movements during reading first will be necessary to understand some issues related to lexical processing that may only occur once visual processing has taken place. The discussion of the reading process in this chapter will be divided into three main parts as follows. Section 1.1 will provide a general overview of the field of eye movements during reading, the major issues related to the research in this area and will briefly introduce the models of eye movement control. Section 1.2 will review research on visual processing of text during reading. Section 1.3 will describe lexical processing in reading, drawing attention to the factors that are found to influence the ease or difficulty by which words are identified. Finally, Section 1.4 will conclude and summarise written language processing and how this relates to the thesis.

1.1 Eye Movements during Reading

This section will provide some general characteristics of eye movements and how the eyes move during reading, and will discuss some related issues in the field of eye movements and two influential models of eye movements during reading.

In normal reading, the eyes make fast movements and pauses as the reader progresses through text. The pauses are called fixations, and the fast movements are called saccades. Information from the fixated word can be acquired during a fixation whereas no new information is gained during a saccade because vision is suppressed during saccades (Rayner, 2009). However, this is not to say that nothing happens during saccades. Instead, cognitive processing continues during saccades (Irwin, 1998; Irwin & Carlson-Radvansky, 1996).

A reader is able to extract an amount of visual and linguistic information in a single fixation. The amount of information that can be extracted from a single fixation is limited, and therefore, the eyes move to a new location in the text. The size and the region from which readers can extract useful information on a fixation during reading is called *perceptual span*. Human vision can be divided into three regions: foveal, parafoveal and peripheral regions. The foveal field of vision is the central region and is characterised by clear visual acuity up to two degrees of visual angle (Rayner & Bertera, 1979) (one degree of visual angle equals to 3-4 letters at a distance of about 60cm; Balota & Rayner, 1991). Outside the fovea is the *parafovea* that extends 5 degrees on either side of the fovea (Rayner & Bertera, 1979), with declining visual acuity compared to the foveal region. Next to the parafoveal region is the *periphery*, in which there is a severe decline in visual acuity. That is, visual acuity attenuates as a function of distance from the fovea. Because of this decline in visual acuity outside the fovea, our eyes make saccades to bring the next word into the foveal visual field so that it can be clearly viewed and therefore processed. Because of this gradient in visual acuity, only information about the currently fixated word and partial information about the words to the right of a fixation (i.e., parafoveal words)[1] is extracted.

The decision of the amount of time to spend in fixating a word (i.e., fixation durations) relates to the question of when the eyes move from the current fixation. The average fixation duration in reading is 225- 275ms[2]. Fixation durations for an individual reader can range from 50ms to 600ms depending on the difficulty or the ease by which the reader processes the fixated words. A fixation that falls around 50-150ms below the average is considered short, reflecting that the fixated word is relatively easy to process. A fixation that falls within 500-600ms beyond the average is considered long, reflecting that the reader experiences difficulties processing the fixated word. Along with the average fixation duration that gives a summary of processing taking place in normal reading, other eye movement measures are used in reporting eye movement data to give a more comprehensive account of the moment-to-moment processing as described below.

When the unit of analysis is a word, the following measures are considered. *First fixation duration* is the duration of first fixation on the word, regardless of whether the target word receives one or more fixations. This measure is used as a computational index of various linguistic phenomena of word processing such as lexical processing related to the orthographic and

---

1 This thesis focuses on normal text reading in English language. As such, terms such as right of a fixation that are mentioned in this thesis will be assumed to be applicable to English and any other language read from left to right. In languages that are read from right to left such as Arabic and Hebrew, the term 'right of a fixation' in this thesis will be equivalent to 'left of fixation' in such languages.
2 It should be noted that any given average fixation durations or saccade length or any other values in this introductory chapter vary as a function of reading skills, target word/ text difficulty, and or orthographic writing system of a specific language.

phonological properties of the fixated words and post-lexical processing related to integrating the meaning of words into the overall sentential context (Inhoff & Radach, 1998). *Single fixation duration* is used when the eyes only make a single fixation on the word during the first pass through the sentence. *Gaze duration* is the sum of all fixations made on a word before the eyes move to another word. This measure is also used as an index of quite early cognitive processing such as lexical and post-lexical processing. Inhoff (1984) argues that the first fixation duration is an indicator of lexical access (early processing influenced by factors such as word frequency) and that the gaze duration reflects not only lexical access, but also text integration processes (e.g., readers' detecting mis-analysis of interpreting earlier parts of the sentence being read). However, other researchers (Rayner & Pollatsek, 1987; O'Regan & Levy- Schoen, 1987) pointed out that the gaze duration does not necessarily indicate text integration processes, given that the eyes sometimes land erroneously in a less than optimal position in the word which necessitates a re-fixation. Also, a considerable number of studies reported that both first fixation duration and gaze duration produced similar statistical significance and converged to similar research conclusions (see e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Reichle, Rayner, & Pollatsek, 1999). If the effect of a variable of interest could be only established in gaze duration, but not in the first fixation duration, then the gaze duration effects suggest that this aspect of processing occurred later than the very first fixation on the word (perhaps at a stage of integration). Therefore, Rayner (1998) recommended using all of these three above-mentioned measures of processing times to give a comprehensive account of the time course of processing.

When the region of interest within the sentence is larger than a word (e.g., a region of 3 or 4 consecutive words), other measures are often reported such as *first-pass reading time* (the sum of all fixations in a region before leaving the region), *second-pass reading* (the sum of all fixations in a region following the first-pass time), *regression path duration* (*a.k.a.* go-past time; sum of all fixations in a region before leaving it to the right of the region, including fixations made during any regressions to earlier parts of the sentence) and *total reading time* (the sum of all fixations in a region which includes both forward and backward movements). In fact, these measures may also be computed for a single word region, however, they are often not reported.

Regression path duration and total reading time are late measures of eye movements. If the effect of a manipulated variable could only be established in regression path duration and or total reading time, then this effect would be associated with the later stages of linguistic processing (e.g., aspects of discourse integration rather than earlier processes such as word identification).

Fixation durations on a word can also be influenced by the characteristics of the previous words, a phenomenon known as *spillover effects*. For example, Rayner and Duffy (1986) found that the fixation durations on a word tended to be longer when the previous word was difficult to process. That is, the decision of when to move the eyes appears to be influenced not only by the characteristics of the currently fixated word, but may also be influenced by the characteristics of the previous word. Also, eye movement data have suggested that information extracted from the upcoming word (the parafoveal word) while the foveal word is still fixated was reported to facilitate the processing of the parafoveal word when it is subsequently fixated (e.g., Pollatsek, Lesch, Morris, & Rayner, 1992), a phenomenon known as the *parafoveal preview effect*. That is, readers extract information from the word to the right of the fixated word (i.e., from the parafoveal word), which suggests that processing of a word begins before the word is actually fixated (when it is still in the parafoveal visual field), and this in turn makes the reading process more efficient. These effects will be further discussed in Section 1.2 and 1.3.

It should be noted that not all words receive a fixation. Few words of about 20-30% of the words in a text are skipped (i.e., not fixated) during reading. This is not to say that the word that is skipped is not being processed. Instead, it is often the case that the skipped word is processed in the previous fixation while it is in the parafovea (Rayner, White, Kambe, Miller, & Liversedge, 2003), especially if it is a function word and a short word (Rayner & McConkie, 1976) as will be discussed further in this chapter. Also, the same word is often fixated more than once in succession (i.e., refixated), especially if it is a long content word or a difficult-to-process word (e.g., Blythe, Liversedge, Joseph, White, & Rayner, 2009), or if the initial fixation on the word lands in a less than optimal viewing location in the word (e.g., near the end of the word rather than towards the centre of the word) (McConkie Kerr, Reddix, Zola, & Jacobs, 1989; McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979). Rayner (1998) recommended using skipping probability and refixation probability along with the other measures of fixation times (fixation durations) in order to obtain a clear insight into the nature of processing in reading.

As mentioned earlier, saccades are eye movements that are made to bring a new region of text into the foveal field of vision. Saccadic targeting relates to the decision about which word is going to be fixated next and where in the word the eyes or the saccade is going to land. In normal reading, our eyes typically tend to land halfway between the middle of the word and the beginning of that word—a location known as *the preferred viewing location* (McConkie, Kerr, Raddix, Zola, & Jacobs, 1989; Rayner, 1979). This location is often contrasted with the *optimal viewing position* located closer to the centre of a word, a little to the right of the preferred viewing location. O'Regan, Levy-Schoen, Pynte, and Brugaillere (1979) found that the optimal viewing position decreased the amount of time spent recognising words presented in isolation. They also reported that when the readers fixated a non-optimal viewing location on a word, the readers tended to refixate those words and they tended to spend approximately 20ms longer in recongising the word for every letter that the reader's fixation deviated from the optimal viewing position. However, this latter processing cost could not be established in normal reading, and the refixation cost was found to be more likely in normal reading when the first fixation landed at the beginning or end of a word than in the middle (McConkie, Kerr, Raddix, Zola, & Jacobs, 1989; Vitu, O'Regan, & Mittau, 1990). Such findings in normal reading indicate that the fixation location or where the reader fixates on a word influences the ease with which the word is processed.

The landing position on a word was also found to be influenced by the amount of visual information extracted before the word is fixated (i.e., when it is on the parafovea). To explain, if readers obtain parafoveal preview of the first three letters of an eight-letter parafoveal word, then the eyes tend to move to the third or fourth letter of the parafoveal word when it is subsequently fixated (Rayner, Sereno, & Raney, 1996; see also Inhoff, Radach, Eiter, & Juhasz, 2003; Juhasz, White, Liversedge, & Rayner, 2008).

A saccade can be either *progressive* (moving the eyes forward in the text) indicating that the processing of a word (or text) is successful, or *regressive* (moving the eyes backward to previous word(s) in the text; *regressions*) often indicating that the reader is experiencing difficulties in comprehending the text. 10-15% of all saccades in reading are regressions. Short within-word regressions (that serve to bring the eyes to the left of the currently fixated word) are possible, and they can be due to oculomotor error or due to the reader finding it difficult to process the fixated word. There is another type of saccade made when the eyes move from the end of one line in a text to the beginning of the following line; this type of saccade is called *return sweep*.

In general, it takes about 175- 200ms for the oculomotor system to decide upon the location of the upcoming saccade target and to program an eye movement (Rayner, Slowiaczek, Clifton, & Bertera, 1983). The duration of the actual movement of the eyes (*saccade duration*) varies according to the distance the eyes move in the text. In reading, the eyes move a distance of approximately two- degrees (on the assumption that a degree equals approximately 3-4 letters). Such a saccade would last about 30ms (Rayner, 1987). That is, the average distance that the eyes travel from one fixation to another fixation (*saccade length*/ size/ amplitude) is about 7- 9 letter spaces for readers of alphabetic languages. However, the length of saccades can vary to be as short as a one-letter space or as long as 15-20 letter spaces. These long saccades occur particularly when the eyes regress to previous words. Saccade length is more influenced by low-level visual factors as will be discussed later in this section. Other measures pertaining to saccades reported in reading research are *fixation position/location* (the letter within a word where a fixation is located), *launch site* (the distance in letter spaces between the location of the prior fixation and the current fixation), *fixation probability* (the frequency with which a word is fixated), *skipping rate* (frequency that a word is skipped), and *refixation probability* (frequency of making at least one additional fixation on the currently fixated word before leaving it).

The point that should be noted here is that the decision of the amount of time spent fixating a word (i.e., fixation durations) or when the eyes move in reading is made independent of the decision of where the eyes are targeted (McConkie & Rayner, 1976; Rayner & Pollatsek, 1981). This is apparent in that these two components of eye movement control—'when' and 'where' to move the eyes-- are influenced by different aspects of the text. Generally, the decision of where to move the eyes and the saccade length appears to be mainly determined by low-level visual information associated with words (e.g., word length). The decision of when to move the eyes, that is, the decision that affects fixation durations, is very largely influenced by linguistic processing of that word and the text (Rayner, 1998); the difficulty or ease of processing a word directly affects how long a fixation will last. The concern of the experiments described in this thesis will focus on the ‘when’ decision and whether a word’s semantic characteristics influence fixation durations on the word during fluent reading.

Computational models of eye movement control in reading have been developed that attempt to simulate oculomotor control during reading and to explain how eye movements relate to the processes underlying reading. Models of eye movement control in reading predict the actual fixation durations and saccade length and fixation locations; thus, they can be directly used to predict human data. The different models differ in a large number of respects, however, broadly, they can be categorised according to two assumptions about (1) the extent to which perceptual, cognitive, and motor control processes guide the eyes through text (i.e., influence the decision of when and where to move the eyes), and (2) how attention is allocated to words during reading. Considering first perceptual and cognitive processing in relation to models of eye movement control, we can further classify models into two types, oculomotor models and cognitive models.

The oculomotor models (e.g., O’Regan, 1992; O’Regan, 1990; O’Regan & Levy-Schoen, 1987; Yang & McConkie, 2004) generally assume that eye movements are driven by low-level visuo-oculomotor processing, and that making eye movements is only indirectly influenced by linguistic processing. Studies investigating the effect of visual factors on eye movements suggest that the initial landing position of the eyes on a word influences the fixation durations that will be made on that word, as well as where the next fixation will be made (McConkie *et al.,* 1988; O'Regan & Levy-Schoen, 1987; Vitu, McConkie, Kerr, & O'Regan, 2001), and the probability that a regression will be made (Vitu & McConkie, 2000).

Based on the effects of low-level visual information that were found to occur for early measures of processing such as first fixation duration, or single fixation duration, O’Regan and colleagues (e.g., O'Regan, 1992, O'Regan, 1990, O'Regan & Levy-Schoen, 1987) proposed the Strategy-Tactics model, according to which readers have developed strategies that are quite automated based on years of reading experience. A key piece of evidence to support this view comes from work investigating the optimal viewing position effects in reading. Specifically, as mentioned earlier, when the eyes land on a location close to the centre of the fixated word (i.e., the optimal viewing position), fixation durations were found to be shortest and the probability of refixating the word were the lowest. Thus, O'Regan and colleagues argue that low-level visual information rather than the linguistic processing is the primary factor that affects eye movement control during reading. According to this view, the eyes are guided by simple strategies driven by low-level visual information, such as targeting the longest word in the parafoveal field of vision extending 20 character spaces to the right of the fixated word, targeting saccades to the centre of words, and using the length of a parafoveal word to inform the decision of whether to skip it, etc. Oculomotor theories are good at accounting for the effect of visual factors on eye movement behaviour during reading, but cannot explain the well-documented effects of linguistic factors such as word frequency, plausibility, syntactic processing, etc. Since there are abundant data in the literature that indicate that linguistic processing influences the decision of when to move the eyes during reading and how long to fixate on a word (Rayner, 1998, 2009), oculomotor accounts have been considered to be quite limited in their explanatory power. Indeed, O'Regan, Vitu, Radach, and Kerr (1994) acknowledge the limitations of the oculomotor models that propose that eye movement behaviour is determined by pure visual processing, suggesting that a coherent theory of eye movement control will need to be a 'hybrid' theory that combines elements of both visual and cognitive processing.

The cognitive models, on the other hand, generally postulate that eye movements are driven by the ongoing linguistic processing underlying reading. Specifically, the cognitive models make a link between eye movements and linguistic processing based on the wealth of eye movement data showing that the linguistic characteristics of words influence the fixation durations even when the words were withheld from readers shortly after readers fixate them (Rayner, Liversedge, & White, 2006). These eye movement data also showed that linguistic characteristics such as word frequency influence how long the eyes remain in the same place of the withheld words, with the eyes tending to remain longer when the word is of low, compared to high frequency, even though the word is no longer present (Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003). Such findings suggest that how long the eyes remain fixated on a given word is largely associated with the ease or difficulty with which a word is identified; words that are difficult to identify are fixated for a longer time than words that are relatively easy to process (Rayner, 1998; 2009). These models themselves can be classified into two groups based on how they consider that attention is allocated during reading: *serial-attention-shift* (SAS) models and *guidance-by-attentional-gradient* (GAG) models. SAS and GAG models differ in the assumption of whether visual attention is distributed serially to only one word at a time or distributed as a gradient of processing that usually encompasses more than one word in parallel at a time. The focus of this thesis is not on the difference between these models. However, an overview of these models will provide a basis for understanding how they address the issue of the time course of when the word meaning can be extracted during lexical processing, which is central to the research questions raised in this thesis, as will be discussed in Chapter 3, 4, 5 and 6. What follows is a description of two of the most influential models of eye movement control.

In attempt to develop a hybrid model, Reichle, Pollatsek, Fisher and Rayner (1998) proposed the E-Z Reader model (later modified by Pollatsek, Reichle, & Rayner, 2006; and Pollatsek, Juhasz, Reichle, Machacek, & Rayner, 2008) in which eye movement behaviour during reading is stipulated to be affected by some visual factors, but primarily driven by cognitive factors. The E-Z Reader model was developed based on the idea of serial allocation of attention (SAS). The basic assumption of this model is that attention moves sequentially from one word to another. That is, the processing of the next word may not begin until the processing of the currently fixated word has finished. Another central assumption to the E-Z Reader model is that accessing the meaning associated with the fixated word (i.e., the completion of lexical identification) signals the eyes to move to the next word (i.e., a linguistic processing event triggers an eye movement). Thus, by shifting the attention from one word to another, readers can process each word in its correct order.

In the E-Z Reader model, there are two stages of lexical processing: L1 stage of lexical processing and L2 stage of lexical processing. The L1 stage starts a familiarity check in which the system assesses how familiar the upcoming word is. The familiarity check is based on the orthographic familiarity (i.e., how often the word has been seen before) and sentential contextual constraints, which is in line with the research that showed that the quality of a word's orthographic form (e.g., reduction in the contrast of letter strings, case alternation, boldface type, etc.) influenced the fixation duration on that word (Reingold & Rayner, 2006). The average time required to complete this stage is influenced by the frequency of the word (in terms of occurrence in language as indexed by corpus studies) and its predictability from the prior context of the sentence (as estimated using cloze tasks). Words that are frequent and/or predictable take less time to be checked compared to words that occur less frequently and/or are unpredictable from the prior context. Whether the familiarity check (i.e., the L1 stage) can be influenced by variables related to word meaning will be of interest to the research undertaken in this thesis. Once the L1 stage is complete, two stages occur simultaneously: (1) the eye movement system begins to program the next saccade, and (2) the L2 stage of lexical processing starts.

In the L2 stage of lexical processing, the meaning of a word, whether it is predictable or not, is activated. This stage is not influenced by visual information or visual acuity since the information being accessed in this stage is semantic, rather than visual. As such, the L2 stage is influenced by higher-order linguistic processing such as the semantic characteristics of the foveal words. This L2 stage in which the semantic meaning of a word is extracted is critical to the questions being raised in this thesis, and will be further explained in the next chapters. When the L2 stage is complete, the reader redirects attention to the next word so that attention is allocated to the parafoveal word but the eyes are still fixated on the foveal word. Thus, the familiarity check of the parafoveal word begins while the eyes are still on the foveal word but only starts after the lexical processing of the foveal word has been completed. This is how the E-Z Reader model explains parafoveal preview. At this point, one of two thing will occur, (i) if the preliminary stage of saccade planning (called the labile stage) is completed before the familiarity check on the next word is finished, then a saccade will be executed and the next word will be fixated, or (ii) if the familiarity check on the next word is completed before the execution of the saccade, then the current planned saccade will be cancelled, and a new saccade will be planned to the word after the parafoveal word in the text (word n+2), thus the parafoveal word (n+1) will be skipped. To be clear, the E-Z Reader model assumes that word n+1 was skipped because it was recognised during the time that attention was shifted to it while the eyes were still on the foveal word, which in turn cued the eye movement system to cancel the saccade to word n+1 and make a new saccade to word n+2. This cancellation of saccade and planning a new saccade takes time and, accordingly, the E-Z Reader model predicts that fixation duration on word n is inflated prior to skipping word n+1.

As is obvious from the description above, the E-Z Reader model can account for various eye movement characteristics such as word skipping and parafoveal preview effects. It is important to note that the E-Z Reader model assumes that the effects on a word (word n) all derive from the extent to which it was processed when the fixations were on the word before (word n-1) (i.e., when it was in the parafoveal visual field). Thus, the E-Z Reader model, to some extent, acknowledges parallel lexical processing of the parafoveal word while fixating the foveal word, but only at the shallow orthographic and phonological levels (e.g., the first three letters and phonemes). To clarify, the E-Z Reader model allows the parafoveal word to be identified while the foveal word is still fixated, but only after the completion of the lexical identification of the foveal word. In this case, the parafoveal word would be skipped. However, if the full identification of the parafoveal word does not occur while the foveal word is still fixated, then the parafoveal word should not influence the fixation durations spent on the foveal word.

A second influential model is the saccade-generation with inhibition by foveal targets (SWIFT) model (Engbert Nuthmann, Richter, & Kliegl, 2005; Kliegl Nuthmann, & Engbert, 2006; Kliegl, Risse, & Laubrock, 2007). This is a guidance-by-attentional-gradient (GAG) model that implements the assumptions of the parallel attention allocation model. The main assumption of this model is that attention is allocated not only to the foveal word, but simultaneously also across the neighbouring words in the perceptual span. That is, this model allows for simultaneous lexical processing of more than one word. To explain, although there is parallel allocation of attention and lexical processing of these words, attention is allocated to words according to a gradient, with foveal word *n* receiving most attention while the other more peripheral words receive less attention. A difficult foveal word limits the number of words attended to in a single fixation. According to SWIFT, when word n is fixated for a longer time, word n+1 will be in the perceptual span for a longer time. As such, word n+1 will be processed more, which, in turn, will lead to an increased probability that the word n+1 will be skipped. Saccades are triggered by a random timer which beings saccade planning at random intervals of time. Linguistic variables, such as the frequency of word occurrence in language, influence fixation times and saccade planning only indirectly by inhibiting the random timer from executing a saccade when the foveal word is difficult to process.

If the parallel processing assumption of SWIFT is correct, then it is possible that fixation durations on the currently fixated word can be inflated by a difficult-to-process parafoveal word. That is, the lexical characteristics of the parafoveal word, such as its orthographic, phonological and meaning properties, can influence the time spent fixating the foveal word (a phenomena known as *parafoveal-on-foveal effects*), according to SWIFT. This represents a fundamental difference between GAG and SAS models as to whether lexical information extracted from the parafoveal word has a direct influence on fixation times on the foveal word. As noted earlier, the E-Z Reader model acknowledges parallel processing of the parafoveal word at the shallow orthographic and phonological levels and uses the extracted parafoveal information to determine the saccade target. According to the E-Z Reader model, this parallel processing occurs only after lexical processing of the foveal word has been completed and the programming of the saccade to the parafoveal word has been initiated. Specifically, shifting attention occurs before the eyes move to the next word; parafoveal processing happens during the time that attention is on the parafoveal word but the eyes are still on the foveal word. To summarise, the E-Z Reader model predicts that the lexical characteristics of the parafoveal word should not influence the fixation duration on the foveal word. As such, finding reliable parafoveal-on-foveal effects would undermine the core assumptions of the E-Z Reader model. In contrast, SWIFT allows parallel processing of multiple words in a single fixation and assumes that saccades target words that have the highest level of excitation (activations). Thus, SWIFT can naturally explain the effects of the characteristics of parafoveal words on the fixation durations on the foveal words.

Experimental work investigating parafoveal-on-foveal effects is mixed and the findings remain controversial. Some researchers have shown that the fixation durations on the currently fixated word are affected by the lexical characteristics of the preview presented in the parafovea (e.g., Hyönä & Bertram, 2004; Inhoff, Starr, & Shindler, 2000; Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006; Starr & Inhoff, 2004). These results provide support to the existence of parafoveal-on-foveal effects, and SWIFT, and may be taken to be challenging to the E-Z Reader model. However, these findings of the studies examining parafoveal-on-foveal effects (or parafoveal pre-processing) are inconsistent (Inhoff, Starr, & Shindler, 2000; Kennedy, 2000; Rayner, White, Kambe, Miller, & Liversedge, 2003). Also, proponents of the E-Z Reader model argue that the inflated fixations on foveal word when the parafoveal word is difficult to process can be attributed to oculomotor errors (Drieghe, Rayner, & Pollatsek, 2008; Rayner, 1975; Rayner, White, Kambe, Miller, & Liversedge, 2003). For example, Rayner, White *et al.* reported that the inflated fixations on the foveal word were observed only when the eyes landed very close to the end of the foveal word when the reader intended to make a saccade to the parafoveal word (i.e., the saccade fell short of the intended target). Thus, they concluded that the mis-located fixations due to oculomotor errors are responsible for the inflated fixations on foveal words, rather than the difficulty associated with the processing of the parafoveal words.

To summarise this section, a skilled reader generally moves the eyes about 7-9 letter spaces (in alphabetical languages such as English) every 225-275ms. Our eyes typically land between the beginning of a word and the middle of the word. Some words receive more than one fixation, especially if they are long or if the initial fixation on the word lands near the end of the word rather than near the centre of the word. The eyes also regress back to a previous word. In addition, skilled readers tend to extract information from the upcoming (parafoveal) words, which aids in their identification when they are subsequently fixated. This section also discussed the debate of whether the properties of the upcoming word can influence the fixation durations on the current word. Here findings are more contentious with some studies suggesting that lexical properties of an upcoming word can affect processing on the current word, whilst others suggest that this is not the case. This section also described two influential computational models of eye movement control that differ in their assumptions about how attention is allocated to words falling in the perceptual span. The E-Z Reader model assumes that the completion of lexical processing of the currently fixated word is what triggers eye movements to the next word, and that attention is allocated to one word at a time. In SWIFT the saccades are, on the other hand, triggered by a random timer and linguistic processing only indirectly influences the timer in relation to the execution of a saccade when the foveal word is difficult to process.

1.2 Visual Processing in Reading

The previous sections discussed some general phenomena of eye movements (e.g., parafoveal preview, word skipping, regressions, and re-fixations). This section extends the discussion by giving a detailed account of the types and amount of visual information that can be extracted from the foveal and parafoveal fields of vision in a single fixation.

Prior to undertaking any linguistic processing of a printed word, visual processing of the word must first take place. When we fixate a word on a page, the light reflects off the page and passes through the pupil (the black hole-like in the centre of the eye surrounded by the coloured part of the eye, the ‘iris’). The light passes through the lens that focuses the light reflecting from the page depending on the distance by which the page is viewed. Finally, the light reaches the retina, stimulating the photoreceptors (rods and cones) in the retina to convert the light into electro-chemical signals. These signals travel through the optic nerve to the optic tract; the signals then project onto the lateral geniculate nucleus (LGN) (the visual areas of the thalamus). At this point, the neurons of the LGN become stimulated and they send axons from the occipital lobe of the cerebral cortex to the primary visual cortex, where visual information is processed. This visual processing starts very quickly at the beginning of a fixation; eye movement studies have demonstrated that when text is masked (e.g., Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981) or disappears after only 50 or 60ms (e.g., Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004; Rayner, Liversedge, White, Vergilino-Perez, 2003) reading behaviour was not affected, suggesting that the visual information required for efficient reading is extracted very early at the onset of a fixation.

Having presented an overview of how we see words, the remainder of this section will be devoted to providing an account of the amount and type of visual information extracted in a single fixation from the effective field of vision (i.e., the perceptual span) and the factors influencing the amount of visual information acquired in a single fixation.

What kind of visual information that can be extracted in a single fixation has been the basis for many eye movement studies. Since letters make up written words, the letters of a word must be first processed before the word as a whole can be identified. As such, one type of visual information that is acquired in a single fixation is the component letters of words. Letters in a word are processed in parallel, rather than in a sequential manner (one letter at a time) during word identification (Paap, Newsome, McDonald, & Schvaneveldt, 1982; Rayner & Pollatsek, 1989). The parallel processing of letters in a word was supported by the findings that showed that letters were reported quicker with more accuracy when they were presented in words rather when they were presented in isolation—a phenomena known as *word-superiority effect* (Reicher, 1969). Letters are detected by analysing their visual features (e.g., horizontal lines, edges and corners, etc.) and are mapped onto unified abstract letter representations (Besner, Coltheart, & Davelaar, 1984; Coltheart, 1981; Evett & Humpherys, 1981; Rayner, McConkie, & Zola, 1980). They are abstract in terms of being independent of surface properties such as case, position, font type, colour, or size. No matter how letters appear in different typefaces or handwriting, mapping the visual features onto abstract letter representations allows readers to recognise words efficiently and rapidly. Support for abstract letter identities comes from research investigating the effect of aLtErNaTiNg CaSe on normal reading (Rayner, McConkie, & Zola, 1980). The reading rate for passages written in alternating case was similar to those for passages written in normal text when the size of all letters was equated in both types of passages (Smith, Lott, & Cronnell, 1969; Perea & Rosa, 2002). This suggests that visual features are encoded as abstract letter identities so that we can recognise the same word in different cases (lower case, UPPER CASE, aLtErNaTe CaSe) and in different font types and handwriting (Clifton, Staub, & Rayner, 2007; Rayner, McConkie, & Zola, 1980; Rayner & Pollatsek, 1989).

Letter order and beginning letters are important visual information that are extracted in a single fixation and that are essential for lexical identification (White, Johnson, Liversedge, & Rayner, 2008). Extracting information about letter position and the initial letters of the fixated word was found to be crucial for a word's identification in normal reading (Lima & Pollatsek, 1983; White, Johnson, Liversedge, & Rayner, 2008). Extracting the initial letters of a word is assumed to limit the number of lexical candidates (the actual perceived word and orthographically similar words) that can become active during lexical identification, an issue that will be further discussed in Section 1.3.1. Without encoding information about the position of the letters in a word, readers would not be able to detect the differences in words that share the same letters (e.g., *gum* and *mug*). In preview studies, it was found that fixation durations on target words (e.g., *clam*) were increased when their parafoveal preview (while still fixating the foveal word) were words with transposed letters (e.g., *calm*; the readers had a parafoveal preview of *clam* when fixating word n, and when their eyes moved to fixate the word n+1, the word changed to *calm*), and that the magnitude of the transposed letter effects was less than that of letter substitution within a word (*wask* was the preview of *work*) (Johnson, Perea, & Rayner, 2007; Masserang & Pollatsek, 2012; Masserang, Pollatsek, & Rayner, 2009). For example, Masserang *et al.* (2009) found that there were no differences in reading times between when the readers had a preview of the parafoveal words (while fixating the foveal word) that contained internal transposed letters (i.e., in the middle of the word such *clam* for clam) and when the preview was identical to target word (*calm* was the preview of *calm*). External transposition and substitution of the first or last letters of the words were found to pose difficulty on processing more than the internal transposition and substitutions of the middle letters of the words did (Johnson, Perea, & Rayner, 2007; Perea & Lupker, 2003; Rayner, White, Johnson, & Liversedge, 2006; White *et al.*, 2008). All of these results lend support to the claim that letter identities and letter position are important for successful word identification. In addition, the results suggest that letter positions are coded poorly/imperfectly during early stages of word identification, except for the first and last letters of words, since reading words that contained letters in the wrong place (e.g., *sturcture*) do not disturb normal reading, thus, facilitate word identification (e.g., Johnson, Perea, & Rayner, 2007; Rayner, White, Johnson, & Liversedge, 2006).

In addition to the abstract letter identities, letter order and initial letters of the currently fixated word, readers also acquire word-boundary information in a single fixation. The spaces between words provide boundary information about the to-be-fixated parafoveal word. Based on such visual information about word boundaries, the eye movement system plans a subsequent saccade to the next word. Word-boundary information includes the length of the currently fixated word and the length (and initial letters) of the next parafoveal word. The number of letters of a parafoveal word was found to influence the likelihood whether a word will be fixated or not (i.e., skipped) as well as to influence where the eyes land in the word (landing position of the eyes) (e.g., Inhoff, Radach, Eiter, & Juhasz, 2003; White, Rayner, & Liversedge, 2005). Specifically, short words were more likely to be skipped than longer words (Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Rayner, 1998; Rayner & McConkie, 1976; Rayner, Sereno, & Raney, 1996; White, Rayner, & Liversedge, 2005). Rayner and McConkie (1976) found that three-letter words were skipped 67% of the time while 7-8 letter words were skipped 20% of the time. Similar results were also reported by Drieghe, Brysbaert, Desmet, and De Baecke (2004). Drieghe and colleagues found that the likelihood of skipping two-letter words embedded in sentences, regardless of whether they were predictable from the preceding context in the sentence or not, was 25% higher than the likelihood of skipping four-letter words. They also observed that this skipping probability was inflated when predictability was taken into account. Particularly, the probability of skipping predictable two-letter words was 72% while the probability for skipping predictable four-letter words was 55%.

Running regression analyses on the data of some eye movement studies, Brysbaert, Drieghe and Vitu (2005) found that word length was the strongest predictor of word skipping as it explained 70% of the variance in the data compared to word difficulty (induced by how frequently a word appears in language as indexed by a language corpus and by word predictability from prior context) that explained 5% of the variance (see also Brysbaert & Vitu, 1998). Long words were also found to be refixated more than short words

(Vitu, O'Regan, & Mittau, 1990). Furthermore, word length of the parafoveal word also affects the landing position on that word when it is subsequently fixated. Rayner, Sereno, and Raney (1996) reported that if the distance to the next word to be fixated was large (about 8 letter spaces), then the eyes tended to move to the left of that word when it was subsequently fixated. In contrast, they showed that if the distance was small (say 2-3 letter spaces), then the eyes tended to move to the right of that word when it is later fixated, which shows that information about the word length of the parafoveal word is extracted while fixating word n and that this type of the information is useful in guiding eye movements, in particular, the decision of where the eyes land in the parafoveal words when it is subsequently fixated. It also shows that where the eyes land in the word is associated with how far to the right of fixation the parafoveal word is processed.

Other kinds of information that can be obtained from the parafovea exist at the orthographic level such as letter identity information. Earlier eye movement studies found that the initial two or three letters of the parafoveal word can be extracted prior to fixation (e.g., Henderson & Ferreira, 1990; Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, McConkie, & Zola, 1980). However, recent eye movement studies have shown that the processing of parafoveal letter identity information can involve more than the first three letters up to 9 letters to the right of fixation (Häikiö, Bertram, Hyönä, & Niemi, 2009), a result that is currently obtained due to improvements in the quality of CRT screens (see Drieghe, Rayner, & Pollatsek, 2005). Also, phonological information can be extracted parafoveally as the preview studies have shown (i.e., studies in which the parafoveal word changes to another word when the eyes move to fixate it). For instance, Pollatsek *et al*. (1992) found that first fixation duration on the parafoveal word when it was fixated was significantly shorter when its parafoveal preview was a homophone preview (*cite-site*) than when the preview was a visually similar matched control (*sake-cake*). Henderson, Dixon, Petersen, Twilley, and Ferreira (1995) also demonstrated that the fixation times on the target words were significantly shorter when their parafoveal previews were phonologically regular initial trigrams (*but* in *button*) than when their previews were irregular trigrams (*but* in *butane*). These issues regarding parafoveal processing will be further discussed in Section 1.3.

How much visual information can be extracted from an area of text (i.e., a sentence in text/ textual information) in a single fixation? Eye movement studies have demonstrated that the number of letters that can be extracted from the right of fixation was different than the number of letters that could be extracted from the left of fixation (i.e., asymmetric perceptual span to the right and left of the fixation). Specifically, 12-18 letter spaces are available to the right of the fovea (Balota & Rayner, 1991; McConkie & Rayner, 1975; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981), and 3-4 letters are available to the left of fovea (McConkie & Rayner, 1975; Rayner, Well, & Pollatsek, 1980). That is to say, the perceptual span is asymmetric, extending further in the direction of reading than in the direction opposite to reading. Other studies found that useful information was not obtained from more than two words to the right of a fixation (Rayner, 1986; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Rayner, Well, Pollatsek, & Beretra, 1982). The findings also revealed that the reading rate of young adults dropped to 60% of their normal reading rate when the readers were provided with only a foveal word while masking all other words in the sentence being read, and increased to 90% of the normal reading rate when they were provided with the foveal word and one word to the right while the reading rate was completely normal when the readers were provided with the fixated word and two words to the right (Rayner, Castelhano, & Yang, 2009; see also Rayner, Well, Pollatsek, & Beretra, 1982). These findings indicate that the perceptual span extends to approximately three words: the fixated word and two words to the right. Words to the left of a fixation point were found to have a little, if any, effect compared to the words to the right of a fixation (Rayner & Pollatsek, 1987).

The asymmetry of perceptual span reflects a psychological process that takes place during reading, rather than reflecting a physiological function of the retina (acuity). Specifically, the asymmetric perceptual span is related to the overall direction of reading, in that readers in English direct much of their attention to the right of fixation to uptake more information about text. This attention- based explanation was recently supported by the study of Miellet, O'Donnell, and Sereno (2009) who examined whether perceptual span was constrained by visual acuity or attentional resources (i.e., whether the extraction of useful information from parafoveal words was constrained by declining visual acuity in the parafoveal regions or was constrained by limited attentional resources allocated to the parafoveal words compared to the foveal word). To examine this, they developed a technique called *parafoveal magnification* in which parafoveal information was enlarged/ magnified for every eye fixation, as a way of equalizing the perceptual impact of the parafoveal word with that of the foveal word. The findings showed that the patterns of fixations for both normal texts (without magnifying the parafoveal words) and parafoveally-magnified texts were similar, suggesting that perceptual span is influenced by attentional resources rather than visual acuity. In addition, the findings from other languages with different orthographic systems were similar to the findings of the perceptual span in English. For instance, the opposite direction of perceptual span was reported for Hebrew-- a language read from right to left (Pollatsek, Bolozky, Well, & Rayner, 1981), providing a convergent evidence for the attention-based explanation of the asymmetry of the perceptual span.

It is worth-mentioning that the above-described estimate of perceptual span is not fixed, but varies according to the influence of some factors related to a reader's age, the difficulty or ease with which a reader processes a foveal word, the target of the next saccade and the orthographic system of a language. Processing foveal and parafoveal information in a single fixation can indirectly be modulated by the age of the reader. Studies showed strong evidence that both beginning readers and old adults had smaller perceptual span compared to young adults, however, this smaller perceptual span was due to different factors affecting the reading of each age group. Smaller perceptual span of beginning readers was attributed to limited attentional resources/ limited capacity of processing parafoveal words while it was attributed to declining visual acuity in the case of old adults (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986; Rayner, Castelhano, & Yang, 2009).

Another factor that influences perceptual span is the difficulty with which the foveal word is processed (a linguistic influence). Linguistic factors such as how often a word is encountered in text as indexed by corpus data (word frequency) and syntactic ambiguity (i.e., the sentence can be read and understood in two different ways) of the foveal word restrict our perceptual span (Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2005). Consider the study of Henderson and Ferreira (1990), for example, which investigated whether the difficulty associated with processing foveal words resulted in decreasing acquisition of useful information from the parafoveal words. In two experiments, they manipulated foveal processing difficulty by varying word frequency (word level) and syntactic ambiguity (sentence level). Their findings showed that both word frequency of the foveal word and syntactically disambiguating the foveal word influenced the amount of information that can be gained from the parafoveal word. Low frequency foveal words increased the initial fixations on the fixated words, and decreased the extraction of useful information from the parafoveal words. Syntactically disambiguating foveal words increased fixation durations on the syntactically disambiguating words.

The amount of parafoveal information we can gain in a single fixation also depends on the target of the saccade (a visual factor). McDonald (2006) demonstrated that if the parafoveal word (n+1) immediately to the right of the fixated word (n) was the target of the next saccade, information about the word (n+1) could be gained before fixating it. However, information about the second parafoveal word (n+2) could not be obtained unless the next saccade targeted word (n+2) and if the word (n+1) was short and skipped. This finding was further supported by Angele, Slattery, Yang, Kliegl, and Rayner (2008), Rayner, Juhasz, and Brown (2007), and Angele and Rayner (2011). The findings of Angele and Rayner suggested that when fixating word (n) information about word (n+2) could not be obtained even when word (n+1) was short in length and highly frequent, and that the frequency of the foveal word did not influence parafoveal processing of word (n+2), except when word (n+1) was skipped.

The estimate of the perceptual span also varies as a function of the orthographic system of one’s language. Logographic languages such as Chinese and Japanese were reported to have smaller perceptual span compared to English (Inhoff & Liu, 1998; Osaka, 1992). Inhoff and Liu showed that Chinese had asymmetric perceptual span that extended to one character to the left of the fixation point and three characters to the right of fixation. This reported smaller span of such languages was more likely due to the density of information that was processed during a single fixation in these languages. That is, the perceptual span is smaller when there is a large amount of information to be processed in a single fixation, suggesting that processing difficulty of a text plays a role in the amount of information that can be extracted in a single fixation.

To summarise, eye movement studies have suggested that in a single fixation, readers extract little information from the left of the fixated word in English (reflecting a psychological process of moving the eyes to the right to uptake information about the text). Because of the drop in visual acuity, no information is extracted from the 14-15 character spaces to the right of fixation. Information about the fixated word along with partial information about the next (right) word (e.g., initial letters and sound of the parafoveal word) is extracted and processed. Only partial information about the parafoveal word can be gained while fixating the foveal word because of the drop in visual acuity outside the fovea, which necessitates making a saccade to the parafoveal word to bring it into the foveal vision. The above section discussed the types of visual information that could be extracted from the foveal word and parafoveal word in a single fixation. Studies provide convergent evidence that the letters of foveal words are encoded in early stages of word identification along with letter order in the foveal word, though the latter may not be encoded perfectly. Studies also suggested that up to nine letters of the parafoveal word to the right of fixation are extracted while the eyes are still on the foveal word, and that this access to the parafoveal information makes the reading process efficient. It was also described how a foveal word that is difficult to process places demands on attentional resources, leaving less attentional resources to be allocated to the parafoveal word, and therefore, leading to a decrease in the amount of useful information that can be gained from parafoveal words while fixating the foveal word. Thus, fixation durations on a word are influenced by the ease or difficulty with which the reader processes the word (Liversedge & Findlay, 2000; Starr & Rayner, 2001), which in turn is determined by the lexical characteristics of the foveal word and the preceding text as will be discussed in the next section.

1.3 Lexical Processing in Reading

Before we can understand the structural relationships between words in a sentence or understand the overall meaning of the sentence, individual words must be first identified, a process known as lexical identification. The process of identifying a word involves accessing its (symbolic) representations that are stored in memory. The stored representations can include its orthographic form (i.e., its spelling), its phonological form (i.e., its sounds) and its semantic representation (i.e., its meaning). This process of accessing the representations of the perceived word in our memory is remarkably quick and occurs with few errors. A normal reader can identify and understand words at a rate of three or four per second (Rayner & Pollatsek, 1989).

Many models have been proposed to account for how these representations are accessed in some systematic way consistent with human word identification. The models of lexical identification often differ in the way the representations are accessed. Some models posit that a perceived word is searched for among other words in a serial manner (serial search models: e.g., Forster, 1976) while other models posit that the perceived word is accessed by activating some possible word candidates in parallel during the course of lexical processing (activation models: e.g., McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989). The models also differ in their assumptions about whether semantic information associated with a word can influence the recognition of its orthographic (or phonological) form. Below is a brief summary of some established models of lexical identification followed by a discussion of the type of information that is extracted from the page in a single fixation. It should be noted that these models were actually designed to explain the recognition of isolated words, rather than describing the process of lexical identification of words during normal silent reading. Note also that models of eye movement control during reading were not designed to explain the nature of lexical processing that takes place during reading. Instead, those models account for eye movement behaviour that occurs during reading.

One of the early models of lexical identification is the logogen model (Morton, 1969) (the word logogen was derived from Greek words: *logos* ‘word’, and *genus* ‘birth’). In this model, each word is represented by units called ‘logogens’. The logogen of a word includes information about the word’s orthographic, phonological, and semantic characteristics. Logogens act like detectors, accumulating evidence from the input received by the sensory system when a word is read or heard. In the case of reading, once the orthographic logogen has accumulated enough evidence from a printed word, a threshold is reached and the logogen fires at which point the word is lexically identified. One variable that influences a logogen's threshold is how frequently the word appears in language (word frequency). Words that occur frequently in language need less activation to reach threshold than words that occur less frequently. Therefore, the logogens of high frequency words are activated quickly, and thus, recognised faster than low frequency words that need more activation for their logogens to reach a threshold. Higher-order linguistic factors such as contextual information also affect the logogen's activation threshold. To explain, the prior context provided by words perceived earlier (e.g., semantic primes, or even potentially, words appearing earlier in the sentence) primes a reader to activate the semantically related words. The activation of the earlier words in the sentence lowers the threshold of upcoming related words, facilitating their identification. In the absence of contextual information that predicts an upcoming word, the semantic logogen of a word only becomes activated once the activation of orthographic (or phonological) logogen has reached a threshold. As such, a word's meaning is retrieved only after its orthographic (or phonological) form has been uniquely identified.

Another class of word identification model is the search models that propose that a word is searched for by comparing its visual properties (a pre-lexical code) to a lexical code in memory until a match is found. A prominent example of this type of model is the Serial Search model proposed by Forster (1976). The central assumption of the serial search model is that the stages of lexical processing are serial in the sense that a stage only begins if processing in the previous stage has finished. The model consists of four forms of mental representations: a peripheral file, an orthographic file containing the orthographic information about all the words we know, a phonological file containing the sound information about all words, and a master file containing all types of information about the words. When a word is visually perceived, the peripheral file creates a pre-lexical code that resembles the orthographic access code in the orthographic file. Then, searching for the word in the orthographic file begins according to the frequency of the word; high frequency words are searched first. When a word in the orthographic file closely matches the perceived word, the location of this entry in the orthographic file is flagged, and the search process continues until a word in the orthographic file matches the perceived word. At this time, the search is terminated and a pointer in the orthographic file is used to retrieve the word entry in the master file. Upon retrieving the word's entry from the master file, a post-lexical process begins whereby the properties of the perceived word are checked against the properties of the word in the master file. If the properties of the perceived word match the properties of the word in the master file, the word is successfully recognised. Like the logogen model, the serial search model assumes that information about the meaning of a perceived word is activated only after its form has been uniquely identified.

In contrast to the Search Model, activation-class models assume word identification occurs via a process of cascaded activation (i.e., as soon as processing takes place in one stage, activation from that stage flows to the second stage before the processing in the first stage is completed). The activation models also assume that activation from one stage is fed forward to the second stage and fed back from the second stage to the first stage and so on (i.e., activation flows back and forth between stages interactively). One influential example of the activation models is the interactive-activation (IA) model developed by McClelland and Rumelhart (1981). The IA model originally consists of three levels of mental representations that are hierarchically organised: the visual feature level, letter level and word level. The visual feature level consists of units corresponding to the visual features of the letters (e.g., horizontal and vertical lines, edges, corners, etc.). The letter level comprises units corresponding to letters of a language. The word level contains units corresponding to the words that are stored in the lexicon. Each level is connected to the level above it and below it in this model by either excitatory or inhibitory connections. For instance, upon seeing the word *'CAT'*, the letter *'C'* would excite the word units 'CAR', *'CAP'* at the word level, but would inhibit *'MAT'* and *'RAT'*. Excitatory connections between levels make the units at the destination level more active while inhibitory connections make them less active. In addition, each unit is connected with each other in the same level by inhibitory connections. To continue with the example of *'CAT'*, the unit corresponding to the letter *'C'* in the initial letter position would become activated via the lower level at which visual features are represented. This activated letter will increase the activation level of word units at the word level corresponding to *'CAR'* and *'CAP'*, but decrease the activation level of

'MAT' and *'RAT'* because the first letter of *'CAT'* shares the same visual features of the first letter in *'CAR'* and *'CAP'*, while it shares almost no visual features of the first letter in *'MAT'* and *'RAT'*. Since the units within the same level are connected by inhibitory connections, the activation of three letter words without a *'C'* at their beginning such as *'MAT'* and *'RAT'* that are active within the word level will be inhibited. At the same time, the word unit corresponding to *'CAT'* at the word level will receive more activation from the continuous feed forward activation from the lower levels (i.e., more visual features and letter units corresponding to the word over processing time cycles).

Activation also flows back from a higher level to a lower level (i.e., from the word level to the letter level, and from the letter level to the feature level). The feedback activation is assumed to boost the activation of the activated units at the lower level. In the previous example, the activated units corresponding to *'CAT'* (and also to a lesser extent *'CAR*' and *'CAP*') at the word level will feed back to the letter level, facilitating the activation of those letter units corresponding to *'CAT'*. In this system, over time, the pattern of activation settles down into a stable state so that only *'CAT'* remains activated and, thus, is recognised. A factor that is assumed to influence activation at the word level in this model is word frequency. The word units corresponding to higher frequency words have higher baseline levels of activation compared to lower frequency words. Therefore, the activation of a high frequency word inhibits the activation of low frequency words within the word level. As a result, high frequency words are identified faster than lower frequency words, a finding that is well established in both the isolated word recognition literature and the eye movements and reading literature (Rayner, 1998).

The original IA model does not account for the role of word meaning in visual word identification. To account for how word meaning can influence lexical processing, Balota, Ferraro, and Conner (1991) recommended that a fourth layer, a meaning level beyond the word level, be added to the original IA model. Stolz and Besner (1996) took Balota *et al.*'s recommendation and described an embellished IA model using the original processing principles of the IA framework of between-level excitatory or inhibitory connections and within-level inhibitory connections. In this embellished IA model, once the semantic units at the semantic level are activated, the semantic units can give support to the active word-level (orthographic) units through feedback activation from the semantic level to the word level. With this semantic feedback, the meaning level can provide an extra source of activation (Balota *et al.*, 1991). As such, within this embellished IA model there is an assumption that the semantic representations of a set of words will become initially activated by the perceived word, and that this will occur prior to the perceived word's orthographic representation being uniquely identified (Balota, *et al.*, 1991; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). It is also assumed that the feedback from the semantic level can influence the speed by which a word's form is identified (Balota *et al.*, 1991; Stolz & Besner, 1996). To explain, in normal reading, the visual information about the orthographic form of a currently fixated word can activate a set of orthographically similar words (orthographic neighbours). The activation of the word unit corresponding to the currently fixated word will inhibit activation of the word units corresponding to its orthographic neighbours at the word level. Concurrently, activation will feed forward from the word level to the semantic level, activating the currently fixated word's semantic representation. The activation of the semantic representation at the semantic level will feed back to the word level within the period that the candidate set at the word level is being reduced via processes of between-level activation and within-level inhibition. In this way, semantics can constrain unique word identification.

What all these models (Logogen, Serial Search and IA models) have in a common is that orthographic and phonological representations are involved in lexical identification. Many studies have been conducted to investigate the effects of a word's orthographic and phonological representation in its identification, and as such, the effects of orthographic and phonological characteristics in lexical identification during normal reading are well documented. However, the effects of the semantic characteristics of words during normal silent reading are poorly understood. What follows is an overview of the types of lexical information or lexical representations that become available during lexical processing in reading. The effects of these variables on eye movement behaviour during reading will also be considered.

1.3.1 Extracting Orthographic Information

Virtually all models of lexical access acknowledge that letter identification is an important part of the process of word identification. Orthographic encoding was introduced in Section1.1, in which it was explained that information about the letters of a foveal word is extracted during the early stages of word identification. This section extends the discussion on orthographic encoding to consider orthographic influences on the process of word identification. When we visually perceive a word, the word identification system activates not only the actual word we perceive, but also other orthographically similar words (e.g., *mat*, *fat*, *rat*, *bat* become active upon perceiving the word *cat*). These orthographically similar words are called *orthographic neighbours*. To be precise, orthographic neighbours refer to the other same-length words that can be generated when changing a single letter within a word, e.g., *mint*, *pint* and *tint* are neighbours of *hint* (Coltheart, Davelaar, Jonasson, & Besner, 1977*).* According to the IA model introduced in the previous section, orthographically similar words become activated because they share some visual features with the actual word. For example, upon seeing the word *'CAT'*, the visual features of *'C'*, *'A'*, and *'T'* will send activation to letters that share the similar visual features (e.g., the visual features of *'T'* can activate *'T'*, *'P'*, and *'L'*). The activated letters, in turn, will send activation to the word level, activating their corresponding word units that are orthographically similar (e.g., *'CAT'*, *'CAR'*, *'CAP'*, *'GAP'*, *'MAT'*, *'RAT'*, etc.).

The effects of orthographic neighbourhood have been tested by either investigating the effect of the number of orthographic neighbours a word has in its lexical identification and/ or by investigating the effect of high or low frequency orthographic neighbours a word has in lexical processing. Andrews (1997) and Perea and Rosa (2000) reviewed the literature on the effect of orthographic neighbourhoods. From these review articles, it is clear that most early experiments were carried out using lexical decision tasks and fewer eye movement experiments were conducted. Andrews, Perea and Rosa conclude that in English words with many orthographically similar neighbours facilitate lexical decision responses (e.g., Andrews, 1989, 1992; Sears, Hino, & Lupker, 1995), especially when the number of high frequency neighbours is controlled (e.g., Carreiras, Perea, & Grainger, 1997; Forster & Shen, 1996). Such findings suggest that the number of low frequency neighbours drives the facilitatory effect of high number of orthographic neighbours (Pollatsek, Perea, & Binder, 1999). These effects can be explained by the IA model of lexical processing as follows: When a target word is a lower frequency word compared to its orthographic neighbours, its respective (high frequency) orthographic neighbours will have higher baseline levels of activation compared to the activation of the (lower frequency) target word. Thus, the activation of higher frequency neighbours during lexical processing will compete with the activation of the (lower frequency) target word, slowing the identification of the target word. On the other hand, when the target word is a higher frequency word compared to its orthographic neighbours, the target word will have a higher baseline level of activation than its respective orthographic neighbours. Therefore, the activation of a (higher frequency) target word will inhibit the activation of its orthographic competitors, hence, the target word will be activated and recognised faster than its orthographic competitors.

Eye movement studies investigating orthographic neighbourhood effects generally suggest that words with many orthographic neighbours are fixated for a longer time than words with few orthographic neighbours, and that words with high frequency orthographic neighbours are fixated longer than words with low frequency orthographic neighbours (Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). For example, Pollatsek, Perea, and Binder (1999) found that the effect of the number of orthographic neighbours was late and was inhibitory in normal reading. They found that the fixation times immediately after leaving a word with many orthographic neighbours were longer than after leaving a word with few orthographic neighbours in reading tasks. They also found that words with many low frequency neighbours were more likely to be skipped.

The effect of the frequency of orthographic neighbours was also explored by Perea and Pollatsek (1998) in a lexical decision task and a sentence-reading task (eye movements during reading). The findings of the lexical decision task were consistent with the previous visual word recognition research in that words with higher frequency neighbours slowed lexical decision (e.g., Carreiras, Perea, & Grainger, 1997). However, this inhibitory effect of higher frequency orthographic neighbours was not replicated in a normal reading task. During normal reading, higher frequency neighbours did not affect the early eye movement measures (first pass reading measures). However, the inhibitory effect appeared one or two words later after the target word was fixated and as the participants went back to read the target words that had higher frequency orthographic neighbours. Perea and Pollatsek concluded that higher frequency orthographic neighbours did not slow the identification of the actual words when they were first encountered in sentences. They also suggested that the actual words were mis-recognised as their respective higher frequency orthographic neighbours, which lead the readers to refixate the actual words as sentential context cued them that they had mis-encoded the actual words. To explain, a high frequency neighbour will have a higher baseline level of activation and, hence, will be activated more rapidly than the actual (lower frequency) word. As a consequence, the reader may mis-recognise the actual word as its respective higher frequency neighbour.

Similar results were also obtained by Slattery (2009) who also investigated how prior sentential context affected the lexical processing of target words with high frequency orthographic neighbours. Slattery demonstrated that when the prior context instantiated the meaning of the higher frequency orthographic neighbour of the target word, a late inhibitory effect was observed, similar to the effect found by Perea and Pollatsek (1998). However, when the prior context did not instantiate the meaning of a higher frequency orthographic neighbour of the target word, the inhibitory effect was not observed. Thus, these results suggest that prior context can rule out the inhibitory effect of higher frequency neighbours (i.e., mis-encoding the actual words as its higher frequency orthographic neighbour).

In summary, orthographic encoding plays an important role in lexical identification. Particularly, extracting information about a foveal word's letters and its orthographically similar words affects the ease and speed with which a word is identified. Section 1.2 reviewed studies indicating that the first letters of a word are important in the process of word identification, and that the order of letters are not perfectly encoded (since we can still read a word with jumbled letters such as *sturcture* relatively quickly). In this section, the effects of orthographic neighbourhoods were discussed. These findings of orthographic neighbourhood particularly suggest that a word's orthographic neighbours are active during word identification, and generally suggest that a word's orthographic properties influence how its representations are accessed during normal reading.

1.3.2 Extracting Phonological Information

Section 1.3.1 described a set of findings from eye movement experiments consistently indicating that orthographic encoding is crucial in accessing the meaning of a word, and that orthographic information was extracted early in lexical identification. Eye movement data also suggest that phonological processing plays an important role during the early stages of lexical identification, even before the word is fixated (when it is in the parafovea). Pollatsek, Lesch, Morris, and Rayner (1992) used a boundary technique to examine the parafoveal phonological processing. The boundary technique (Rayner, 1975) is a methodological technique used in studies of eye movements during reading. In this technique, the target word (or non-word) that appears in the parafoveal region changes to another word as the reader moves the eyes to fixate it. Pollatsek *et al*. (1992) found that fixation times spent on a target word (e.g., *beach*) were shorter when the preview was a homophone (e.g., *beech*) compared to when the preview was an orthographic control word (e.g., *bench*). Similar results were obtained from other eye movement studies using a fast priming technique[3] (e.g., Lee, Binder, Kim, Pollatsek, & Rayner, 1999; Lee, Rayner, & Pollatsek, 1999; Rayner, Sereno, Lesch, & Pollatsek, 1995). All these studies indicate that phonological information about a word is extracted from the parafovea before it is fixated.

Research on the effects of phonological processing in lexical identification was also carried out to find whether phonological information of words could be accessed via their orthographic representations. The findings of eye movement data showed that fixation times spent on phonologically irregular words (e.g., *pint*) were longer than phonologically regular words (e.g., *tent*), and this phonological regularity effect emerged early in eye movement records (Inhoff & Topolski, 1994; Sereno & Rayner, 2000). In addition, Sereno and Rayner (2000) found the regularity effect was larger for low frequency words. Thus, it seems that skilled readers first access the orthographic rather than phonological representations to access the meaning of the word during normal reading. Phonological representations, on the other hand, are used to access

[3] In fast priming (Sereno & Rayner, 1992), the preview is a word (or non-word) that is unrelated to the target word. When the eyes move across the boundary, this preview is changed into a prime. The prime is displayed for 35ms, and then is replaced by the target word.

the meaning of certain types of words such low frequency words and phonological irregular words (Sereno & Rayner, 2000) and phonologically ambiguous words such as *wind* and *tear* (Carpenter & Just, 1981). Since phonological processing is not central to this thesis, presenting a detailed account of the phonological processing will not be considered beyond the discussion given above.

1.3.3 Extracting Morphological Information

Morphemes are the smallest unit of meaning in language. Some words contain only one morpheme (e.g., *book*), while many longer words contain two or more morphemes (e.g., *encoding* contains three: **en-cod-ing**). Morphemes can be classified as prefixes attached at the beginning of words (e.g., *en-* as in ***en****cod*e), suffixes attached at the end of words (e.g., *–ing* as in *cod****ing***), and root morphemes that can stand alone (e.g., *code* as in *en****cod****ing*). Compound words contain at least two root morphemes (e.g., *basketball*, *blackberry, etc.*). Eye movement data indicate that readers decompose multi-morphemic words into their constituent morphemes, and this morphemic decomposition occurs in the early stages of lexical identification. Evidence for the involvement of morphemic analysis of the constituent parts of the multi-morphemic words in the encoding of these words was established in Finnish studies. These studies (e.g., Hyönä & Pollatsek, 1998; Pollatsek & Hyönä, 2005; Pollatsek, Hyönä, & Bertram, 2000) compared different types of Finnish compound words. For example, Pollatsek and Hyönä (2005) and Pollatsek, Hyönä, and Bertram (2000) used pairs of words that were matched on the frequency of the whole word, but which differed in the frequency of the constituent morphemes. The findings showed that the frequency of the first constituent morpheme influenced first fixation duration on the word with longer fixations on low frequency compared to high frequency morphemes. The frequency of the second morpheme only influenced later fixations. Similar results were also obtained for English compound words (e.g., Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff & Placke, 2003) and for English prefixed words (e.g., Niswander-Klement & Pollatsek, 2006). These results suggest that readers decompose multi-morphemic words into their constituents and this morphemic decomposition occurs early in lexical identification. The results also suggest that a word's initial constituent morpheme is active before the activation of the end morphemes. However, these results are especially true in the case of long multi-morphemic words. Studies suggest that shorter words (including compounds) are not analysed into their constituent morphemes, but instead, can be processed as whole words, while it is necessary to decompose longer multi-morphemic words into their constituent morphemes (e.g., Bertram & Hyönä, 2003; Juhasz, 2008). Central to the focus of this thesis, these studies provide evidence that morphological analysis is fundamentally based upon orthographic processing and that the meaning of a word's constituent morphemes can constrain the identification of a word.

In English, eye movement data suggest that no morphological information is extracted from the parafoveal words while the eyes are still fixating the foveal words (Inhoff, 1989; Kambe, 2004; Lima, 1987). That is, previewing the morphological information of parafoveal words while the eyes are still fixating the foveal word did not affect the fixation times on the parafoveal words when they were later fixated. Eye movement studies investigating English reading compared the preview of parafoveal words with prefixes (e.g., *mistrust*) and pseudoprefixed words (e.g., *mistress*) (Lima, 1987; Kambe, 2004) and compared the preview of compound words (*cowboy*) with a preview of only first morpheme (*cowxxx*) and with the preview of pseudocompound words (e.g., *carpet*) (Inhoff, 1989). The findings showed that prefixed words were fixated for a shorter time than pseudoprefixed words; however, the preview benefit of parafoveal words was not different for prefixed and pseudoprefixed words (Lima, 1987). Similarly, Inhoff's (1989) findings indicated the preview benefit was the same for all three cases. These findings suggest that in English, information about the first morpheme of the parafoveal word is not acquired while the eyes fixate the foveal word.

To summarise, readers extract morphological information from the foveal word during word identification. Particularly, they decompose morphologically complex (foveal) words, especially longer words, into their constituent morphemes, and the processing of these constituents feeds into the processing of the word as a whole. On the other hand, little or no morphological information from the word to the right of fixation in English is extracted during word identification. In the experiments that will be reported in this thesis, the experimental stimuli were short words (with 4 to 7 letters) and were mono-morphemic words (except in Experiment 2). According to the findings from the studies above, the processing of the experimental stimuli in this thesis should not require any morphemic decomposition. Therefore, a fuller account of the issues related to how morphological decomposition occurs, as part of semantic processing, will not be given here.

1.3.4 Extracting Meaning

First fixation and gaze duration on a word not only reflect processes of encoding the visual form, orthographic processing, and phonological processing of a word, but can also reflect processes associated with accessing a word's meaning and evaluating its congruency with the prior context. Eye movement studies investigating lexically ambiguous words (i.e., words that have at least two distinct meanings) provide strong evidence that the fixation time spent on a word also reflects the time to access the meaning of the word, and therefore, that word meaning can influence word identification. In normal reading, lexical ambiguity studies have investigated the differences between the effect of *biased ambiguous words* (i.e., words with at least two distinct, unrelated meanings such as *port* where one meaning is much more frequent than the other) and *balanced ambiguous words* (i.e., words with two unrelated meanings such as *straw* with both meaning have almost equal frequency in language). Duffy, Morris, and Rayner, (1988); Rayner and Duffy (1986) and Rayner and Frazier (1989) showed that when the prior context was relatively neutral, gaze duration on balanced ambiguous words was longer than the gaze duration on single meaning words (control words) matched on length and frequency. They also found that when a disambiguating prior context preceded the balanced ambiguous words, the gaze duration on the balanced ambiguous words was not statistically different from that on the control words as the context constrained the meaning to be activated for the ambiguous words. Gaze duration on biased ambiguous words was found to be relatively similar to the fixation times spent on control words in a neutral context (Duffy *et al.*, 1988; Sereno, O'Donnell, & Rayner, 2006), which indicated that the frequent meaning of an ambiguous word was accessed first when the context was neutral. When the reader encountered disambiguating information that instantiated the less frequent meaning later in the text, readers tended to make long fixations and more regressions to the biased ambiguous words. When a disambiguating prior context that instantiated the less frequent meaning preceded biased ambiguous words, gaze duration (Duffy *et al.*, 1988; Rayner & Frazier, 1989) and first fixation duration (e.g., Sereno *et al.*, 2006) were longer on the biased ambiguous words than the control words (a phenomena known as *subordinate bias effect*).

Other eye movement studies that provide some insight on how word meaning can be extracted in a single fixation are semantic plausibility studies. Semantic plausibility refers to whether the meaning of a sentence as a whole makes sense in terms of real-world knowledge. Whether semantic plausibility has an early effect or a late effect in lexical identification has been a subject of debate. Rayner, Warren, Juhasz, & Liversedge (2004) examined the earliest records of eye movements at which semantic plausibility had an effect on a target word. They presented participants with sentences in three conditions: plausible (*John used a knife to chop the large carrots for dinner last night*), mild implausible (*John used an axe to chop the large carrots for dinner last night*), and anomalous (*John used a pump to inflate the large carrots for dinner last night*). They analysed the early and late measures of eye movements on the target words (*carrot*). The results showed that the mild implausibility influenced only late measures (regression path duration), indicating that it affected the integration of the critical word (*carrot*) into the whole meaning of the sentence. The anomalous condition, however, showed its effect on earlier measures as evident in gaze duration (*carrot* in the anomalous condition was read 20ms longer than in the plausible condition and 17ms longer than in the implausible condition), which suggested that strong semantic anomaly could cause immediate and substantive disruption to reading. The findings of this study were replicated by Warren and McConnell (2007), who compared two sets of sentences describing events that were possible in both sets but the events in one set were plausible and the other were implausible (e.g., possible-plausible: *the man used a strainer to drain the thin spaghetti yesterday evening*; possible-implausible: *the man used blow-dryer to dry the thin spaghetti yesterday evening.*). Warren and McConnell also found the effect of anomaly appeared in early eye movement records (first fixation duration on the target words) and the effect of implausibility appeared in later eye movement measures (regression out and regression path duration) for the target.

However, Staub, Rayner, Pollatsek, Hyönä, and Majewski (2007) reported that the effect of mild semantic implausibility could be present even in the earliest measures of eye movements that were associated with word identification using a more rigorous manipulation of semantic plausibility. The meanings of all sentences were semantically plausible in all conditions. The target words were compounds consisting of two nouns (e.g., *mountain lion*), and the compound was fully plausible in the sentence when the meaning of the sentence is understood as a whole. The researchers created implausibility by manipulating the analysis of the first noun of the compound (implausible head noun analysis condition, novel compound: *Jenny heard the mountain lion pacing in its cage*. Plausible head analysis, familiar compound: *Jenny looked out on the huge mountain lion pacing in its cage*). The findings revealed that there were significant differences between reading the target word (*mountain*) in both conditions as evident in the early measures of eye movements. The study indicated that readers used semantic information such as the semantic information of the verb (*looked out on*) to restrict the anticipation of what came next (*mountain*) and when what came next (*mountain*) violated the semantic restriction of the verb (*heard*), processing and reading the word (*mountain*) was disturbed.

Filik (2008) also examined whether context can modulate the processing disruption associated with an event that would have been implausible or anomalous without the contextual support. Filik compared the fixation times on sentences like (*He glared at/picked up the lorry and carried on down the road.*), where *he* could refer to either a *Hulk* or a man called *Terry*. The findings showed that the effect of implausibility was not observed in the eye movement record on the target word (*lorry*), but was observed one or two words downstream (*and carried*). Processing was not disturbed in the condition in which the Hulk picked up the lorry, suggesting that readers used contextual information during the early stages of processing implausibility. That is, contextual information can prevent disruption to what would otherwise be a violation of world knowledge.

To conclude this section, the two lines of eye movement studies on lexical ambiguity and semantic plausibility emphasise that the meaning associated with a word can be attained relatively early, along with the assessment of the congruency of word meaning with prior context. For the purposes of this thesis, it is obvious that information about the meaning of the fixated word is available during its identification and can affect fixation durations.

1.3.5 Factors Influencing Lexical Processing

The lexical characteristics of a word influence the ease or difficulty with which the word is identified. The previous sections introduced some of these variables such as word length and orthographic neighbourhoods. What follows is a review of some of these variables that have not been mentioned earlier, and that have been well documented to influence eye movement measures associated with lexical processing in normal reading. Therefore, the experiments reported in this thesis will be controlled for these variables by means of statistical control and/ or experimental control. Such control will ensure that any differences in the fixation times on the experimental stimuli will not be due to these extraneous variables, but can be attributed to the manipulated semantic variable of interest in this thesis (i.e., semantic neighbourhood density as will be discussed later in this section).

One of the well-documented lexical factors in the eye movement literature is word frequency (i.e., how often a word is encountered in text as indexed by corpus data). High frequency words were widely reported to be fixated for a shorter time than low frequency words (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Sereno & Rayner, 2000; Schilling, Rayner, & Chumbley, 1998), even after controlling for word length that confounds word frequency. The findings of eye movement studies also demonstrated that high frequency words were more likely to be skipped than low frequency words if these words were short and if eyes landed close to the target (to-be-skipped) word before skipping it (Henderson & Ferreira, 1993; Radach & Kempe, 1993; Rayner & Fischer, 1996; Rayner, Sereno, & Raney, 1996). In addition, the likelihood that a low frequency word receives more than one fixation (or is re-fixated) was observed to be higher than the probability of refixating a high frequency word

(Inhoff & Rayner, 1986; Rayner, Sereno, & Raney, 1996). The word frequency effect seems to attenuate as a function of word repetition in the text. Rayner, Raney and Pollatsek (1995) found that fixation times on low frequency words were shortened with repetition while reading a short passage. They found that there was no difference between low frequency words and high frequency words on the third encounter of the words in the text. In summary, the effect of word frequency emerged early in the eye movement records, suggesting that the word frequency effect influences lexical identification in reading. High frequency words are fixated for less time than low frequency words.

Another factor that has been reported to influence lexical processing is word predictability. A predictability effect appears when the prior context in the sentence provides the reader with a good idea of what the next word would be (i.e., the prior context predicts the target word). Predictability was found to influence only how long a word is fixated, but not where the eyes land in a word (Rayner, Binder, Ashby, & Pollatsek, 2001; Vainio, Hyönä, & Pajunen, 2009). Eye movement studies showed that highly predictable words were skipped more often than words with low predictability (Altarriba, Kroll, Scholl, & Rayner, 1996; Balota, Pollatsek, & Rayner, 1985). Predictable words were also found to be fixated for a shorter time than unpredictable words (Altarriba *et al.,* 1996; Ehrlich & Rayner, 1981; Inhoff, 1984; Rayner, Binder, Ashby, & Pollatsek, 2001). In addition, the findings also showed that readers tended to go back to the highly predictable words less than they tended to go back to the unpredictable words (Ehrlich & Rayner, 1981; Inhoff, 1984) while less predictable words were more likely to be refixated than highly predictable words (Balota, Pollatsek, & Rayner, 1985).

Rayner and Well (1996) manipulated the degree of word predictability (high predictability, medium predictability, and low predictability) to examine the effect of predictability on early and late measures of eye movements. They found that predictability influenced early stages of word identification; high predictable words were fixated shorter and skipped more often than low and medium predictable words. Their findings also revealed that both word predictability and the number of letters a word had played a role in word skipping while word frequency did not.

These results of the effect of word predictability on fixation times were replicated in Drieghe, Rayner, and Pollatsek's (2005) study. One of the interesting findings of Drieghe and colleagues (experiment 1) was that the difference in the probability of skipping was a function of the previous landing position (*launch site*). Predictable words were more frequently skipped than orthographically similar nonwords (different from the predictable word with only a single letter) when the launch site for a saccade came from a close distance to the skipped word. There was no observed difference between predictable words, unpredictable words and nonwords when the saccade came from a far distance to the skipped words. These results suggest that context can speed lexical identification.

The effects of orthographic (e.g., orthographic neighbours and orthographic neighbourhood frequency), phonological (e.g., regularity), morphological (e.g., multi-morphemic words) and contextual factors (e.g., predictability, plausibility and lexical ambiguity) in lexical identification are well researched and understood. These effects have already been discussed in the earlier sections of this thesis (effects of orthographic neighbourhood: Section 1.3.1; phonological effects: Section 1.3.2; contextual effects: Section 1.3.4). As mentioned in Section 1.3.4, lexical ambiguity and semantic plausibility studies suggest that meaning also appears to constrain unique identification. However, both types of studies also used contextual information to investigate whether meaning is involved in lexical processing. Thus, it is not clear whether and how the meaning representation of a word can influence its lexical identification in normal reading. A direct way to study such a question is to investigate the effect of the semantic characteristics of a word in its identification. Such a question has received relatively little attention. Notable exceptions include studies investigating the influence of the number of words associated with the meaning of the word in question (Duñabeitia, Avilés, & Carreiras, 2008) and the number of semantic features related to colours, taste, texture, etc. of the word (Cook, Colbert-Getz, & Kircher, 2013) and the number of contexts in which a word appears (Plummer, Perea, & Rayner, 2014). Generally, these studies demonstrated that words with enhanced semantic characteristics (e.g., high number of semantic associates, high number of semantic features, and high contextual diversity) were fixated for less time than words with weak semantic characteristics, and that their influences appeared in early eye movement records. These studies attributed the facilitatory effect of enhanced semantic characteristics to the enhanced semantic representations a word has because of activating many associates, semantic features or contexts, which feeds back to the orthographic level and boosts the activation of the word in question. A detailed account that can provide explanations for these findings will be given in Chapter 3, 4, 5, and 6.

Another semantic influence that remains completely unexplored with respect to eye movements and reading is the effect of the words that tend to occur with a particular word (i.e., a word's semantic neighbours). Semantic neighbours can be defined as words that are situated in close proximity to each other in texts (first-order- co-occurrence, henceforth) as well as words that have the same co-occurring words in common (i.e., occur with the same other words), regardless of whether they appear in close proximity to each other in text (second-order- co-occurrence, henceforth) (Lund & Burgess, 1996; Shaoul & Westbury, 2012). The primary semantic neighbourhood variable that has been manipulated is how densely the semantic neighbours of a word in semantic space (i.e., the average distance of the co-occurrence neighbours from the target word; *semantic neighbourhood density*, henceforth). In this thesis, it was assumed that semantically similar words (i.e., semantic neighbours) are active during the processes of word identification, and that the density of a word's semantic neighbourhood affects word identification in normal reading. This assumption raises the question of whether a high degree of the semantic neighbourhood density of a given words facilitates or inhibits the processing of the given word during normal reading. To gain insight of the possible direction of this effect in normal reading, the thesis drew on studies that used a different methodological paradigm that resembles the task of normal (silent) reading as will be discussed in detail in Chapter 2.

The effect of a word's semantic neighbourhood density (SND) in its identification has been examined by studies in which participants were asked to respond to a single word in isolation. The findings of these studies indicated that words with denser semantic neighbourhoods were responded to faster than words with less dense semantic neighbourhoods (e.g., Buchanan, Westbury, & Burgess, 2001; Shaoul & Westbury, 2010a; Yates, Locker, & Simpson, 2003). Isolated word recognition tasks are known to exert some experimental demands that are not necessarily part of normal reading

(Kuperman, Drieghe, Keuleers, & Brysbaert, 2013). Since no such demands are required in a task in which participants are asked to read some text silently whilst their eye movements are recorded, one could argue that recording eye movements during normal reading is a nonintrusive method that reflects lexical processing as it happens, and therefore is a better method of examining word identification. Another merit for using eye movements recorded whilst reading is that measures of eye movements index whether the observed effect occurs relatively 'early' or 'late' in word identification. This was the method that was used in this thesis.

How the co-occurrence-based semantic neighbourhood density (SND) would affect normal reading is not yet investigated; therefore, this thesis explores its effect in word identification during normal reading. The thesis will contribute to resolving some issues regarding the SND effects: does the SND effect influence lexical identification processes or task-specific processes (i.e., only limited to lexical decision tasks and other visual word identification tasks)? If the SND characteristics of words influence normal reading, is this effect facilitatory or inhibitory? If the SND effects are found (i.e., if the fixation times on target words are affected by SND), then one has clear evidence that the SND effects are not restricted to laboratory word identification tasks, but is actually influencing lexical identification during normal reading. Such findings will lend support to the assumption that word meaning can constrain the unique identification in normal reading. All of these issues will be further discussed in Chapter 2, 3, 4, 5, 6, and 7.

1.3.6 Summary

Identifying individual words in normal reading should occur first in order for other, later processing (e.g., understanding the structural relationships between individual words and constructing the meaning of the whole sentence) to occur since this later processing depends on lexical characteristics that become available via lexical identification. It was discussed that fixation times on a word reflect accessing its orthographic and phonological information as well as accessing the meaning of the word as evidenced from the findings of lexical ambiguity and semantic plausibility studies. The description of lexical processing in reading also involved discussing the well-founded variables such as word frequency and predictability that were consistently found to influence the fixation times spent on target words. It was also discussed the effects of semantic characteristics associated with currently fixated words have received little attention in the eye movement and reading literature and that the present thesis will investigate the effects of semantic neigbourhood density in lexical identification during normal reading.

1.4 Summary and Conclusion

In this chapter, an account of some aspects of language processing in skilled reading was given. Language processing entails some consecutive, coordinated sub-processes, and these processes begin with the process of visual encoding of the printed word in order that full word identification may then take place. All models of word identification agree that a printed word's lexical identification involves retrieving its orthographic and/ or phonological representations stored in the long-term memory. However, these models differ in their assumptions about the influence of a word's semantic representation in its lexical identification. Some models assume that a word's meaning is retrieved only after its orthographic form has been uniquely identified, while other models assume that semantic information about a word can constrain the unique identification of its form.

In particular, the focus of this chapter was on lexical processing that is contingent on visual processing. Reviewing abundant eye movement data showed that the linguistic factors such as word frequency, word predictability and lexical ambiguity influence the decision of when and, to a lesser extent, where to move the eyes during reading. While the influences of many orthographic, phonological and morphological factors in lexical processing are well documented in the field of eye movements during, the influence of the semantic characteristics of a word in its identification during normal reading is less well understood.

In this thesis, it was assumed that a word's semantic representation could affect its unique identification during normal reading. Specifically, it is assumed that semantically similar words (i.e., semantic neighbours) are active upon the activation of a word's orthographic representation. It is also assumed that density of a fixated word's semantic neighbourhood can constrain the identification of its orthographic form in normal reading. These assumptions raise the question of whether semantic neighbourhood density might facilitate or inhibit the identification of a word in normal reading. While this is unexplored with respect to eye movements and reading, there is a large literature related to semantic neighbourhoods and their effects using tasks that do not necessarily reflect normal processing that occurs during natural reading. In the next chapter, the theoretical foundation of semantic neighbourhood density effects will be reviewed, an overview of the computational models that are used to arrive at a words' semantic neighbourhood will also be given, and some studies that investigated the effects of semantic neighbourhood density on isolated word identification will be reviewed.

Chapter 2: Semantic Neighbourhoods

Chapter 1 introduced some models of word identification that differ in their assumptions about how the meaning of a word can influence the unique identification of its orthographic (or phonological) forms. It was discussed that some models assume that a word's meaning is activated only after its form has been uniquely identified (e.g., Forster, 1975) while other models assume that a word's meaning can be activated before the competition between orthographic representations (those of the actual word and its orthographic neighbours) is resolved (e.g., Stolz & Besner's (1996) embellished interactive-activation model (McClelland & Rumelhart, 1981)). In normal silent reading, the influence of a word's semantic characteristics in the identification of its orthographic form is under researched. One reason for this under-researched area is that researchers differ in their views of what constitutes the meaning of a word, unlike their views of the constituents of the orthographic and phonological representations. To explain, semantic representations have been defined in different terms of: (1) observable properties of objects (e.g., colour, taste, shape, etc.), (2) semantic categories (e.g., animal, fruit, bird, etc.), (3) semantic associations (e.g., *hair- brush*), or (4) co-occurrences in text (e.g., *cat*, *kitten*). Thereby, words can be semantically related (i.e., semantic neighbours) in different ways (Buchanan, Westbury, & Burgess, 2001; Mirman & Magnuson, 2008).

Chapter 2 will thoroughly describe the concept of semantic neighbourhoods in terms of contextual co-occurrence, highlighting the theoretical basis of semantic neighbourhoods and reviewing the isolated word recognition studies that explored the semantic neighbourhood effects in word identification. Section 2.1 will briefly give background information on the theoretical context concerning the effects of semantic neighbourhood density. In so doing, this section will describe the two major views of the nature of semantic representations, the object-based view and the language-based view. Section 2.2 will focus on one of the language-based theories, in particular, the co-occurrence-based theory of semantic representations. This section will describe the theoretical foundation of co-occurrence-based semantic representations and how computational models are used to capture word meaning. This section will also give an overview of two computational models, one of which will be used in this thesis to define the semantic neighbourhoods of the target words that appeared in the experimental stimuli used in the experiments reported in this thesis. Section 2.3 will review the empirical studies that have been conducted on the effects of semantic neighbourhoods, focusing on the studies that used co-occurrence-based definitions of semantic neighbourhoods. Section 2.4 will address how the issues presented in this chapter (and also Chapter 1) are relevant to the current work in this thesis and will reinstate the purpose of the current work. It will also present the research questions that are central to this thesis. Section 2.5 will summarise semantic neighbourhoods and will state the structure with which the remainder of this thesis is organised.

2.1 Views of Semantic Representations

Our daily lives, as humans, are full of tasks that require exploiting world knowledge that we have accumulated throughout our lifetime. This world knowledge includes, for example, information about how to perform some functions in our daily lives such as driving cars, eating fruit and vegetables and information about the behaviour of some creatures such as the barking of dogs. Based on our cumulative experiences, we are able to extract and store such knowledge about the world in *semantic memory*. Semantic memory refers to human memory of word meaning and includes many types of information about concepts (McRae, 2004). For example, from our past encounters with the concept *cat*, we know that it is an *animal*, it has *whiskers* and it is related to other concepts such as *kitten* and *tiger*. We also know from our experiences that the meaning of the word *test*, for example, is related to the meaning of *experiment*, *trial*, *quiz*, and *exam*. These meanings we know about the words from our past experiences are what can be referred to as the semantic representations of the words.

There have been many different views and models explaining the nature of semantic representations and how they are stored and retrieved. Virtually, all developers of semantic memory models to a great extent agree that the humans' semantic system exhibits a general structure and some regularity that is assumed to be shared by individual humans; they also acknowledge that individual differences to a lesser extent may influence the semantic structure (Buchanan, Westbury, & Burgess, 2001). The models of semantic representations attempt to capture the structural regularities in either the objects found in the world or the structural regularities in the relationships between words found in language. Accordingly, these models represent two major views based on the type of information they stipulate about word meaning, object-based theories and language-based theories. The object-based view represents word meaning in terms of some observable properties or features (e.g., colour, taste, smell, etc.) or categories (e.g., animal, plant, bird, etc.). The language-based view defines the meaning of a word in relation to other words in language (i.e., how a word is used in language); words can be related to each other by means of associations (i.e., words that are semantically associated with each other, e.g., *hair* and *brush*) or by means of co-occurring in similar contexts in text (e.g., *movie* and *game* both appear in the contexts of entertainment and enjoyment). What follows is a brief description of the object-based view, highlighting the issues related to the feature-based view, which is considered a representative view of the object-based theories. Then, this description will be followed by a brief comparison to the language-based theories.

The feature-based view of semantics postulate that the meaning of a word (especially concrete words) comprises multiple types of knowledge including: visual knowledge (e.g., *shape*, *size*, *colour*, *characteristic motion*), knowledge associated with sounds that the objects/entities produce (e.g., *loud*, etc.), how they smell (e.g., *smelly*, *smells nice*, *is scented*), taste (e.g., *musty*, *sweet*, *sour*, etc.) and feel (e.g., *hard*, *damp*, *cold*, etc.). Also, the meaning of a word includes knowledge about the typical behaviour of creatures (e.g., *meows*, *barks*, etc.), situational/event-based knowledge (e.g., what the objects are used for such as *cutting*, where they can typically be found such as *in the kitchen*, and who typically used them such as *farmers*).

The above-mentioned types of conceptual representations of words are called *semantic features* (McRae, Cree, Seidenberg, & McNorgan, 2005; Rosch & Mervis, 1975). The feature-based conceptual representations are empirically derived from feature production norms using feature-listing tasks. In such tasks, human participants are asked to list the semantic features of some basic-level concepts (e.g., *orange*, *cat*). Recently, McRae, Cree, Seidenberg, and McNorgan (2005) developed feature production norms for 541 basic-level concepts of living and non-living things over three years. In their norms, each subject enlisted the features of 20 or 24 concepts and each concept was presented to 30 participants. For each concept in their feature production norms, information is provided about the number of features, number of distinguishing features (e.g., *barks*, *meows*), ratings of the distinctiveness of the features, the likelihood that a feature would appear in a certain concept, distributional statistics about feature correlations (i.e., the tendency of two features to occur in the same basic-level concepts). Under the feature-based view, words are considered semantically similar (i.e., semantic neighbours) if they have several overlapping semantic features. In this sense, *robin* and *canary* are thought of as semantic neighbours because they share some features in common such as '*bird'*, '*can fly'*, '*small'*, '*can sing'*, etc.) (Cree & McRae, 2003).

Featural representations were hypothesised to underlie our implicit statistical knowledge of feature correlations and our explicit theory-based knowledge (Holyoak & Spellman, 1993; Lin & Murphy, 1997). To explain, both humans and connectionist networks were found to naturally encode the extent to which certain pairs of features co-occur across concepts (Cree, McRae, & McNorgan, 1999; McRae, de Sa, & Seidenberg, 1997; McRae, Cree, Westmacott, & de Sa, 1999), suggesting that our brain's neurons (and those of the connectionist network) learn correlations (Saffran, Aslin, & Newport, 1996; Jusczyk, Cutler, & Redanz, 1993). For example, people tend to list feature pairs such as <*has wings*> and <*has feathers*> when they are asked to list the features associated with some concepts such as *robins*, *pigeons* and *canary* (McRae, 2004). Also, people sometimes have explicit theory about why two features co-occur based on the relationship between features. To illustrate, people reported that <*has wings*> was causally related to <*flies*> (Ahn, Marsh, Luhmann, & Lee, 2002; Murphy & Medin, 1985).

The conceptual representations derived from the feature production norms could account for some experimental phenomena such as semantic similarity priming (Cree, *et al.*, 1999; McRae *et al.*, 1997), feature verification (Solmon & Barsalou, 2001; McRae *et al.*, 1999), categorisation (Smith, Shoben, & Rips, 1974) and conceptual combination (Smith, Osherson, Rips, & Keane, 1988). As such, the feature production norms have been argued to provide some insight into important aspects of word meaning (Medin, 1989; McRae, 2004).

Some major challenging questions, however, were raised against the feature-based view of semantic memory. One challenging question is how the various semantic features that represent the meaning of a concept bind together as a coherent unified whole, so that the features are effortlessly perceived as being aspects of a single concept (Roskies, 1999; von der Malsburg, 1999). For example, how the shape feature of an object binds with the feature of the location of that object so that both shape and location features provide a unified representation of the object. To solve this binding problem, a number of hypotheses were proposed including von der Malsburg's (1999) temporal synchrony of neuronal firing rate (i.e., temporally synchronising the activity of different neurons), Simmons and Barsalou's (2003) hierarchy of convergence zones (i.e., a set of processing units that encode activity among multiple input units, Damasio, 1989), and Patterson, Nestor, and Rogers' (2007) single convergence zones. It is beyond the scope of this thesis to review these hypotheses. To date, it is not clear which solution is most viable (McRae & Jones, 2013).

Another question that was raised against the feature-based view is how words that have no observable or physical properties such as abstract words and sophisticated verbs are represented in terms of semantic features. The feature-based view is mainly based on research that has been carried out on concrete words and observable actions (verbs) (McRae & Jones, 2013). The feature-based representation view cannot sufficiently capture the meaning of abstract concepts (Shallice & Cooper, 2013), particularly as abstract words do not have sensory referents in the world (Paivio, 1986). In an attempt to provide a resolution to this issue, some researchers suggested that the cognitive organisation of abstract concepts might be partially different from the cognitive organisation of concrete concepts (Plaut & Shallice, 1993). Thus, Plaut and Shallice along with other researchers proposed some models that instantiated the meaning of abstract words and sophisticated verbs using a mechanism that was different from the mechanism used to instantiate the meaning of concrete words (McRae & Jones, 2013). To date, however, there have not been any feature-based models that specify one common mechanism for constructing the semantic representations for all types of words.[4]

Another limitation of the feature-based view is that the feature production norms were collected for a few hundred words that were mostly concrete nouns and observable actions. This is due to the hand-coding (and laborious) nature of the feature-listing tasks that were used to develop the norms. This seems to be an obstacle for the feature- based view that provides the semantic representations for a limited number of words and limited topics. All these issues are resolved when considering the language-based models, especially those that derive semantic representations from large-scale text corpora instead of embodied experiences collected from participants.

In contrast to the object-based view, the language-based view postulates that word meaning does not have to be represented in terms of structural regularities of the semantic features of words themselves. Instead, the language-based view captures word meaning by the patterns of word usage in language (i.e., how a word is used in relation to other words in language). As such, the history of a word's usage in language is what gives the word its meaning. The history of word usage can be derived either from tasks whereby participants are asked to write the first related word that came into mind when they saw another word (association-based semantics) or from recording the systematic patterns of the words that co-occurred around a given word (co-occurrence-based semantics). Two words are considered semantically similar (or semantic neighbours) under this view if they are semantically associated with each other (association norms, Nelson, McEvoy, & Schreiber, 1998) or if they appear in similar contexts in large samples of text (i.e., their distributions in text or their global co-occurrence, Lund & Burgess, 1996). The associational-based semantic representations, just like in the object-based view, were developed using tasks to elicit associations from human participants. In such tasks, the number of distinct responses that two or more participants enlisted is tallied (e.g., the association norms developed by Nelson *et al.*, 1998). Because of the (hand-coding) nature of the tasks from which associations were

[4] It is worth mentioning that McRae and colleagues are currently conducting research on 'the structure and the content of abstract concepts in the human mind and brain' (April 2012- March 2017), which will give more insight on the organisation of abstract words under the feature-based view.

developed, associates were collected for a limited number of words (5,019 words in Nelson *et al.*'s (1998) association norms), which is one of the limitations of the associational-based semantic representations. For this reason, the global co-occurrence-based semantic (distributional-based view) can be argued to be more successful in terms of deriving semantic representations for millions of words from large corpora of text. This thesis will focus on co-occurrence based semantic representations, and the remainder of this chapter will be largely devoted to understanding this view.

2.2 Distributional Semantic View

The distributional semantic view stipulates that the representation of the meaning of a given word can be derived from the other words that tend to co-occur with in large samples of text (i.e., its distribution in text). These other words that co-occur with a given word are called the *semantic neighbours* of the word. Precisely, semantic neighbours under the distributional semantic view can be defined as words that are situated in close proximity to each other in texts (*first-order co-occurrences*, henceforth) as well as different words that have in common the same words that co-occur with them, regardless of whether they appear in close proximity to each other in a text (*second-order co-occurrences*, henceforth) (Lund & Burgess, 1996; Shaoul & Westbury, 2012). To illustrate, consider the phrase in the previous sentence, '*situated in close proximity to…*'. '*Situated'* and '*proximity'* appear close to each other in this phrase, and as such they are thought of as first-order co-occurring words. Now consider another phrase, '*situated in a close location to…*', for the sake of argument. You will notice that both '*proximity'* in the former phrase and '*location'* in the latter phrase appear in similar linguistic contexts (i.e., share some co-occurring words), and therefore they are considered second-order co-occurring words.

Having very briefly introduced the concept of semantic neighbours under the distributional semantic view, the remainder of this section will specify the theoretical grounding of this view that will be necessary to understand the rest of this thesis. Then, how distributional semantic models are generally built will be laid out followed by a description of two examples of distributional semantic models. Understanding these distributional semantic models will be necessary to grasp the idea of how words can be semantically related and how a word's semantic neighbourhood is derived under this view, which, in turn, is necessary to understand the research undertook in this thesis.

2.2.1 Distributional Semantic View: Theoretical Foundation

The distributional semantic view is theoretically originated in structural linguistics and is motivated by *distributional methodology* (Harris, 1954) that postulates that if two linguistic units (e.g., unit *A* and unit *B*) both occur with a third linguistic unit *C* (i.e., *A* and *B* have similar distributional properties), then *A* and *B* are considered related. This distributional methodology was later extended to theorise about semantic representations so that the meaning of a given word depends on the aspects of meaning shared between the given word and the words that comprise the contexts in which it appears. According to the distributional hypothesis, two words are similar in meaning if they appear in the same contexts (i.e., appear with same neighbouring words). That is, the degree of semantic similarity between two words can be seen as a function of the overlap among their linguistic contexts (i.e., words that co-occur with in a language). In this way, semantic similarity is linked to co-occurrence (or distributional) similarity as Harris stated.

'The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear' (Harris, 1954, pp. 2-3).

'If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different than the distributions of A and C. In other words, difference in meaning correlates with difference in distribution' (Harris, 1954, pp. 2-3).

To explain these quotes, the similarity or difference in the meanings of words is reflected in the words' distributions (i.e., the words that co-occur with) in a large text. As such, if two words occur frequently in similar contexts, it is more likely that these two words are similar in their meanings (Firth, 1975). For example, the word *movie* may appear in the context of (i.e., co-occur with) *enjoy* and *watch*. It is, therefore, argued that it can be inferred that these words are semantically similar – at some level, they share some aspect of meaning. It also can be inferred that *movie* is similar to other words such as *game*, itself a words that appears in the context of words like *enjoy* and *watch*, even though *game* may not co- occur with *movie*.

To summarise, the distributional semantic view relies on the co-occurrences found in a text corpus to construct semantic representations. Under the co-occurrence-based view, two words are considered semantic neighbours based on their co-occurrences in similar contexts in a large-scale text corpus; the contexts in which a word appears entail some important aspects of its meaning.

2.2.2 Distributional Semantic View: Modelling

Some distributional semantic models are used to derive semantic representations by analysing a text corpus. Before explaining how the models of distributional semantics were built to do this, it is important to lay down some basic terminologies used in the literature of distributional semantic models.

2.2.2.1 Distributional Semantic Modelling: Terminologies

Semantic space is a space that is used to spatially represent word meaning as presented in Figure 2.1. As can be seen from this figure, words are represented as points in this space. The distance between one point (i.e., a word) and another reflects the degree of semantic similarity between the two words. Words that are close to one another in this space are considered semantically similar. So, what is actually being modelled in semantic space is the semantic similarity between words as a function of their proximity from one another in an *n*-dimensional space where *n* can reflects the number of dimensions (i.e., the number of co-occurrence words). Typically, the number of dimensions used in the distributional semantic models is very high (e.g., 100000 dimensions). In Figure 2.1, only a two-dimensional space is visualised for simplification.

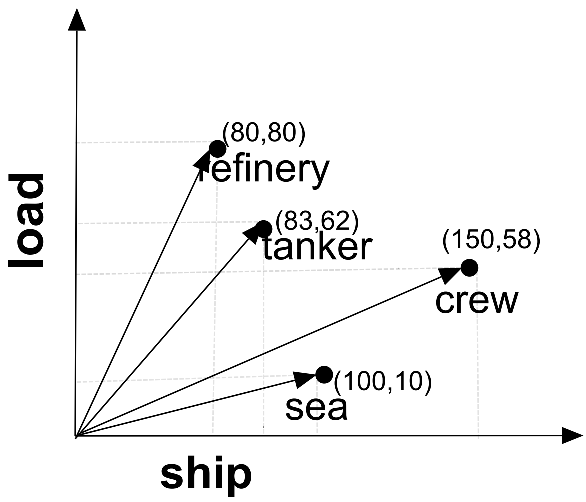

Figure 2.1 A geometric representation of a hypothetical two-dimensional space. The words (*refinery*, *tanker*, *crew*, and *sea*) are represented as points in two dimensions (i.e., co-occurring words) of *load* and *ship*. The spatial proximity between words reflects how the words are close or similar in their meanings. For instance, in this space *tanker* is close to *refinery* while it is relatively distant from *sea*. Therefore, one can infer that the meaning of *tanker* is more similar to the meaning of *refinery* than to the meaning of *sea*.

To arrive at the geometric representations of semantic space (as illustrated in Figure 2.1), distributional semantic models are used to first collect distributional information (profiles) for words in a *matrix of co-occurrence counts* (see Tables 2.1 and 2.2), and then transform such distributional data to geometric representations. The distributional information of a word refers to the '*the sum of all its environments*' (Harris, 1970, p.775). The environments of a given word can be the words that surround the given word in a line, sentence, or phrase (i.e., neighbouring words), or it can be the documents in which the word appears. Thus, the distributional semantic models are used to populate a *word-by-word matrix* or *word-by-document matrix*. A word-by-document matrix is used to assess the relationships between words and number of documents in which they appear (i.e., the similarity between documents) while a word-by-word matrix is used to directly measure co-occurrences between different words (i.e., the similarity between words). Since the focus of this thesis is on the effect of the degree of similarity between a word and the other words that co-occur with in language, distributional semantic models that are used to produce word-by-word matrices (i.e., that defines the context/ environment in which a word appears as its neighbouring words in text) will be central to this thesis and will be further discussed in the remainder of this chapter.

To give a simple example of a word-by-word matrix, consider the example of, *'Tankers offload oil to refineries'*. If we consider the context of a target word as one word ahead and one word behind the target word, then the context of *'offload'* will be *'tankers'* and *'oil'*. Producing a co-occurrence matrix for the previous sentence according to a one-word ahead and one-word behind criterion should look like the co-occurrence matrix presented in Table 2.1.

Table 2.1 *A One-Word Ahead and One-Word Behind (Raw) Co-Occurrence Matrix.*

	Co-occurring Words				
WORDS	tankers	offload	oil	to	refineries
tankers	0	1	0	0	0
offload	1	0	1	0	0
oil	0	1	0	1	0
to	0	0	1	0	1
refineries	0	0	0	1	0

A distributional semantic model is used to build a co-occurrence matrix based on large samples of text from a large-scale corpus of hundreds of millions or billions of words. Thus, after summing the co-occurrence counts for each word in the corpus and after applying some mathematical and statistical techniques that will be discussed later in Section 2.2.2.2 in this chapter, the resultant co-occurrence matrix will be somewhat similar to the simplified co-occurrence matrix presented in Table 2.2. This table shows that the word *'tanker'*, in this hypothetical example, co-occurs 83 times with the word *'ship'*. The rows in this matrix represent target words and the columns represent contexts (i.e., dimensions) or words that co-occur with the target word in text.

Table 2.2 *A (Hypothetical) Co-Occurrence Matrix. In this Matrix, 'tanker' Co-Occurs 83 Times with 'ship' and 62 Times with 'load'. Also, 'tanker', 'oil' and 'refinery' Have Similar Co-Occurrence Counts in Each of the Three Dimensions ('ship', 'load' and 'carry').*

	Co-occurring Words			
	carry	load	ship	....
crew	54	58	150	....
oil	61	58	85	....
refinery	50	80	80	....
sea	4	10	100	....
tanker	67	62	83	....

A co-occurrence matrix consists of distributional *vectors* containing the values found in the cells of a row. For example, the distributional vector of *'tanker'* in the co-occurrence matrix presented in Table 2.2 is X_{tanker} = (67, 62, 83, …). Each value in the vector is called a *dimension* or *feature*. To reiterate, the values in the matrix represent co-occurrence counts (frequencies; e.g., the number of times *tanker* co-occurs with *ship*). Thus, each value in a vector specifies one attribute or characteristic of the word in the space. The vector of a word specifies the location of the word in an *n*-dimensional space (usually a very high-dimensional space, say, with100000 dimensions as mentioned before). However, knowing that the location of the word *tanker* is (67, 62, 83) in a three-dimensional space, for example, is not informative of anything, except its location in semantic space. As such, knowing the location itself is meaningless. When we consider the location of a word (e.g., *tanker)* in relation to its proximity to the locations of other words (e.g., *refinery*, *oil*, *sea*, etc.) in semantic space, then these locations become meaningful with the respect to specifying which words are closer and therefore semantically similar to a given word than other words. The ultimate goal of a word-by-word distributional semantic model is to represent semantic similarity between words, by spatially modelling word meaning, in terms of the proximity between words in a high-dimensional semantic space.

To measure how similar or different the meanings of words in a high-dimensional semantic space, similarity or distance measures are used. Similarity measures indicate how similar two vectors are and give high scores for similar vectors. Distance measures, on the other hand, indicate how different two vectors are, and give low scores for similar vectors. One example of the similarity measures is *cosine similarity*, which is the angle between two arrows/vectors (see the angles between the vectors in Figure 2.2). In Figure 2.2, the angle between *'tanker'* and *'sea'* is larger than the angle between *'tanker'* and *'refinery'*. As such, the cosine angular distances indicate that vectors of *'tanker'* and *'refinery'* are more similar than the vectors of *'tanker'* and *'sea'*. Cosine similarity measures how similar two vectors in a scale of [0,1], where 0 indicates no similarity and 1 indicates maximal similarity. An example of the distance measures is the *Euclidean distance*, which measures the straight distance between two points (see the dashed lines between the vectors in Figure 2.2). In Figure 2.2, the Euclidean distance between *'tanker'* and *'sea'* is larger than the distance between *'tanker'* and *'refinery'*. As such, the vectors of *'tanker'* and *'sea'* are more different than the vectors between *'tanker'* and *'refinery'*. It is worth mentioning that the results of applying similarity and distance measures are equivalent if the vectors are normalised (as will be explained in Section 2.2.3.1 and 2.2.3.2); both distance and similarity measures give similar account of how close two words are in semantic space (Evert & Lenci, 2009).

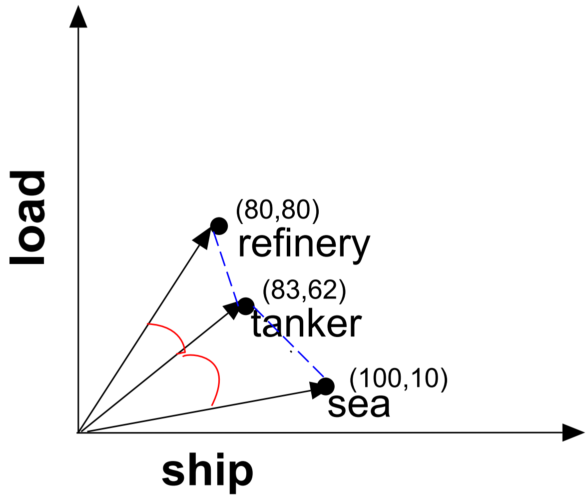

Figure 2.2 A (hypothetical) two- dimensional semantic space; in this space, the vectors of three words (*'tanker', 'refinery'* and *'sea'*) are geometrically represented in terms of their co-occurrences with two dimensions (*'ship'* and *'load'*). In this hypothetical example, *'sea'* co-occurs 100 times with *'ship'* and 10 times with *'load'*. The illustration of this space also shows that words that have similar values in the same dimensions are located close together in the space. For example, both *'tanker'* and *'refinery'* have similar values of 80 and 85 respectively in the dimension of *'ship'* and 62 and 80 respectively in the dimension of *'load'*. Thus, the vectors of *'tanker'* and *'refinery'* are much closer to each other in this space compared to *'sea'* that has very different values in these two dimensions. The Euclidean distance between *'sea'* and *'tanker'* (the dashed line) is larger than the distance between *'refinery'* and *'tanker'*. Also, the cosine angular distance (the angle) between *'sea'* and *'tanker'* is larger than the cosine angular distance between *'refinery'* and *'tanker'*.

As can be seen from Figure 2.2, the angle and the distance between *'refinery'* and *'tanker'* is less than the angle and the distance between *'sea'* and either *'refinery'* or *'tanker'*. Therefore, it can be seen that the distributional profiles of *'refinery'* and *'tanker'* are similar in this semantic space and, hence, the words *'refinery'* and *'tanker'* are semantically similar while *'sea'* is less related to any of these words. Thus, the distance between two vectors indicates how similar the contexts of usage of the two words represented by the vectors are.

2.2.2.2 Distributional Semantic Modelling: Building Steps

Building a word-by-word distributional model generally involves three main steps. The first step involves selecting the corpus from which co-occurrence information is extracted. In the second step, the corpus is linguistically processed so that it can be used by the model. This step involves detecting and eliminating unwanted text (e.g., removing non-English documents from the corpus), converting all words in the corpus to upper case letters so that the differences in capitalisation is eliminated (e.g., *Door* and *door* are converted to *DOOR*), adding a space to separate the possessives (*'s*) from the words, and replacing the hyphens in the hyphenated words with a space (e.g., *first-class* is converted to *FIRST CLASS*). In this way, the linguistic processing helps the model to detect identical words (e.g., *Door*: *door*; *first-class*: *first class*), and treat them as equivalent.

In the third step, mathematical and statistical processing for the linguistically processed corpus takes place. The mathematical and statistical processing involves building a matrix of co-occurrence frequencies, weighting the co-occurrence frequencies, smoothing the matrix by reducing its dimensionality, and measuring the similarity or distance between vectors. The basic mathematical and statistical processing is explained below.

To build a co-occurrence matrix, the number of times another word co-occurs with a target word is counted (e.g., how often does *'tanker'* occur in the context of *'load'*?). Words are considered to have co-occurred with a target word if they appear immediately adjacent to the target word as well as if they are separated from the target word by a number of intervening words in a line of a written text. The maximum number of intervening words that are considered to co-occur with a target word is called *window size*. The frequencies of co-occurrences are populated one window at a time in a way that the window slides forward one word at a time until all words in the corpus are processed (see Figure 2.3). Distributional models differ in the type of window (e.g., windows of words, sentences, paragraphs, or whole documents); the discussion in this thesis will be limited to only word-based distributional models as mentioned before. The models also differ in the size of the window (i.e., how many words fall in the window) and its extension (i.e., how many words to the left and to the right of the target word).

By recording every window movement, the co-occurrence matrix is compiled. For every target word in this matrix, there is a row and some columns as presented in Table 2.1 and 2.2. If the window is used to count co-occurrences symmetrically in the both directions (to the left and right of the target word) within the window, then the resultant word-by-word co-occurrence matrix is symmetric in the sense that the rows and columns for a target word both contain the same co-occurrence counts. If the window is used to count co-occurrences in only one direction (to the left or right words from the target word), then the resultant matrix is directional in the sense that the rows and columns contain co-occurrence counts in different directions. To explain, a left-directional co-occurrence matrix gives co-occurrence counts with the preceding (left in English) words within the window; the values in a row are co-occurrence counts of the target word with the left words in the window while the values in a column are the co-occurrence counts with the right words in the window. A right-directional co-occurrence matrix populates counts of co-occurrences with the succeeding (right in English) words within the window; a row contains co-occurrence counts of the target word with the right words within the window while a column contains counts of co-occurrence counts with the left words within the window.

It should be noted that the above-described directional information is discarded in the final stages of applying an algorithm to concatenate the row and column vectors (Sahlgren, 2006). Thus, it does not matter whether a directional or symmetric word-by-word matrix is used, as it will be the case that words that have occurred with the same other words in a particular corpus being analysed will have similar representations when comparing their vectors (Sahlgren, 2008).

		5	4	3	2	1		1	2	3	4	5	
and	criticism	are	obvious	parts	of	any	$\boxed{\text{interactive}}$	teaching	materials	but	the	balance	must

			5	4	3	2	1		1	2	3	4	5
and	criticism	are	obvious	parts	of	any	interactive	$\boxed{\text{teaching}}$	materials	but	the	balance	must

AHEAD	are	obvious	parts	of	any	interactive	teaching	materials	but	the	balance	must
interactive	0	0	0	0	0	0	5	4	3	2	1	0
teaching	0	0	0	0	0	0	0	5	4	3	2	1

BEHIND	are	obvious	parts	of	any	interactive	teaching	materials	but	the	balance	must
interactive	1	2	3	4	5	0	0	0	0	0	0	0
teaching	0	1	2	3	4	5	0	0	0	0	0	0

Figure 2.3 A visualisation of a sliding window (five words ahead and five words behind the target word) with inverse linear ramp weighting. In this example, the first target word is the word *'interactive'* and the second target word is the word *'teaching'*. The tables below show the vectors that appear ahead and behind the target words; these vectors would be contained in the co-occurrence matrix after weighting the counts from the sliding window (but before normalising the rows) (based on Shaoul & Westbury, 2012).

Once the co-occurrence matrix is populated from the whole corpus, the co-occurrence counts in the cells are weighted using a weighting scheme that assigns weights to the context words based on their distances from the target word in the window. Applying a weighting function involves multiplying co-occurrence frequencies by a number reflecting the distance of the context word from the target word in the window. One of the weighting functions used in some distributional models is called *linear ramp*, which gives more weight to the co-occurrence neighbours that are located closely to the target word. To illustrate, consider that we have a five-word window (i.e., five words to the left and five words to the right of the target word). If a linear ramp function is applied as a weighting scheme to the co-occurrence counts, then the co-occurrence frequency of the neighbouring word that appears directly close to the target word in either direction will be multiplied by five in this five-word window. The co-occurrence frequency of the next neighbouring word out in either direction will be multiplied by four, while the co-occurrence frequency of the word that appears at the edge of the window in either direction will be multiplied by one. Another type of weighting functions that is used in other models is the *inverse linear ramp*, which gives more weight to the co-occurrence neighbours that are located far away from the target words. Implementing the inverse linear ramp as a weighting scheme gives less weight to the closer words that often tend to be function words, and therefore, high frequency words. Since function words convey little semantic information, the raw co-occurrence statistics could simply reflect frequencies that are correlated with syntactic functions along semantic relationships between words (Durda & Buchanan, 2008; Rhode, Gonnerman, & Plaut, 2005). Through the application of this type of inverse weighting schemes, the effect of function words is minimised (for a brief review of other types of weighting functions, see Shaoul & Westbury, 2010a).

The weighted co-occurrences are then stored in a raw co-occurrence matrix that contains the weighted frequencies of co-occurrences for all possible combinations of words in all possible positions in the window (before and after the target word in the window) (see Figure 2.4, Panel (C) for a simple visualisation of such matrices). At this point, due to passing the sliding window over a large corpus, the consequent weighted (raw) co-occurrence matrix is very large, which can be computationally laborious and impossible, and also very sparse at the same time since most words rarely co-occur with each other in a corpus (i.e., most of the cells in the co-occurrence matrix will contain co-occurrence counts of zeros).

To solve the issues of the high-dimensionality and the sparseness of the vectors of the data, the sparse (raw) co-occurrence matrix is compressed by reducing its dimensions (columns). Dimensionality reduction is achieved by filtering out some words in the matrix based on linguistic or statistical criteria. Filtering out words based on linguistic criteria involves removing the words that belong to a closed grammatical class (e.g., function words) as these words are assumed to have little semantic information. These closed class words, constituting a small number of words in language, have orthographic frequencies (i.e., how often the words appear in a corpus) much higher than the orthographic frequencies of the rest of open class words (Bayaan, 2001;

Zipf, 1949). Accordingly, the closed class words have very high co-occurrence frequencies with almost all words in the corpus. Therefore, the vectors of closed class words are very dense with large values, making them much closer to all other words than low frequency words. The problem with the linguistic criterion as a form of dimensionality reduction is that it removes only few words from the data because the majority of words in language belong to open grammatical classes. Statistical criteria involve removing words with some undesired statistical characteristics, for example, very high and very low frequency. Removing very high and very low frequency words, thus, to some extent resembles linguistic filtering since very high frequency words tend to belong to closed grammatical classes. The statistical filtering not only removes closed class words, but also succeeds in removing words belonging to open grammatical classes. The result of reducing the dimensionality of the matrix is a low-dimensional space with denser information.

Finally, similarity between words (vectors) is compared using similarity or distance measures. The similarity measures give the mean distance between a target word and all its co-occurrence neighbours. Thus, semantic similarity indexes how near or similar the target word's neighbours to the target word in terms of similarity of their contextual usage in language. As mentioned in Section 2.2.2.1, an example of a similarity metric that is implemented in many models is the *cosine similarity* that measures the angle between two vectors. It was suggested that semantic similarity measures (both similarity and distance metrics) should be normalised for (or take out the effect of) vector length (i.e., the number of dimensions contained in each vector) because similarity measures result in making the words with many and large co-occurrence counts too similar to most other words while distance measures result in making words with many co-occurrence counts too far from other words (Widdows, 2004). This problem can be avoided by directly using cosine similarity since it normalises vectors for their respective length; thus, cosine similarity is a popular technique to compute normalised vector similarity (Sahlgren, 2008).

2.2.3 Examples of Distributional Semantic Models

The previous section discussed the general steps used in word-by-word distributional semantic models to build up semantic space that reflects the proximity between words and, therefore, the semantic similarity between the words in the space. This section extends the discussion on the building steps by providing a through description of one of the most influential word-by-word distributional models, *Hyperspace Analogue to Language* (HAL, Lund & Burgess, 1996). This description of HAL is followed by a through description of a recent HAL-based model, *High Dimensional Explorer* (HiDEx, Shaoul & Westbury, 2006, 2010a) that was empirically tested to provide empirical justifications for the optimal parameters that should be considered when building up distributional semantic models. A detailed description of the approach specified in HAL and HiDEx will be given in this section since it was on this basis that target words of the experiments reported in this thesis were selected. Hence, in order to have a clear appreciation of what the experimental manipulation will present, it is vital to fully understand how HAL and HiDEx computed semantic neighbourhood density (SND) metrics (the measures of i.e., how similar or close the neighbours to the target words in semantic space).

2.2.3.1 Hyperspace Analogue to Language (HAL)

Hyperspace Analogue to Language (HAL) is one of the most influential word-based distributional models. It was developed by Lund and Burgess (1996) to derive the meanings of words from lexical co-occurrences using 100000 words from a 160 million-word corpus that they derived from USENET newsgroups. In the original HAL, a 10-size window was used to collect co-occurrence frequencies. That is, words were considered to have co-occurred if they appeared within a window of 10 words ahead of the target word or 10 words behind the target word. Like the general steps of building co-occurrence matrices, HAL's moving window slides one word a time until all the words were processed. After recoding the first-order co-occurrences from the whole corpus, the frequencies of co-occurrences are weighted using a linear ramp function as a weighting scheme (see Section 2.2.2.2 for a description of this weighting scheme).

The weighted co-occurrence values were then stored in a raw matrix that included all possible combinations of words in all positions in the window. As a result, the co- occurrence matrix was very large (100000 target words * 100000 co-occurrence words * 20 positions [10 words ahead and 10 words behind the target word] = 200 billion vectors). In addition to the very high-dimensionality of the vectors, the matrix was also very sparse at this point with most cells in the matrix containing co-occurrence frequencies of zeros as discussed in Section 2.2.2.2. To reduce the high dimensionality of the vectors, the co-occurrence frequencies in the window were summed in a way that all the co-occurrence values appearing before the target word were summed in one cell, and all the co-occurrence values appearing after the target word were summed in another cell (see Figure 2.4). So, before summing the co-occurrence frequencies, the co-occurrence counts for each target word required 20 cells in the matrix (10 cells for the 10 context words ahead and 10 cells for the 10 context words behind the target word). After summing the co-occurrence counts, the co-occurrence counts for the target word only required two cells in the matrix (one cell ahead and one cell behind the target word).

	conduct -5	conduct - 4	conduct -3	Conduct -2	conduct -1	conduct +1	conduct +2	conduct +3	conduct +4	conduct +5
study	36	15	64	150	300	210	12	200	36	25

(36 + 15 + 64 + 150 + 300 = 565) (210 + 12 + 200 + 36 + 25 = 483)

	conduct (backwards)	conduct (forwards)
study	565	483

(A)

	a	<--->	concern	conduct	copy	<--->	zoo	a	<--->	concern	conduct	copy	<--->	zoo
study	14890	<--->	150	565	190	<--->	15	1200	<--->	301	483	58	<--->	12

(B)

	a	<--->	concern	conduct	copy	<--->	zoo	a	<--->	concern	conduct	copy	<--->	zoo
student	3487	<--->	565	677	357	<--->	2	1230	<--->	489	698	680	<--->	0
students	1250	<--->	678	999	780	<--->	3	1120	<--->	687	897	590	<--->	23
studentship	1984	<--->	23	87	0	<--->	0	1320	<--->	143	87	0	<--->	0
study	14890	<--->	150	565	190	<--->	15	1200	<--->	301	483	58	<--->	12
stuff	89065	<--->	123	9	230	<--->	2	1230	<--->	20	3	340	<--->	34

(C)

Figure 2.4 A (hypothetical) example of the way HAL word vectors are aggregated: by (A) summing weighted co-occurrences of the context word *'conduct'* that appeared ahead and behind the target word *'study'* into a single value each; (B-C) then the forward and the backward summed values are inserted into a global co-occurrence matrix; finally, the values are normalised after all vectors have been inserted in the global co-occurrence matrix (this normalisation step is not shown here) (based on Shaoul & Westbury, 2012).

The method used in HAL to remove the effect of orthographic frequency (see Section 2.2.2.2, for a description of this effect) was vector length normalisation by means of dividing each value in a vector by the vector length (i.e., the number of co-occurring words in that vector). However, vector length normalisation has been criticised as it can lead to systematic frequency bias (Durda & Buchanan, 2008, Shaoul & Westbury, 2006; Song, Bruza, & Cole, 2004). To explain, high frequency words tend to co-occur with many words in language, thus, they have large co-occurrences with many other words (Shaoul & Westbury, 2006), making their vector length larger compared to the vector length of low frequency words that co-occur with few words in language. However, normalising word vectors by dividing their elements by the length of the target word vector may not lead to eliminating the influence of orthographic frequency due to the following confound. Vector length is correlated with co-occurrence frequencies for each word, but it may not be correlated with the orthographic frequencies of the words. To explain, two words can have the same orthographic frequency; however, the vector length of these two words could be very different if the number of words with which they co-occur is different and if the position with which they co-occur differs (Shaoul & Westbury, 2006).

Before calculating semantic similarity, the sparse vectors in the original HAL were eliminated by keeping a number, N, of vectors with the highest row variances (i.e., getting rid of words that co-occur very often or very rarely with the target word). In this sense, if the rows with top 10000 most variant words were only used, then rows would contain 20000 elements/ cells (10000 words ahead +10000 words behind) instead of 200000 cells. The resultant matrix would be smaller and denser compared to the previous matrices. It was then that the distance between two vectors in the space was calculated using Euclidean distance metric, which was used to reflect how similar two words were in terms of the contexts of usage (co-occurring words). If two word vectors had similar values in the same dimensions, then the words represented by these vectors would be close to each other in the semantic space (see Figure 2.2).

To define a word's semantic neighbourhoods, the original HAL used a fixed N (usually ten) of the closest semantic neighbours to define *the number of semantic neighbours* a word had. It also used the average distance between the vector for the target word and those for the closest N neighbours in semantic space (see e.g., Buchanan *et al.*, 2001) to define *semantic neighbourhood density.* However, Shaoul and Westbury (2010a) pointed out that this way of conceptualising semantic neighbourhoods and semantic neighbourhood density is not ideal for two reasons. One reason is that averaging the distance of the N neighbours ignores the distribution of the distances of these neighbours from the target word. To explain, two words may have different distributions of neighbours around them in the space, however, averaging the distance as a density measure may result in concluding that the two words have identical density. A second reason is that using the 10 closest words to the target word was not empirically tested as a better measure of semantic neighbourhood density than say the 20 or 5 closest semantic neighbours. Indeed, Shaoul and Westbury (2010a) empirically tested different window sizes and found that other window sizes than the 10 word-sized window used in HAL were more optimal in predicting human performance on lexical decision tasks.

In addition, Shaoul and Westbury (2006) found that words' orthographic frequencies correlated with HAL's measures of semantic neighbourhood size (i.e., the number of words that are considered members of a word's semantic neighbourhoods) and semantic neighbourhood density (SND; i.e., the average distance between a target word and its semantic neighbours). As HAL did not satisfactorily eliminate the effect of orthographic frequencies, the original HAL's SND measure was confounded with orthographic frequency. Therefore, Shaoul and Westbury (2006, 2010a) modified HAL so that the measures of semantic neighbourhood size and SND were not sensitive to orthographic frequencies. To do this, Shaoul and Westbury changed the method of vector normalisation so that the influence of orthographic frequency is eliminated, and then they defined new measures of semantic neighbourhood size and SND. Shaoul and Westbury's modified version of HAL will be discussed in the next section.

To improve HAL's performance in predicting human data, many researchers developed different versions of HAL by varying some of its parameters (e.g., corpus size, window size, and semantic similarity measures, etc.) (Shaoul & Westbury, 2012). The researchers used the same design of the original HAL, but changed some of HAL's parameters so that they could compare the performance of the original HAL to that of their new versions of HAL in predicting human's performance in tasks such as lexical decision tasks and semantic categorisation tasks (Shaoul & Westbury, 2010a, 2012). Before giving examples of some models that modified some of HAL's original parameters, it is worth mentioning that the selection of a particular parameter was not based on theoretical motivation (Levy, Bullinaria, & Patel, 1998; Sahlgren, 2008). Instead, whether a particular parameter (e.g., a 10-word window) was psychologically plausible was only determined by experimentally testing whether this parameter could explain human performance on some tasks better than another parameter could (e.g., a 5 word window) (Levy *et al.*, 1998).

One of the modified HAL models is *COALS* (*Correlated Occurrence Analogue to Lexical Semantic*) model (Rohde, Gonnerman, & Plaut, 2007). Rohde and colleagues used correlation to normalise vectors produced by their model and to obtain similarity between vectors instead of using HAL's vector normalisation technique and similarity measures. They also removed closed class words from the corpus, and used *singular value decomposition (SVD)*[5] as a technique of dimensionality reduction of the co-occurrence matrix. With these changes to the original HAL, Rohde and colleagues found that COALS performed really well on word semantic similarity tasks. When Bullinaria and Levy (2007) changed HAL's Euclidean distance metric to *PMI (pointwise mutual information)*[6] to measure the distance between word vectors, they found that the accuracy of HAL's performance on semantic tasks improved. In addition, Durda and Buchanan (2008) used the design of HAL to develop a model called *WINDSORS* (Windsor Improved Norms of Distance and Similarity of Representations of Semantics). WINDSORS eliminated any real correlation with orthographic frequency through the use of many statistical and mathematical methods, and Durda and Buchanan found that WINDSORS was able to model semantic priming tasks and word similarity tasks. There are other models that were developed based on HAL, but were far more complex than HAL (e.g., *BEAGLE* by Jones & Mewhort, 2007). These were developed to account for

[5] SVD was used to calculate the approximation of a lower-dimensional co-occurrence matrix to the original high-dimensional matrix

[6] PMI was used to calculate the ratio between the probability of two words co-occurring given their joint distribution versus the probability of their co-occurrence given their individual distributions and assuming independence.

sentence completion and semantic categorisation. Below is a thorough description of one of the recent implementation of HAL that enabled researchers to use different parameters to explore the effects of changing these parameters on the model's performance in predicting human data. This new implementation of HAL was empirically tested and found to account for the variation in humans' response latencies obtained in single word recognition studies (Shaoul & Westbury, 2010a).

2.2.3.2 High Dimensional Explorer (HiDEx)

Recently, Shaoul and Westbury (2006, 2010a) developed a model called *High Dimensional Explorer* (HiDEx) by changing some parameters of the original HAL to better explain the variance in response data obtained from lexical decision and semantic decision tasks. They also used a bigger corpus of one billion words from texts found in USENET (text from 2005 to 2007) instead of 160 million words used by HAL. Shaoul and Westbury chose this bigger corpus because many words in their lexicon had one or very few occurrences in a corpus of 160 million words; thereby they opted for a larger corpus as a way of obtaining many occurrences for all words. After the linguistic processing of the corpus as per Section 2.2.2.2, the co-occurrence matrix was built with HiDEx using a sliding window described in earlier sections in this chapter. Shaoul and Westbury (2010a) found that a window size of 10 words behind and 0 or 5 words ahead of a target word was the best in capturing variance in lexical and semantic decision responses to that target word.

After recording the co-occurrence frequencies, the co-occurrence frequencies in the window were weighted using a weighting scheme. The findings of Shaoul and Westbury (2010a) indicated the most optimal weighting scheme was the one that gave less weight to the closer words that often tended to be function words, and therefore, high frequency words. One example of such weighting scheme is the inverse linear ramp (see Section 2.2.2.2, for a description of this weighting scheme). Then, the weighted co-occurrences were summed in each window as per HAL's technique of summing co-occurrences described in Section 2.2.3.1.

To eliminate the effect of orthographic frequencies, Shaoul and Westbury (2010a) normalised each word vector by dividing all the elements/ co-occurrence values in each vector by the orthographic frequency of the word represented by the vector instead by the vector's length, as Buchanan *et al.* (2001) recommended. In this way, the co-occurrence values for high frequency words shrank while they were amplified for low frequency words. To illustrate, consider that there are two target words (*T1* and *T2*) that both co-occur with a word (*W1*). Say *T1* occurred 10 times in a selected corpus and co-occurred once with *W1* in this corpus. And say *T2* occurred 100 times in the same selected corpus and co-occurred 10 times with *W1*. If these co-occurrence values for each target word vector were divided by its respective target word's orthographic frequency (1/10 in the vector of *T1* and 10/100 in the vector for *T2*), then this division will give the same value for the element *W1* in each vector of *T1* and *T2*. Thus, the influence of orthographic frequency would be eliminated in HiDEx. The consequent weighted co-occurrence matrix was very sparse at this stage because most words rarely co-occurred with each other.

HiDEx dealt with the sparseness issue by retaining only a certain number of vectors (usually 14000) for words with the highest orthographic frequency (instead of the greatest variance used in HAL). Finally, the word vectors were used to compute contextual similarity. To do this, HiDEx calculated the mean distance between the vectors of a word and the words comprising its neighbourhood. Shaoul and Westbury (2010a) specified that HiDEx use a multiple of the standard deviation of the distances between a word's vector and vectors for each of its semantic neighbours to compute a threshold. This threshold was then used as a cut off, and the number of semantic neighbours that fell within this threshold determines the word's semantic neighbourhood size (NCount). This procedure is explained next in slightly more detail. A set of word pairs that have co-occurred at least once in the corpus was created. Then, 5%- 10% of the total number of pairs (constituting billions of word pairs) was randomly selected. Then, the distances between each selected word pair (usually represented as the cosines of the angles between the vectors representing the words) were calculated to find the standard deviation of all distances, and this in turn was used to define a threshold of neighbourhood membership (i.e., which words were, and which words were not, considered semantic neighbours of the target word in question). The neighbourhood membership threshold was set at 1.5 standard deviations below the mean distance (which was about 6.7% of the average distance between any two words). Since most words have weak or no relationships, this cut-off point would ensure that a neighbour was counted as part of a word's neighbourhood only if it was one of the closest 6.7% of millions of the randomly selected pairs. Due to the thresholded nature of the neighbourhood, some proportion of a word's co-occurrence neighbours would not be categorised as semantic neighbours, and in this way, the number of semantic neighbours a particular word had will vary, with some words having more semantic neighbours than others, and some words even having none.

After obtaining a measure of a word's semantic neighbourhood in HiDEx, a measure can then be obtained called the *Average Radius of Co-occurrence* (ARC) which is the average cosine or distance between a target word and all its semantic neighbours within the threshold (see Figure 2.5). Since ARC is an average distance measure, it reflects how close or distant the semantic neighbours are to a target word in semantic space, and to this extent, ARC indexes semantic neighbourhood density (SND). A word that has more close semantic neighbours is more similar to its neighbours in terms of contextual usage (i.e., they appear quite frequently in similar contexts). A word that has more distant neighbours indicates that this word is less similar to its neighbours. In this way, ARC captures and represents the average similarity of a word to its co-occurrence neighbours that fall inside its neighbourhood threshold. The resulting average of cosine similarity (i.e., ARC values) ranges from 0 to 1, where 0 indicates no similarity and 1 indicates maximal similarity, with no negative values since frequencies of co-occurrence cannot be negative. Words that have semantic neighbours that are more similar to them have higher ARC values. When a word has no semantic neighbours (as a consequence of the thresholded neighbourhood), the word is assigned an ARC value that reflects the distance between the word and its closest semantic neighbour (the first co-occurrence neighbour outside the threshold). In sum, the ARC values are an index of a word's SND, and the influence of SND in word identification during normal reading is the focus in this thesis.

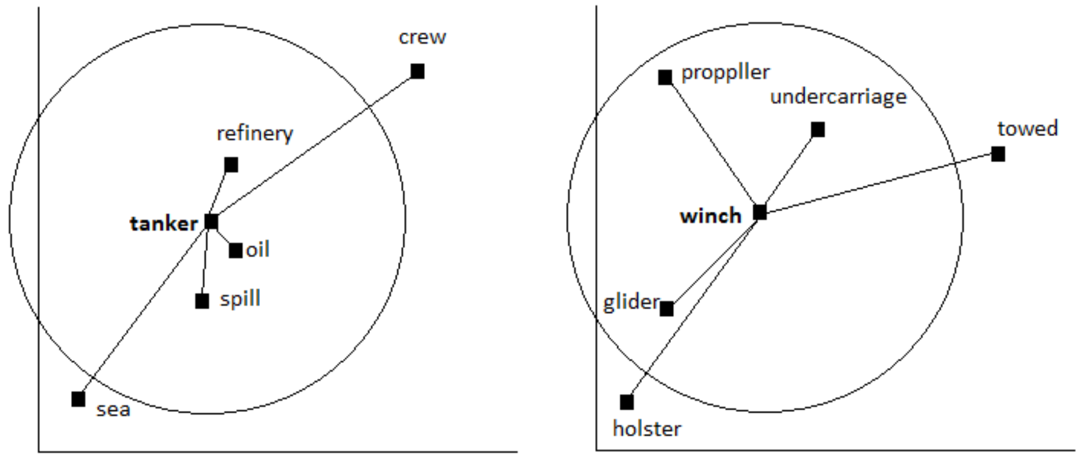

Figure 2.5 A two-dimensional visualisation of the neighbourhood membership threshold. The words *'tanker'* and *'winch'* in this example have three semantic neighbours (based on Shaoul & Westbury, 2010a). The semantic neighbours are close to *'tanker'*, whereas the semantic neighbours are distant from *'winch'*. Thus, *'tanker'* has a higher ARC value than *'winch'*.

2.2.4 Summary

To summarise, the distributional-based view of semantics assumes that a word's meaning can be captured from its distribution in large samples of text by analysing the context (e.g., neighbouring words) in which the word occurs. Words that share a set of words with which they commonly co-occur are also assumed to have similar meaning. Many distributional semantic models were developed to represent semantic similarity between words in terms of spatial proximity of the words in a spatial representation of meaning (i.e., sematic space). Semantic space has a large number of dimensions with points (vectors) that represent the location of the words in the space. The position of a word vector in relation to the positions of other word vectors in the space indicates the extent to which some aspects of meaning are shared among the words. Particularly, it was discussed that the distance between the vectors reflects how similar their meanings are. Words that are more related in their meanings tend to cluster closer together in semantic space, whereas words that are semantically less related are more distant from each other in this space. It was also discussed that some mathematical and statistical techniques are implemented in the distributional semantic models such as HAL and HiDEx to arrive at semantic representations and the semantic similarity between words. The next section will be devoted to discussing some empirical findings of the effect of co-occurrence-based semantic neighbourhoods and semantic neighbourhood density (SND) in lexical processing.

2.3 Empirical Studies on the Effects of Semantic Neighbourhood Density (SND)

Several studies have been conducted to examine the language-based SND effects in lexical processing; these studies varied in terms of the specific SND measures used to define semantic neighbourhoods and the type of behavioural tasks used in their experiments. To reiterate, the current thesis will use the term dense semantic neighbourhood words to refer to words whose semantic neighbours are close and, thus, semantically similar to them in semantic space, and will use the term sparse semantic neighbourhood words to refer to words whose semantic neighbours are distant and, thus, semantically different from them in semantic space. Generally, the findings from many studies produced convergent evidence that denser semantic neighbourhood words were responded to faster than sparser semantic neighbourhood words in tasks that rely on the familiarity of the presented words to make responses such as lexical decision tasks[7] (e.g., Buchanan, Westbury, & Burgess, 2001; Shaoul & Westbury, 2010a). On the contrary, the findings about the SND effects in tasks that require excessive[8] processing of the meaning of the presented words are mixed, with some findings indicating a facilitaotry effect (Siakaluk, Buchanan, & Westbury, 2003) while other findings showing an inhibitory effect (Shaoul & Westbury, 2010a), and still other findings demonstrating a null effect of SND

[7] In a lexical decision task, human participants are asked to make yes/no response as to whether some presented visual stimuli constitute words or nonwords, as quickly and as accurately as possible.

[8] Excessive meaning processing in the sense that participants are asked to do some deep analysis of the meaning of the presented words, an analysis that does not normally take place in silent reading tasks in which the meanings of words are automatically extracted without asking participants to make a deep analysis of the meanings of the presented words.

(Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012; Yap, Tan, Pexman, & Hargreaves, 2011). An example of such tasks is the semantic categorisation task whereby participants are asked to make responses as to whether some presented visual words are concrete or non-concrete (living or no-living; or animal or non-animal), as quickly and as accurately as they can.

In this section, some of these studies will be reviewed; the review will focus on the studies that used a co-occurrence-based definition of semantic neighbourhoods, with limited reference to the findings of the associational-based SND effects [9]. First, this section will discuss the findings of the earlier studies conducted on the co-occurrence-based SND effects, and then the review will turn into describing the findings of the most recent studies that explored the SND effects. This will be followed by a discussion of Shaoul and Westbury's (2010a) study in which they used the SND measures derived from their model (HiDEx) (see Section 2.2.3.2). In so doing, the consistencies and inconsistencies between the findings of the studies will be highlighted, mentioning the differences in the methodologies of the reviewed studies. Finally, the section will discuss how the SND effects were interpreted in the context of visual word recognition models.

One of the earliest and most influential studies that explored the SND effects was conducted by Buchanan, Westbury, and Burgess (2001). Buchanan and colleagues conducted a series of experiments in which they examined the effects of semantic neighbourhood size (i.e., the number of semantic neighbours) in lexical processing using lexical decision tasks and naming tasks. One of the measures of semantic neighbourhood size they used was HAL's semantic distance, which they defined as the mean distance between the target word and its 10 closest semantic neighbours in semantic space. Using hierarchal regression analyses, Buchanan *et al*. tested whether HAL's semantic distance could predict the speed with which words were recognised in lexical decision tasks and naming tasks. They removed the role of other lexical variables by entering them first before HAL's semantic distance in the regression analyses in the following order: log frequency, number of

[9] Recall that associational-based semantics is part of the language-based semantics. It will be worthwhile to give a brief overview of the findings of how the associational-based SND affects lexical processing and compare these findings with those of the co-occurrence-based SND effects.

orthographic neighbours, word length, number of semantic associates, and semantic distance. They found that HAL's semantic distance predicted lexical decision latencies and, to some extent, naming latencies. The findings also showed that there was a positive partial correlation (and a semi-partial correlation) between lexical decision latencies and semantic distance, reflecting that as semantic distance decreased the response latencies decreased as well. Thus, their finding in the first experiment clearly showed that words with denser semantic neighbourhoods (decreased semantic distance between words and their respective closest semantic neighbours) resulted in quicker lexical decision latencies compared to words with sparser semantic neighbourhoods.

To assess whether the facilitatory SND effects were not due to a confounding effect of a traditional semantic variable (imageability ratings), Buchanan *et al.* (2001) included imageability in a hierarchal regression analysis. Their findings indicated that semantic distance accounted for a unique variance in lexical decision latencies even after partialling out the contribution of imageability. Then, Buchanan and colleagues followed their regression analyses with some factorial experiments to further examine the effects of semantic distance. They observed that words with large semantic neighbourhood size (i.e., words with low semantic distance) were responded to faster than words with small semantic neighbourhood size (i.e., words with high semantic distance), even after partialling out the effect imageability from their analyses. They also found that the effect of semantic distance was larger for low frequency words as opposed to high frequency words. That is, lexical decision latencies for low frequency words appeared to be influenced by semantic distance more than those for high frequency words. One limitation of Buchanan *et al.*'s (2001) study as they themselves noted is that they used a cutoff point in terms of a fixed number of the closest semantic neighbours (10 semantic neighbours in their case) to define the semantic neighbourhoods of the words they used in their experiments. Instead, they recommended defining semantic neighbourhoods by using a cutoff point in terms of distances and then counting the number of semantic neighbours falling within the specified distance, which was how Shaoul and Westbury (2006, 2010a) defined semantic neighbourhoods in their HiDEx model.

To test whether the findings of Buchanan *et al.* (2001) could be extended to other visual word recognition tasks, Siakaluk, Buchanan, and Westbury (2003) used two types of semantic categorisation tasks that were assumed to be more sensitive to semantic effects and that were thought to require accessing word meaning before making responses to the presented stimuli (Forster & Shen, 1996). The two types of semantic categorisation tasks were a *yes/no- task* in which participants were asked to respond to both experimental words (nonanimals) and non-experimental words (animals) and a *go/no-go task* in which participants were asked to respond to only experimental items. They also used Buchanan *et a*l.'s (2001) definition of semantic distance. Using a one-way analysis of variance (ANOVA) to analyse the data of their first experiment (a yes/no task), the results showed that the main effect of semantic distance was just significant in the subject analysis and was not significant in the item analysis. A post hoc analysis showed that the responses to low semantic distance (i.e., denser semantic neighbourhood) words were 15ms faster than responses to matched high semantic distance (i.e., sparser semantic neighbourhood) words. To determine that the lack of effect of semantic distance was not due to another semantic variable (subjective frequency, *a.k.a.* familiarity), they entered subjective frequency into a regression analysis, and found that semantic distance accounted for only a modest amount of variance (3%) above and beyond subjective frequency. Thus, the researchers concluded that this lack of semantic distance effect in their first semantic categorisation task was not due to a possible confounding variable of subjective frequency.

Instead, Siakaluk and colleagues hypothesised that this lack of semantic distance effect in their first semantic categorisation task might be due to the differences between the responses made in lexical decision tasks and responses made in their yes/no semantic categorisation tasks. To explain, *'yes'* responses were expected to be made to the experimental stimuli (words) in the case of lexical decision tasks, whereas *'no'* responses were expected to be made to the experimental stimuli (animal-ness) in the case of yes/no semantic categorisation tasks. Thereby, Siakaluk and colleagues conducted a second experiment in which they employed a go/no-go semantic categorisation task, which was assumed to require participants to make *yes-like* responses to only the experimental stimuli similar to the responses made in lexical decision tasks and, hence, increasing the chance of observing the effect of semantic distance. The findings of their second experiment indicated that the effect of semantic distance was significant; low semantic distance words (i.e., denser semantic neighbourhood words) were categorised 41ms faster than words with high semantic distance (i.e., sparser semantic neighbourhood words).

These two studies of Buchanan *et al*. (2001) and Siakaluk *et al*. (2003) so far suggest that the effects of SND do appear in tasks that require making responses based on the familiarity of the orthography of words (and accessing words' meanings to some extent) rather than in tasks that require deep semantic processing that does not necessarily reflect processes taking place in lexical processing during normal reading. This observation is supported by the findings of other more recent studies that replicated Buchanan *et al.*'s (2001) facilitatory effect of SND on lexical decision latencies using different measures of SND (Pexman, Hargreaves, Siakaluk, Bonder, & Pope, 2008; Shaoul & Westbury, 2010a; Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012; Yap, Tan, Pexman, & Hargreaves, 2011). However, the facilitatory SND effect on semantic decision latencies found by Siakaluk *et al*. (2003) could not be replicated later. Indeed, the findings of the SND effects on semantic tasks are inconsistent, with some researchers observing an inhibitory effect (Shaoul & Westbury, 2010a) and others demonstrating a non-significant effect (Pexman *et al*., 2008; Yap *et al*., 2011; Yap *et al*., 2012).

These above-cited studies used different measures of SND. For example, Pexman *et al*., (2008) found that that number of semantic neighbours derived from co-occurrence information in a high dimensional semantic space (Durda, Buchanan, & Caron, 2006) significantly predicted only the lexical decision latencies, but not the semantic categorisation latencies. This finding was later replicated by Yap *et al.* (2011) and Yap *et al*. (2012) using different measures of co-occurrence-based semantic neighbourhoods. Yap *et al*. (2011) used the mean cosine similarity between a target word and its closest 5000 neighbours in a high dimensional space as a measure of SND, and Yap *et al.* (2012) used the ARC metric described in Section 2.2.3.2 in this chapter. All these studies found that the effects of semantic neighbourhoods were only present in lexical decision tasks, with denser semantic neighbourhood words (i.e., words with semantic neighbours that are more similar to them) were responded to faster than words with sparser semantic neighbourhoods.

Another piece of recent empirical evidence that showed the consistency of the findings of the facilitatory SND effects on lexical decision latencies and the inconsistencies of the SND effects on semantic decision latencies is the study carried out by Shaoul and Westbury (2010a). In particular, Shaoul and Westbury tested whether HiDEx's indices of SND could explain differences in lexical and semantic decision data better than the original HAL's parameters (window size and weighting scheme). When using inverse ramp as a weighting function and a window size of 10 words behind and 0 or 5 words ahead, they reported that SND predicted response latencies better than using the original HAL's parameters of a linear ramp weighting scheme and a 10-word window. Specifically, they demonstrated that words with higher SND produced shorter lexical decision responses, consistent with the findings of the previous studies (Buchanan *et al.*, 2001; Pexman *et al.*, 2008; Yap *et al.*, 2011; Yap *et al.*, 2012). Shaoul and Westbury also investigated the SND effect in other tasks that required more extensive semantic processing. Specifically, Shaoul and Westbury used two semantic decision tasks in which they asked participants to make explicit semantic judgments about whether two words in a pair were or were not related (yes/no semantic decision task), and to make responses only to word pairs that were semantically related (a go/no-go task). They observed an inhibitory effect of increased SND (i.e., increased SND resulted in longer decision latencies in both tasks), contrary to the facilitatory SND findings observed by Siakaluk *et al.*, (2003) in their semantic tasks. Recall that Siakaluk *et al.*'s semantic tasks are slightly different from those of Shaoul and Westbury's.

The inconsistency between Shaoul and Westbury's findings and those of Siakaluk *et al.*'s results can probably be attributed to differences in the types of semantic decision tasks used in the two studies. Shaoul and Westbury's decision tasks involved judgments as to the semantic relatedness of sequentially presented word pairs while Siakaluk *et al.*'s decision tasks involved judgments about single words. Although Shaoul and Westbury's task was not, strictly speaking, a semantic priming lexical decision task, the format of presenting a target word to which a response was required immediately after a preceding word is certainly a close approximation to a priming paradigm. It is at least possible that the response latencies in their experiments may have reflected the influence of the preceding word on processing of the subsequent target word (see Moss & Tyler, 1995 for a review of semantic priming effects). In addition, as Shaoul and Westbury themselves noted, their decision latencies were much longer than those found in the other studies employing semantic categorisation tasks in which participants made responses to single words (e.g., Binder, Westbury, McKiernan, Possing, & Medler, 2005; Siakaluk *et al.,* 2003). It is therefore possible that decision times in this study reflected post-lexical processes associated with decision formation. It seems likely that methodological differences caused the differing patterns of effects.

Broadly, on the basis of the complete body of research discussed above, it appears that consistent and pronounced facilitatory SND effects are obtained in tasks that tap into simple word identification processes rather than those that require extensive processing of the meaning of words. In all of these studies, the researchers interpreted the facilitatory effects of SND within an interactive model of word recognition in which all the (orthographic, phonological and semantic) levels are connected by bi-directional activation links (i.e., an Interactive- Activation based framework as per McClelland & Rumelhart, 1981). Such models assume feedforward and feedback activation between their distinct units that are dedicated to processing orthography, phonology, and semantic information. These models also assume interactivity between the units of processing in the sense that activation from one unit (or more) can affect the processing of other units. These models explain the SND effects as follows: a target word with a denser semantic neighbourhood (i.e., the average similarity between the word and its neighbours is high) will receive more activation from its close co-occurrence neighbours at the semantic level, and the increased semantic activation is fed back from the semantic level to the orthographic (word) level. Consequently, the orthographic representation of the target word will be facilitated. This in turn results in speeded lexical decision responses (e.g., Buchanan *et al.*, 2001). In the present thesis, the findings that will be obtained in the experiments will be considered within this interactive context.

At this point, it is fair to mention Mirman and Magnuson's (2008) study that was often cited as providing a contrast to the facilitatory effect of the increased SND in lexical decision tasks found in the above-reviewed studies. However, the findings of Mirman and Magnuson's study should be read with a caveat as the SND measures used in their study was defined in terms of a feature-based measure rather than a co-occurrence-based measure. Specifically, in their second experiment, Mirman and Magnuson (2008) showed that a facilitatory effect could arise as a consequence of distant (i.e., less similar) semantic neighbours rather than close (i.e., more similar) semantic neighbours. In their study, near semantic neighbours slowed semantic and lexical decision times whilst distant semantic neighbours speeded decision times. Training an attractor dynamic network, Mirman and Magnuson studied the effect of near and distant neighbours by examining the correlation between the number of near and distant neighbours with errors in settling into a correct activity pattern of semantic units for a concept. Their findings showed a strong positive correlation with number of near neighbours (i.e., high number of near neighbours was linked to making more settling errors), indicating an inhibitory effect of near neighbours. The findings also revealed that there was no reliable correlation with the number of distant neighbours, except for a dip to the negative side (i.e., more distant neighbours was associated with fewer errors), indicating a facilitatory effect of distant neighbours. Interpreting their findings in terms of attractor dynamics, the researchers suggested that distant neighbours are far away from the target word, creating a gravitational gradient for faster settling into attractor basins, whilst near neighbours slowed the settling process because their basins of attraction are closer to the target word's basin of attraction.

However, it should be noted that Mirman and Magnuson, in their second experiment, defined near vs. distant neighbours in terms of the cosine or distance between the target's semantic features (e.g., taste, colour, function, etc.) and the semantic features of other words in the corpus they used, rather than in terms of the distance between the co-occurrence neighbours and their respective target words. As mentioned earlier, the feature-based view of semantics defines semantic similarity (and presumably semantic neighbours) as a function of shared semantic features, while the distributional-based view defines semantic similarity in terms of co-occurrence within similar contexts. For instance, *movie* and *play* are considered semantic neighbours under the distributional view because they tend to occur within similar semantic contexts, while they are not considered semantic neighbours under the feature-based view because they do not share semantic features. Thus, the feature-based semantic similarity (i.e., how much a target word shares semantic features with other words) is based on a different theoretical account than the distributional hypothesis, and is therefore not synonymous with the semantic similarity of co- occurrence neighbours. As such, feature- based SND does not speak directly to co-occurrence- based SND, and consequently, the effect of semantic similarity of neighbours, defined in terms of shared semantic features, found in Mirman and Magnuson's (2008) study may not apply to the effect of semantic similarity of neighbours as defined in terms of shared co- occurrence neighbours. Since the focus of this thesis is on co-occurrence-based SND, Mirman and Magnuson's findings SND effects may not be comparable to those findings of the studies that used co-occurrence-based semantic representations to define semantic neighbourhoods.

While the focus of this thesis is on the co-occurrence-based SND, a brief review of the findings pertaining to the associational-based SND is necessary in order to compare both language-based definitions of SND and, hence, provide an overview of how the language-based SND in general influences lexical processing. Buchanan *et al.* (2001) investigated the effect of semantic neighbourhood size as defined by number of associates (Nelson, McEvoy, & Schreiber, 1998) along with HAL's semantic distance, and found that the effect of the associational-based semantic neighbourhood size was facilitatory, however, was weaker than the effect of semantic distance they derived from HAL. This finding of Buchanan *et al.* was replicated later by Yap *et al.* (2011) who found that the number of associates did not predict response latencies in lexical decision and naming tasks. Other researchers found a robust effect of semantic neighbourhood size as defined by the number associates on lexical decision latencies (e.g., Locker, Simpson, & Yates, 2003, Yates, Locker, & Simpson, 2003; Duñabeitia, Avilés, & Carreiras, 2008). Particularly, these studies found that words with larger semantic neighbourhoods (large number of associates) were responded to faster than words with smaller semantic neighbourhoods in lexical decision tasks (Locker *et al.*, 2003, Yates *et al.*, 2003; Duñabeitia, *et al.*, 2008), naming, progressive demasking and sentence reading in Spanish (Duñabeitia *et al.*, 2008). The study of Duñabeitia, *et al.*, (2008) clearly indicated that Spanish words with a high number of associates were read for a shorter time than matched words with low number of associates as evident in gaze duration (21ms shorter) and total reading time (23ms shorter). Thus, their study suggests that semantic neighbourhood effects may appear in early measures of lexical processing (such as gaze duration, see Section 1.1 in Chapter 1 for a description of this measure), indicating that semantic neighbourhoods, as defined by at least the associational-based models, influence lexical processing during normal reading in Spanish.

In sum, the visual word recognition studies produced convergent evidence that word with denser semantic neighbourhoods (i.e., words with more similar semantic neighbours) are recognised faster than words with sparser semantic neighbourhoods (i.e., words with less similar semantic neighbours) in tasks that depend on the familiarity of words to make responses (e.g., lexical decision tasks). The SND effects are less clear and have proved inconsistent across studies that used tasks that require participants to do excessive processing of the meaning of words before making responses (e.g., semantic categorisation tasks) due to the nature of such tasks as discussed in this section. The facilitatory effect was explained within interactive models that assume strengthened feedback from the semantic level to the orthographic level.

2.4 Purpose of the Present Thesis

All of the previous studies mentioned in Section 2.3 used isolated word recognition tasks. Isolated word recognition tasks are known to exert some experimental demands that are not necessarily part of normal reading (Kuperman, Drieghe, Keuleers, & Brysbaert, 2013) as discussed in Chapter 1. How the SND effects would affect normal reading has not yet been explored, and for this reason, the experiments reported in this thesis undertook the current investigation. A relevant study that investigated the effect of semantic neighbourhoods, defined by the number of associates a word has, on reading in Spanish was reported by Duñabeitia, Avilés, and Carreiras (2008) as discussed in the previous section. Duñabeitia *et al.* found that the associational-based SND effects appeared in early records of eye movements during Spanish sentence reading such as gaze duration, suggesting the language-based SND effects, as operationalised at least in the associational-based terms, can influence lexical processing during normal reading.

This thesis investigates the effect of co-occurrence-based SND, as defined by the mean distance between a word and its semantic neighbours falling within a specified threshold (Average Radius of Co-occurrence; ARC, Shaoul & Westbury, 2010a) in normal reading. The main question of this thesis is how a word's SND influences its lexical identification during normal reading in English. Particularly, this thesis addressed four related questions: (1) whether the average distance between a target word and its closest semantic neighbours falling within a specified threshold (i.e., SND) predicted fixation times on the target word, (2) whether words that had more similar semantic neighbours (i.e., high SND words) were fixated for a shorter time than words with less similar semantic neighbours (i.e., low SND words), (3) whether target word frequency interacted with target words' SND, and (4) whether target words' SND interacted with their orthographic neighbourhood size.

To answer these questions, the present thesis directly examines Shaoul and Westbury's (2006, 2010a) method for defining semantic space, using linear-mixed effect (LME) models, which will be introduced in Chapter 3, to analyse the fixation times on the selected target words with high and low ARC values. The predictions of the present study findings are derived from the findings of Duñabeitia *et al.* (2008) and the findings of lexical decision tasks, which are thought to resemble normal reading tasks in that no excessive processing of the meaning of words is required (Buchanan *et al.*, 2001; Pexman *et al.*, 2008; Shaoul & Westbury, 2010a; Yap *et al.*, 2011, Yap *et al.*, 2012). Based on these findings, it was predicted the SND effects in lexical processing occurring during normal reading would be facilitatory for early reading time measures (e.g., first fixation, single fixation and gaze duration). Words whose semantic neighbours that are more similar to them (i.e., words with high SND (high ARC values)) are expected to be read faster than words whose semantic neighbours are less similar to them (i.e., low SND words (low ARC values)).

Providing answers to the questions of this thesis will clarify whether and how word meaning can constrain unique word identification during lexical processing occurring in normal reading. If the SND characteristics are found to influence word identification, then such findings will clearly demonstrate that the SND effect is not the artefact of laboratory single word recognition tasks, such as lexical decision tasks, but actually influences word identification during reading. Additionally, if the SND metric of the Average Radius of Co-occurrence (ARC, Shaoul & Westbury, 2010a) is found to predict the fixation times spent on reading target words in normal reading tasks that are assumed to activate semantic representations, such findings then will support the claim that ARC can capture informative aspects of words' semantic representations. This, in turn, will give support to Shaoul and Westbury's (2010a) conceptualisation of semantic representations.

2.5 Summary and Thesis Structure

A word's meaning can be realised in relation to the meanings of other words that appear in similar contexts (i.e., by its distribution in text). Some distributional semantic models have been developed to capture the meaning of words in terms of their similarity to other words in semantic space. The semantic space is built by means of using co-occurrence information found in large text corpora. The words are represented as points or vectors in semantic space. A vector contains dimensions constituting statistical information about the number of times a given word co-occurs with other words in the corpus, and these numerical values specify the location of the given word in semantic space. If two word vectors have similar values (i.e., co-occurrence frequencies) in the same dimensions, then the words represented by these vectors will tend to be close to each other in semantic space. Thus, they are considered semantically similar or semantic neighbours in this space. If word vectors have very different values in the same dimensions, on the other hand, then the words represented by these vectors will tend to be distant from each other in semantic space and, hence, will be semantically less related to each other (i.e., distant semantic neighbours). The influence of how close or distant sematic neighbours a word has (i.e., SND) in lexical processing has been examined in studies that used isolated visual words recognition tasks. These studies reported that words with denser semantic neighbourhoods (i.e., words with more close or similar semantic neighbours) were responded to faster in lexical decision tasks than words with sparser semantic neighbourhoods (i.e., words with distant semantic neighbours). However, these studies used tasks that do not necessarily reflect lexical processing that occurs during normal reading.

To date, there have been no studies that have examined the SND effects with the respect to normal reading and eye movements; thereby, the SND effects are not yet understood in normal reading using eye movements as a methodology. Thus, the experiments reported in this thesis will be the first to explore the SND effects in normal reading. SND in this thesis was defined in terms of the mean distance between a target word and its close semantic neighbours that fall within a specified threshold (Average Radius of Co-occurrence; ARC, Shaoul & Westbury, 2006; 2010a). The SND effects in this thesis were investigated in four experiments that will be described in the next chapters. If these experiments could establish the SND effects in the early measures of eye movements during reading, such findings would provide a basis for understanding how the semantic characteristics of words can constrain unique word identification in normal reading. In this way, the findings of this thesis will contribute to the literature of eye movements during reading in terms of informing our understanding about how word meaning influences lexical identification. In addition, such findings will also provide support to Shaoul and Westbury's (2010a) conceptualisation of semantic representations as capturing informative aspects of semantic memory.

The rest of the thesis will be organised in the following structure. Chapter 3, 4, 5 and 6 will describe the experiments that were carried out to answer the research questions of this thesis alongside a discussion of the findings. Chapter 3 will describe the first experiment constituting a preliminary investigation of the main SND effects in normal reading while controlling for other variables that are well known to influence lexical processing. Target words in this experiment were placed in single line sentences, and linear mixed effect (LME) models were used for data analyses. The findings demonstrated that target words' SND influenced the lexical processing of the target word and the subsequent words. The second experiment described in Chapter 4 passively explored the interactive effects of SND and other variables without strictly controlling for the variables that are known to influence lexical identification. In this experiment, target words were embedded in three passages of text, and LME models were also used to analyse the data. With the use of LME models, no strict control of the extraneous variables was required, thus, maximising the number of the target words used in this experiment and, as consequence, maximising the statistical power. The findings of this experiment indicated that target words' SND interacted with target word frequency.

The third experiment described in Chapter 5 was a follow-up experiment based on the findings of the second experiment; Experiment 3 directly examined the joint effect of SND and word frequency using single-line sentences. LME models were used again for data analyses. Based on the findings, it was concluded that a word's semantic representation could influence lexical processing prior to the word's full identification. The fourth experiment described in Chapter 6 was conducted to provide further evidence to the conclusion made in Chapter 5 that a word's SND characteristics can constrain unique word identification in normal reading. For this experiment, another word-level variable, namely, the number of orthographic neighbours, was orthogonally manipulated along with the SND metric. In this experiment, a sentence reading task was also used with LME models for data analyses. Chapter 7 will summarise the SND effects that were found in normal reading in this thesis, and discuss the theoretical implications of the findings and directions for further studies.

Chapter 3: Basic Effects of Semantic Neighbourhood Density in Normal Reading

Chapter 1 reviewed abundant eye movement data showing that the fixation times on a word reflect processes associated with accessing the word's orthographic and phonological representations, as well as accessing the word's meaning and assessing its congruency with the prior context. However, most eye movement studies suggesting that word meaning influences lexical identification used contextual information (e.g., lexical ambiguity studies: Duffy *et al.*, 1988; Sereno *et al.*, 2006; semantic plausibility: Filik, 2008; Rayner *et al.*, 2004; Staub *et al.*, 2007). Therefore, these studies provide only a limited insight into how the semantic representation of a fixated word is accessed during its lexical identification. Whether the semantic characteristics of a word influence its lexical identification has received little attention in the literature on eye movements during reading. The few eye movement studies concerning the influence of a word's semantic characteristics in lexical processing in reading indicate that the rich semantic properties of foveal words (e.g., high number of associates: Duñabeitia *et al.*, 2008; high number of semantic features: Cook *et al.*, 2013; high contextual diversity: Plummer *et al.*, 2014) facilitate lexical processing in normal reading as evident in decreased fixation times on words with rich semantic representations.

Another semantic influence that remains completely unexplored with respect to eye movements and reading is semantic neighbourhood density (SND) effects, defined as the effects of the average distance between a word and all its semantic neighbours in semantic space (Lund & Burgess, 1996). All studies that investigated the SND effects in lexical processing to date have been carried out using single word recognition tasks that are associated with particular task demands that are not necessarily present during normal reading, as discussed in Chapter 2. To reiterate, these isolated word studies have generally indicated that words with denser semantic neighbourhoods are processed faster than words with sparser semantic neighbourhoods in lexical decision tasks (e.g., Buchanan *et al.,* 2001; Siakaluk *et al.,* 2003). However, it is not clear that the SND effects observed in behavioural tasks are necessarily artefacts of isolated word recognition tasks, or that they will necessarily generalise to eye movement behaviour associated with normal reading. This thesis raises the question of the extent to which the previous SND findings may be related to laboratory behavioural tasks by investigating SND effects, defined according to the Average Radius of Co-occurrence (ARC, Shaoul & Westbury, 2010a), in normal reading tasks. Evidence from eye movement experiments will provide more ecologically valid confirmation as to whether Shaoul and Westbury's (2010a) conceptualisation of semantic representations can capture informative aspects of semantic memory. In addition, the thesis will contribute to the literature on eye movements during reading by examining whether and how a word's semantic characteristics can influence its lexical identification during normal reading.

3.1 Experiment 1

In Experiment 1, a standard experimental approach was adopted so that target words with either high or low SND were embedded within the same sentential contexts. Target words and sentence frames were matched across conditions. In this way, the design of this experiment exerted tight experimental control on the SND manipulation. Experiment 1 also provided the opportunity to evaluate whether the SND effects demonstrated in previous isolated visual word recognition studies, especially those employing lexical decision tasks (e.g., Buchanan *et al.*, 2001; Pexman *et al.*, 2008; Yap *et al.*, 2012), generalised to a normal reading situation.

Experiment 1 was a sentence reading experiment in which participants read single-line sentences. Pairs of target words matched on word frequency, word length, number of phonemes, orthographic neighbourhood size, semantic plausibility and target word predictability were embedded within the same sentence frame. Target words were positioned in the middle of the sentences. It was ensured that the word before the target word was always three or more letters long (*mean* length = 5.21) maximising the chances that it was fixated (Radach & Kempe, 1993; Radach & McConkie, 1998). Participants read six practice sentences prior to the start of the experiment proper. A comprehension question was displayed for 15% of sentences to ensure that the participants were understanding the sentences.

Stolz and Besner's (1996) embellished interactive-activation (IA) framework (McClelland & Rumelhart, 1981) that includes a semantic level was considered to derive the predictions of the results of this experiment. This embellished IA model assumes that a word's semantic representation will initially become activated by the perceived word and this will happen prior to the perceived word's orthographic form being uniquely identified. It also assumes that feedback activation from the semantic level influences the speed with which a word is lexically identified. Employing this model to explain the lexical processing in normal reading, the visual information of the orthographic form of a currently fixated word can partially activate a set of orthographically similar word units (i.e. orthographic neighbours) along with the word unit of the fixated word itself. The word unit corresponding to the perceived word inhibits the activation of its orthographic competitors at the word level. Concurrently, activation feeds forward from the word level to the semantic level, activating the semantic representation of the perceived word. If the perceived word has high SND characteristics, then this word will have rich semantic representation at the semantic level due to the presence or the activation of its semantic neighbours that are closer (and more semantically similar) to it at the semantic level. As such the rich semantic representation associated with the high SND word will be activated and a greater amount of activation will feed back from the semantic level to the word level within the period that the candidate set is being reduced via processes of between-level activation and within-level inhibition. Thus, it was predicted that if a word's SND influences lexical identification, then decreased reading times with increased SND would be observed.

If the perceived word has low SND characteristics, then this word will have weaker semantic representation because there will be a network of distant (and semantically dissimilar) neighbours within the semantic level, which will only provide weak activation of the target word, with reduced, feedback of activation from the semantic level to the word level. As such, low SND will not have a comparable impact on lexical processing. Therefore, it was predicted that readers would exhibit significantly longer reading times on low than high

SND words in line with the findings of the previous lexical decision studies (e.g., Buchanan *et al.* 2001; Pexman *et al.*, 2008; Yap *et al.*, 2012) and in line with the predictions based on Stolz and Besner's (1996) embellished IA model (McClelland & Rumelhart, 1981).

Since SND was predicted to influence lexical processing, then the SND effects should be reflected in the fixation durations on the target words themselves, and potentially, subsequent words in the text if the effect spills over (as per Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Pollatsek, Juhasz, Machacek, & Rayner, 2008). If the SND manipulation only influences the ease with which a fixated word is lexically identified, then the SND effect should be short lived and should appear in only early reading times measures on target words (e.g., single fixation, first fixation and gaze duration). If the SND manipulation has a stronger longer lasting effect, and influences later stages of lexical processing and even produces effects that carry over into post-lexical processing, then this effect should also appear in late reading time measures (e.g., regression path duration and total reading time) on target words, and also may spill over onto the words following target words.

If the claim that a word to the right of fixation can influence the durations on the currently fixated word (i.e., parafoveal-on-foveal effects) is correct, then the SND characteristics of the target word should influence the fixation durations on the pre-target word. Such findings will be consistent with the parallel processing models of eye movement control during reading (e.g., the SWIFT model). If no parafoveal-on-foveal effects of the target words' SND are established, then such findings will give support to the serial processing models such as the E-Z Reader model.

3.1.1 Method

The analyses of Experiment 1 examined the basic SND effect in lexical processing while controlling for the extraneous variables that are well known to influence lexical processing during normal reading. Linear-mixed effect (LME) models (e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008) were used for data analyses in all the experiments reported in this thesis. LME models are a generalisation of linear regression that can include both *random factors* and *fixed factors* in one analysis. To explain, fixed factors (or fixed effects) refer to the factors/ covariates that researchers investigate. These factors can be categorical (e.g., whether the previous word is fixated) or can be continuous (e.g., the values of word frequency). Random effects refer to the random variations in data that are not investigated or are not the main interest of researchers, such as those variations related to target words selected for experiments (items) and those variations related to participants. LME models were used for analysing the data obtained from the experiments in this thesis for the following reasons. First, the inclusion of both fixed and random factors in one analysis allows us to assess whether a significant, or non-significant, effect is caused by the differences between individual participants or items (e.g., target words). That is, the use of LME models allow researchers to detect whether, for example, the investigated effects (e.g., the characteristics of words such as SND in this thesis) influence reading times independent of other un-investigated (or uncontrolled) factors (e.g., slow readers vs. fast readers). LME models make it also possible to include SND and other variables as continuous variables in the analysis and, therefore, considerable loss of statistical power resulting from dichotomising the variables, that would have been necessary as a requirement of ANOVA, is avoided (Cohen, 1983; MacCallum *et al.*, 2002). In addition, LME analyses are more flexible with missing data (i.e., the number of observations between participants and items are different), which is typical in eye-tracking research.

3.1.1.1 Participants

Forty-two participants took part in Experiment 1. All were students at the University of Southampton, with an age range of 18-30 years, were native English speakers, and all had normal vision or corrected-to-normal vision. None of the participants were dyslexic. Participants were awarded either course credits or given £2 for taking part.

3.1.1.2 Apparatus

EyeLink1000 eye tracking system (SR Research Ltd, Canada) was used to record eye movements. The sampling rate was 1000Hz. EyeLink1000 allows binocular recording of both eyes, but only the right eye was tracked in this experiment. The text was displayed just like a normal text (with lowercase letters, except where the uppercase letters were appropriate) in 14-point Courier New font on 21-inch ViewSonic CRT monitor at 1024 × 768 resolution. The participants were seated 70 cm from the monitor; approximately, 3 characters subtended one degree of visual angle. The text was presented in black on a white background.

3.1.1.3 Materials and Design

22 pairs of target words were manipulated for SND (22 words with high and 22 low SND). The index of SND used in this thesis was the Average Radius of Co-occurrence (ARC) provided by Shaoul and Westbury's (2010b) Neighbourhood Density Measures for 57,153 English words using the following settings in HiDEx: context size (co-occurrence words) = 10000 words; window size = 5 words ahead + 5 words behind; weighting scheme = inverse ramp; normalisation method = PPMI, similarity metric = cosine). In all experiments reported in this thesis, the British National Corpus (BNC) frequency, number of letters (word length), the number of orthographic neighbours were calculated for each word using *N-Watch* software (Davis, 2005).

Target words were embedded in 22 experimental sentence frames. Two lists of these sentences were created (list A, list B) with each list contained eleven sentences with high SND words and eleven sentences with low SND words. Each pair was matched on word frequency, word length, number of orthographic neighbours and number of phonemes (see Table 3.1). The high and low SND target words significantly differed in SND (t (42) = 11.57, p < 0.05), and did not differ on these controlled variables (all ps > 0.05) as can be shown in Table 3.1. An example of the stimuli with the manipulation of SND is given below (*badge* is a high SND word and *scarf* is a low SND word). The full stimuli set used in this experiment can be found in Appendix A.

She put her pink **badge**/**scarf** on the desk.

No sentence frame was read twice by any participant, and in total, each participant read twenty-two sentences. Before conducting Experiment 1, all the sentences were pre-tested for plausibility and predictability using pen and paper questionnaires. In the plausibility ratings, the participants were asked to rate how likely it was that the event in the given sentences would occur. These ratings were made on a 7-point Likert scale (1 = *very implausible*, 7 = *very plausible*) by two participant groups (twelve participants in each group that were drawn from the same population as those tested in the main reading experiment). The results showed that the sentences in both lists were plausible (list A *mean* = 5.63; list B *mean* = 5.64), and that they did not differ from each other in terms of plausibility (t (21) = -0.08, p = 0.93). A third group of twelve participants completed a predictability cloze test in which they saw the beginning of the sentences up to the word preceding the target words and were asked to complete the sentences with the most obvious word that came to mind. The result of the cloze test showed that none of the participants predicted the target words (total number of predicted target words = 0).

Table 3.1 *Means, Standard Deviations (in Parentheses) and t-Test of the Characteristics of the Target Words in Experiment 1.*

Variables	Low SND		High SND		*t*	*p*
	M	(SD)	M	(SD)		
SND	0.57	(0.04)	0.41	(0.06)	11.57	0.0005
Log BNC Frequency	4.80	(2.22)	4.70	(2.42)	0.97	0.33
Word Length	5.21	(1.35)	5.21	(1.25)	0.0001	1.0001
ON	4.13	(4.25)	4.13	(4.23)	0.0001	1.0001
Number of Phonemes	4.04	(1.16)	4.29	(1.08)	-0.77	0.44

Note. SND: semantic neighbourhood density; BNC: frequency from British National Corpus; word length: number of letters; ON: number of orthographic neighbours. M: mean; SD: standard deviation.

3.1.1.4 Procedure

Participants were first given an information sheet that contained general information about the experiment and a set of instructions, and then they were asked to sign a consent form if they agreed on taking part. Participants were asked to read the sentences normally, and to answer the comprehension questions that appeared after some sentences as accurately as possible by pressing a button box to indicate '*yes/no*' responses.

An initial calibration of the eye tracker was carried out. Viewing was binocular, but only the movements of the right eye were recorded. A chin rest and head rest were used to minimise head movements of the participants. A 3-point calibration presented horizontally across the middle of the screen where the sentence appeared, followed by a validation procedure. Calibration was rejected if the average error for all points was greater than 0.58 degree. A single point drift correction was performed before each sentence was read. The experimenter was able to view the text that the participant was reading on a separate monitor. If the experimenter detected that the gaze-tracking accuracy declined, a full calibration was performed before the next screen. Participants were instructed to read each sentence for comprehension. After reading the text on the screen, participants pressed a button on the back of the button box to move to the next screen. This button press caused either the next sentence or a comprehension question to be presented. The entire experiment lasted approximately 10-15 minutes.

3.1.2 Results

Prior to analysis, fixations less than 80ms and above 800ms were excluded. Fixation times above or below 3 standard deviations from the mean were also excluded. In total 1.09% of the data was removed.

The dependent variables for the target words were *first fixation duration* (the duration of first fixation on the word, regardless of whether the target word received one or more fixations), *single fixation duration* (the duration of the fixation when only a single fixation is made on the word), *gaze duration* (the sum of all fixations made on a word before the eyes move to another word), *regression path duration* (sum of all fixations from the first fixation on a word until a fixation to the right of the word), *total reading time* (the sum of all fixations on the word), and *skipping rate* (the probability that the target word does not receive a direct fixation during first-pass reading). A normal Quantile-Quantile plot (Wilk & Gnanadesikan, 1968) was obtained to check whether the fixation durations (the dependent variables of this experiment) were normally distributed. The plot indicated that all fixation durations were not normally distributed. Therefore, the fixation durations were log-transformed to approximate a normal distribution.

The SND effects were estimated using linear-mixed effect (LME) models (e.g., Baayen, 2008; Baayen, Davidson, & bates, 2008). The LME models were tested with lmer program of the lme4 package (Bates & Sakar, 2008) in the R environment for statistical computing (under the GHU General Public License, version 2.15.2, 64-bit build, R Development Core Team, 2012). LME were fitted using the restricted maximum likelihood method, specifying participants and items as crossed random effects. Then, the target word frequency, word length and orthographic neighbourhood size followed the SND metric (ARC, Shaoul & Westbury, 2010b) were entered in the models as fixed effects, one variable at a time. Entering the extraneous variables in the model before entering the SND metric ensures that any small variation in these variables will be ruled out before examining the effect of SND. Finally, interaction terms (frequency * ARC, and length * ARC, orthographic neighbourhood size * ARC) were subsequently added to the resultant model as fixed effects. All the fixed variables including the SND variable were entered as continuous variables, and were all centred at the means to minimise collinearity (whereby there are very high correlations among predictors) in the analysis of data. The following statistics will be reported in the results of this experiment: the regression coefficients (*b*), standard errors (*SE*), *t* values (or the *Z* value in the case of the skipping probability) together with *p* values based on Markov Chain Monte Carlo sampling 10,000 samples (Baayen *et al.*, 2008).

The overall mean comprehension rate was 98.9%, indicating that the participants read and understood the sentences. Note that in all experiments using Linear Mixed Modelling reported in this thesis, an effect is referred to as reliable, significant, or robust if the fixed effect coefficient has a *t* value of 2 or more; an effect is termed marginal if the coefficient has a *t* value between

1.645 and 1.96; an effect is termed unreliable or non-significant if the coefficient has a *t* value of less than 1.645 (as per Schad, Nuthmann, & Engbert, 2010). For main effects, the models' positive fixed effect coefficients indicate that the higher values of the fixed effects are associated with longer reading times while negative coefficients indicate that the higher values of the fixed effects are associated with shorter reading times.

All sentences were divided into four regions as shown in the examples below.

REGION	1	2	3	4
	She put her/	pink/	**badge**/	on the desk/.

The particular regions of interest for the analyses of this experiment were Region 2 constituting a pre-target word, Region 3 constituting a target word, and Region 4 constituting post-target words. If the post-target word immediately after the target word was a three-letter (or less) word, the next word was included as part of the post-target region. If the next word was also a three- (or less) letter word, then the following word was included as part of the post-target region. This criterion for determining the post-target region maximised the chances that the post-target region was fixated (Radach & Kempe, 1993; Radach & McConkie, 1998).

The reported results will include 1) target words' SND effects on the fixation times on pre-target words (i.e., parafoveal-on-foveal effects), 2) the immediate effects of target words' SND characteristics on their fixation times, and 3) the spillover effects of SND on the fixation times on post-target words.

To investigate whether the SND characteristics of the target words influenced the fixation durations on the pre-target words, LME analyses were conducted. The reading time measures for the pre-target words were the dependent variables; participants and items were entered as random effects, and the predictor of target words' SND along was entered as fixed effects. There were no reliable effects for these analyses (all *t*s < 0.6) providing no evidence of any parafoveal-on-foveal effects (see Table 3.2 for the mean values associated with these analyses).

Table 3.2 *Means of Fixation Times on the Pre-Target Words Preceding High and Low SND Target Words in Experiment 1. Standard Deviations are Given in Parentheses.*

	High SND	Low SND
Single Fixation Duration	211 (55)	217 (75)
First Fixation Duration	215 (73)	215 (78)
Gaze Duration	242 (108)	240 (113)
Regression Path Duration	281 (168)	282 (171)

Note. SND: semantic neighbourhood density.

Next, the immediate effects of SND on the fixation times on the target words were examined. Table 3.3 and 3.4 list the results of the LME analyses carried out on the fixation times on the target words. As can be seen from these two tables, in the baseline models the number of orthographic neighbours was not a significant predictor in any of the early or late reading time measures. The effect of word length was not significant in all reading time measures, except in a late reading time measure of regression path duration (b = 0.047, SE = 0.017, t = 2.8, p < 0.05), with long words fixated for a longer time compared to short words. The effect of word frequency was marginally significant in some early reading time measures (first fixation duration: b = -0.661, SE = 0.356, t = -1.86; gaze duration: b = -0.761, SE = 0.404, t = -1.89) and the late reading measures of total reading time (b = -2.744, SE = 1.647, t = -1.667); reading times were decreased when the words were of high frequency compared to when they were of low frequency. These findings of not being able to instantiate robust effects of the lexical variables are not surprising given that the target words selected for this experiment did not vary in word frequency, word length or orthographic neighbourhood size as the target words here were closely matched on these variables.

To examine the SND effects in this experiment in which lexical variables were tightly controlled, the SND metric of the target words (ARC, see Chapter 2, Section 2.2.3.2 for a through description of this metric) was then entered in the LME models. The results showed that the target words' SND was a significant predictor in all early reading time measures (single fixation duration: b = -123.37, SE = 23.68, t = -5.21, $p < 0.05$; first fixation duration: b = -54.005, SE = 16.0002, t = -3.38 $p < 0.05$; gaze duration: b = -58.242, SE = 20.187, t = -2.89, $p < 0.05$) and the late measure of total reading time (b = -193.49, SE = 85.221, t = -2.270, $p < 0.05$). As the negative coefficients of the model indicated, the increased SND was facilitatory. That is, reading times decreased when the target words were high SND words, compared to when they were low SND words, consistent with the predictions. However, SND did not predict the skipping probability (after running a logistic LME) (b = 3.004, SE = 1.67, Z = 1.79, $p > 0.05$) as well as SND did not predict the refixation probability (b = 3.66, SE = 1.31, Z = 1.12, $p > 0.05$). Recall that the decision to skip a word must be made early in processing (when the word is in the parafovea). Given this and consistent with the findings of parafoveal-on-foveal effects reported earlier in this section, it is likely that information about the target words' SND was not obtained parafoveally (before fixating the word). As such, SND did not influence to the skipping probability.

To gain a general insight into the nature of the significant SND effect, the pattern of effects observed in the mean of reading times for each measure was considered (see Table 3.5). Recall that it was predicted that high SND words would be processed faster than low SND words. As explained earlier, a high SND word has a strong semantic representation due to having closely packed semantic neighbours (i.e., semantically similar neighbours). A low SND word, on the other hand, has a weaker semantic representation due to having distant semantic neighbours (i.e., semantically less similar neighbours). As such, the high SND word will benefit from having enhanced semantic feedback activation (provided by high SND characteristics) sent to the word level. This enhanced semantic feedback activation contributes to resolving competition between orthographic neighbours at the word level, and, thus, helps in constraining word identification. A low SND word, on the other hand, will have weaker semantic feedback activation (provided by the low SND characteristics) sent to the word level, and thus will not have a comparable impact on word identification.

As can be seen from Table 3.5, the average reading times for high SND words were in general less than those for low SND words in all reading time measures. Specifically, the differences between fixation times on high SND and low SND words were shown in early reading time measures (14ms in single fixation duration, 9ms in first fixation duration, and 20ms in gaze duration), with high SND words fixated for less time than low SND words. The same pattern of findings was also obtained in the later reading time measures (29ms in regression path duration and 48ms in total reading time); again reading times were longer for low SND words compared to high SND words. The pattern of fixation times based on the LMEs and the means is consistent with predictions made in the Introduction of this experiment.

Examination of the interactive effects in this experiment indicated that all interaction terms failed to reach significance in this experiment, which is not striking given that this experiment tightly controlled for word frequency, word length and orthographic neighbourhood size. Thus, interactive effects may not be apparent in this experiment.

Table 3.3 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Single Fixation, First Fixation and Gaze Duration Recorded on the Target Words as Dependent Variables (Experiment 1).*

	Log Single Fixation Duration			Log First Fixation Duration			Log Gaze Duration		
	Variance	Std. Dev.		Variance	Std. Dev.		Variance	Std. Dev.	
Random effects									
Item (intercept)	183.2	13.54		33.59	5.796		76.81	8.764	
Subject (intercept)	423.6	20.58		289.36	17.011		324.99	18.028	
Residual	1393.6	37.33		1458.78	38.194		1433.47	37.861	
AIC			4064.725			6984.946			5970.723
BIC			4080.630			7003.046			5988.189
logLik			-2028.362			-3488.473			-2981.361
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects									
Intercept	203.371	5.251	38.73	202.766	3.644	55.64	206.010	4.102	50.22
Freq. n	-0.605	0.526	-1.15	-0.661	0.356	-1.86*	-0.761	0.404	-1.89*
Length n	2.312	2.562	0.90	1.264	1.459	0.87	0.712	1.852	0.38
ON n	-0.231	0.738	-0.31	0.007	0.489	0.02	-0.081	0.593	-0.14
ARC	-123.37	23.68	-5.21*	-54.005	16.0002	-3.38*	-58.242	20.187	-2.89*
Interactions									
Freq. n * ARC	4.216	5.887	0.72	5.201	4.049	1.28	7.827	4.888	1.60
Length. n * ARC	17.441	18.580	0.94	21.277	12.945	1.63	22.4698	16.053	1.40
ON. n* ARC	-7.536	5.247	-1.44	-8.367	4.575	-1.62	-5.605	5.598	-1.00

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients ($1.645 \leq t < 1.96$); robust significant coefficients ($t \geq 2$). No significant coefficients ($t < 1.645$).

Table 3.4 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Regression Path Duration and Total Reading Time Recorded on the Target Words as Dependent Variables (Experiment 1).*

	Log Regression Path Duration			Log Total Reading Time		
	Variance	Std. Dev.		Variance	Std. Dev.	
Random effects						
Item (intercept)	0.012	0.111		2162	46.49	
Subject (intercept)	0.027	0.166		2847	53.36	
Residual	0.162	0.403		18052	134.36	
AIC			951.030			10315.26
BIC			969.896			10334.03
logLik			-471.515			5153.63
Fixed effects						
Intercept	5.581	0.040	136.34	323.437	14.718	21.976
Freq. n	-0.006	0.004	-1.56	-2.744	1.647	-1.667*
Length n	0.047	0.017	2.8*	9.839	7.067	1.392
ON n	0.0001	0.007	0.02	-0.599	2.438	-0.246
ARC	-0.273	0.214	-1.28	-193.49	85.221	-2.270*
Interactions						
Freq. * ARC	0.057	0.053	1.06	17.525	21.198	0.827
Length * ARC	0.166	0.163	1.02	57.9269	67.1231	0.863
ON* ARC	-0.044	0.050	-0.89	4.2299	23.9165	0.177

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients (1.645 ≤t<1.96); robust significant coefficients (t≥2). No significant coefficients (t<1.645).

Basic SND Effects

Table 3.5 *Means of Reading High SND Target Words and Low SND Target Words in Experiment 1. Standard Deviations are Given in Parentheses.*

	High SND	Low SND
Single Fixation Duration	221 (19)	235 (30)
First Fixation Duration	217 (14)	226 (20)
Gaze Duration	241 (31)	261 (20)
Regression Path Duration	275 (55)	304 (49)
Total Reading Time	303 (57)	351 (61)

Note. SND: semantic neighbourhood density.

The last aspect of the data that was considered concerned whether effects due to the SND characteristics of the target words spilled over onto subsequent words. LME models were conducted. The LME models included the reading times of the post-target words (n+1) as dependent variables, participants and items as random effects, and the target words' SND metric (ARC) as a fixed effect. The results showed that the target words' SND effects spilled over onto the subsequent words. Specifically, the target words' SND was a significant predictor in gaze duration (b = -0.664, SE = 0.210, t = -3.15, $p < 0.05$), regression path duration (b = -0.873, SE = 0.289, t = -3.02, $p < 0.05$), while this spillover effect was marginal in single fixation duration (b = -0.493, SE = 0.295, t = -1.67) and first fixation duration (b= -0.443, SE = 0.264, t = -1.67). As indicated by the negative coefficients, fixation times on the post-target words decreased when the previous words were high SND words. This pattern of effect found in LMEs is consistent with the pattern of effect observed in the means of reading times of the post-target words (see Table 3.6). In all of the reading time measures, the mean fixation times on the post-target words were shorter following high SND target words than low SND target words.

Table 3.6 *Means of Fixation Times on the Post-Target Regions following High and Low SND Target Words in Experiment 1. Standard Deviations are Given in Parentheses.*

	High SND	Low SND
Single Fixation Duration	246 (101)	252 (102)
First Fixation Duration	240 (108)	247 (78)
Gaze Duration	263 (121)	275 (125)
Regression Path Duration	327 (228)	332 (248)

Note. SND: semantic neighbourhood density.

3.1.3 Discussion

Experiment 1 investigated 1) whether target words' SND effect could influence the reading times on prior words (i.e., parafoveal-on-foveal effects), 2) whether the main effect of target words' SND could influence their lexical processing during normal reading, and 3) whether this SND effect would influence the reading times on subsequent words (i.e., spillover effects).

3.1.3.1 Parafoveal-on-Foveal Effects of SND

The analyses of Experiment 1 were carried out to investigate the possibility of finding the parafoveal-on-foveal effects. The results of the current analyses showed that the SND characteristics of the target words did not influence the fixation times on the previous words. That is, information about the meaning of the parafoveal words (i.e., the target words in this experiment when they were in the parafovea while fixating the pre-target words) was not found to affect the lexical processing of the foveal words (i.e., the pre-target words in this experiment). The results of Experiment 1 provide no evidence for the claim that the processing of a currently fixated word is affected by the semantic characteristics of the parafoveal word, which is in line with the findings of many eye movement studies (e.g., Altarriba *et al*., 2001; Rayner, Balota, & Pollatsek, 1986; Rayner & Morris, 1992; Rayner, White, Kambe, Miller & Liversedge, 2003).

The findings that no parafoveal-on-foveal effects were observed in this experiment might be considered to be inconsistent with the predictions of the SWIFT model. Since the SWIFT model allows parallel processing of multiple words at a time, it predicts that the characteristics of a parafoveal word can influence the fixation durations on the foveal word, which was not the case in the findings of this experiment. In contrast, these findings are in accord with the assumptions of the E-Z Reader model. In this model, parafoveal processing occurs only after the lexical processing of the foveal word has been completed and the programming of a saccade to the parafoveal word has been initiated. Specifically, parafoveal processing happens during the time that attention is on the parafoveal word but the eyes are still on the foveal word. Accordingly, the E-Z Reader model predicts that the lexical and semantic characteristics of the parafoveal word would not influence the fixation durations on the foveal word, a prediction that was met by the findings of Experiment 1.

3.1.3.2 Immediate SND Effects

This part of the analyses was carried out to examine whether target words' SND influenced their reading times, especially early measures of eye movements that are associated with lexical processing. Consistent with the predications, it was found that the main effect of SND was significant and facilitatory. That is, high SND words were fixated for a shorter time than low SND words. This SND effect appeared in the early and late measures of eye movements, suggesting that the SND characteristics of the target words influenced lexical identification processes.

The findings of this experiment are consistent with those of visual word recognition studies that have generally showed that increased SND is associated with quicker response latencies in lexical processing paradigms (e.g., Buchanan, *et al.*, 2001; Siakaluk *et al.*, 2003; Yates *et al.*, 2012). The present findings also contribute to resolving the debate in these visual word recognition studies as described in Section 2.3 in Chapter 2. To reiterate, these studies consistently found that denser semantic neighbourhood words were responded to faster than sparser semantic neighbourhood words in lexical decision tasks (e.g., Buchanan *et al.* 2001; Shaoul & Westbury, 2010a). On the other hand, the findings of SND effects were inconsistent in tasks that required comparatively deep semantic analysis of the meaning of the presented words (e.g., semantic categorisation tasks). With these paradigms some researchers have reported a facilitatory effect of increased SND (Siakaluk *et al.*, 2003), others have reported an inhibitory effect of increased SND (Shaoul & Westbury, 2010a) and still others found a null SND effect (Pexman *et al.*, 2008; Yap *et al.*, 2011, Yap *et al.*, 2012). Since one of the aims of the visual word recognition studies is to investigate how different variables influence normal word identification, studying the SND effects using eye movement recording whilst reading normally is important to draw conclusions about whether these effects occur in an ecologically valid task without the artefacts of isolated visual word recognition tasks. The results of Experiment 1 indicate that the SND characteristics of the target words are actually influencing their lexical processing in normal (silent) reading and that the effect of increased SND is facilitatory rather than inhibitory and is not the consequence of the specifics of laboratory tasks.

The findings of this experiment is also in line with the recent findings of three eye movement studies that investigated the effect of some aspects of word meaning in lexical processing during normal reading. Duñabeitia, Avilés, and Carreiras, (2008) found a significant and facilitatory effect of the number of semantic associates in gaze duration and total reading time. Similar findings were also obtained by Cook, Colbert-Getz, and Kircher (2013) who reported a significant effect of the number of semantic features (e.g., colour, taste, etc.) on gaze duration and total reading time. Interestingly, Plummer, Perea, and Rayner (2014) found that the semantic effect of contextual diversity (i.e., the number of passages in which a word appears) significantly influenced all early and late reading time measures (first fixation, single fixation, and gaze duration along with regression path duration, and total reading time) when controlling for word frequency. Particularly, reading times for words with high contextual diversity were significantly shorter than for words with low contextual diversity. The present study also observed a significant and facilitatory effect of SND, as defined by the average semantic similarity of a word and all its semantic neighbours that fell within a specified threshold

(Shaoul & Westbury, 2010a), in early and late measures of eye movements, suggesting the effects of SND influenced lexical processing and integrating the word meaning into the sentential context. These eye movement studies along with the present study all examined different aspects of semantic characteristics, and they provide preliminary evidence that the semantic properties of words can influence lexical processing in reading. The present study specifically demonstrated that the effects of the SND emerged early in reading time measures such as single fixation and first fixation duration, which is not surprising given that Plummer *et al.*'s (2014) study established the early effects of some aspects of the semantic characteristics a word had in its lexical identification in normal reading.

The increased SND advantage can be explained by Stolz and Besner's (1996) embellished interactive-activation (IA) model (McClelland & Rumelhart, 1981). To recap on the assumptions of the IA model, there is bidirectional cascaded processing between levels (feedforward activation and feedback activation between levels), and the connections between levels are either excitatory or inhibitory while the connections between the units within the same level are always inhibitory (i.e., within-level competition). The activations of the units within a level compete with each other so that the strong candidates inhibit the weak candidates until there is one most active candidate. Stolz and Besner (1996) added a semantic level at the top of the levels (the word level, letter level, and visual feature level) in the original IA model as explained in Section 1.3 in Chapter 1. In this thesis, it was assumed that the SND effect resides at the semantic level concurrently with feedback activation from the semantic level to the word level. To explain, a word's semantic neighbours, for example, have to be activated first so that they can have an impact. These semantic neighbours are only activated via the feedforward activation from the word level to the semantic level. Therefore, their effects should appear only after this feedforward activation has taken place.

As such, this model can accommodate the facilitatory effect of SND in lexical processing in normal reading as follows. When a word is perceived, the visual information of the orthographic form of the currently fixated word partially activates a set of orthographically similar word units along with the fixated word unit at the word level. The activation of the word unit corresponding the perceived word inhibits the activation of its orthographic neighbours. At the same time, activation feeds forward from the word level to the semantic level, activating the semantic representation. If the perceived word has high SND characteristics, then its sematic neighbours will be closer (and more semantically similar) to the word at the semantic level. These closely packed semantic neighbours will provide a greater amount of activation at the semantic level. Therefore, a great amount of activation will feed back from the semantic level to the word level within the period that the candidate set is being reduced via processes of between-level activation and within-level inhibition. As such, high SND can facilitate word identification. If the perceived word has low SND characteristics, on the other hand, then its semantic neighbours will be distant at the semantic level. These distant neighbours will only provide weak activation at the semantic level and, thus, reduced feedback activation from the semantic level to the word level. That is, the effect of low SND will not have a strong impact on unique word identification, as the high SND effect will. Based on this theoretical account, a high SND word is identified faster than a low SND word, a prediction that was met in the present study.

An explanation of why words with high SND characteristics have rich sematic representations (and, thus, enhanced semantic activation) at the semantic level can be based on Reichle and Perfetti's (2003) suggestions. According to Reichle and Perfetti, a word has a strong representation depending on the frequency with which (orthographic, phonological or semantic) information is encoded, as well as on the word's similarity to other words in long-term memory. They also assumed that over one's lifetime of experience and skilled reading, one learns to use the word's form and semantic relations that are shared among words in one's language to make the reading process efficient and rapid. Accordingly, a high SND word is semantically similar to many other words that are stored in a reader's long-term memory. As such, the high SND word is well inter-connected to many other words, which strengthens its semantic representation.

The current findings of the immediate SND effect can be explained by both the E-Z Reader model (e.g., e.g., Reichle, Pollatsek, & Rayner, 2006) and the SWIFT model (e.g., Engbert, *et al.*, 2002, 2005) that were introduced in Section 1.1 in Chapter 1. According to the E-Z Reader model, the lexical processing of word n occurs in two separate stages: L1 and L2. The L1 stage is a familiarity check stage in which the familiarity of the fixated word is assessed. Once the L1 stage is complete, two stages occur simultaneously: (1) the eye movement system begins programming the next saccade, and (2) the L2 of lexical processing stage starts. In the L2 stage, the fixated word is fully identified by accessing its semantic and contextual-appropriate meaning. When the L2 stage is complete, the reader redirects attention to the next word while the eyes are still on the currently fixated word. It was suggested that single and first fixation duration should reflect processing taking place during L1 processing on word n (Reinglod, Yang, & Rayner, 2010; Sheridan & Reingold, 2013). Since the present study found that the immediate SND effect on word n appeared in single fixation and first fixation duration, one can infer that the SND effects influence L1 processing on word n. To explain, the L1 lexical processing (familiarity check) on word n starts early even before the word n is fixated (during the parafoveal preview while the eyes are still on the prior word), and the L1 processing continues when the parafoveal word is later fixated. Therefore, the orthographic and phonological processing of word n can occur when word n is in the parafovea (Inhoff & Topolski, 1994; Sereno & Rayner, 2000). This allows enough time for the enhanced semantic representation of word n (particularly high SND characteristics) to be activated during the L1 lexical processing and to be used to assess the familiarity of word n when it is later fixated. The semantic representation of a low SND word will be weak to influence the initial stage of assessing the fixated word's familiarity due to the nature of the fixated word's semantically dissimilar neighbours. As such, high SND can facilitate the completion of L1 stage of lexical processing while low SND should not influence the preliminary stage of lexical processing.

The findings of the immediate SND effects can also be accounted by the SWIFT model (Engbert *et al.*, 2002, 2005). Recall that in the SWIFT model, the target of a saccade is selected from an activation field (i.e., words of a sentence) which evolves over time depending on linguistic and visual processing, and that a word with the highest activation (e.g., a difficult-to-process word) in the activation field is selected as the next saccade target. Based on this principle, if a foveal word is highly activated at the time of saccade target selection, then the random timer will be inhibited from executing a forward saccade. Therefore, the foveal word will receive a refixation. Accordingly, if word n is a low SND word (i.e., a difficult-to-process word), then the activation of the low SND word will be high at the time of selecting a saccade target. As a consequence, the high activation of the low foveal SND word will inhibit the random timer from executing a forward saccade. Thus, the foveal low SND word will be refixated (i.e., this word will be fixated for a long time). If the currently fixated word is a high SND word (i.e., an easy-to-process word), on the other hand, the activation of this high SND word will be lower than the case of a low SND word. As a consequence, the random timer will not be inhibited by the foveal word. Instead, the timer will initiate a new saccade program to the subsequent words with the highest activation within the activation field. That is, a high SND word should be fixated for less time than a low SND word.

3.1.3.3 Spillover Effects of SND

The analyses of the Experiment 1 were also carried out to examine whether the effects of target words' SND characteristics spilled over onto subsequent words. The analyses of Experiment 1 indicated that the target words' SND characteristics influenced the fixation times on the subsequent words in early and late reading time measures, with shorter reading times on the subsequent words following high SND words.

The spillover of SND effects found in this experiment can be explained by the E-Z Reader model (e.g., Reichle, & Rayner, 2006; and Pollatsek *et al.*, 2008) as follows. Since the L2 lexical processing is assumed to reflect processes of accessing the meaning of word n, then the SND characteristics are assumed to influence the L2 processing of the currently fixated word (n). If the currently fixated word n has high SND (i.e., an easy-to-process target word), then the L2 processing on this word will be completed faster than if the word has low SND characteristics because of the enhanced semantic activation associated with high SND words. Therefore, the quick completion of the L2 processing (as a consequence of having high SND characteristics) will allow more parafoveal preview of word n+1 (a post-target word) (i.e., the time between attention has shifted to word n+1 and before the eyes start to move away from word n) compared to low SND words. During this parafoveal preview, the familiarity check on word n+1 is carried out while still fixating the target word. Following a high SND word, the post-target word will have a head start when it is fixated, as a great amount of its familiarity check will have already been carried out while fixating the previous word.

The SWIFT model of eye movement control (Engbert *et al.*, 2002, 2005) can also account for the spillover effect found in this experiment. According to the SWIFT model, the fixation durations on a word are influenced by the amount of foveal activation. If the currently fixated word (n) is a difficult-to-process word, it is more likely that this word will be refixated by means of the foveal inhibition mechanism as mentioned in the previous section. Because of this foveal inhibition associated with processing a difficult fixated word, parafoveal processing of word n+1 will be reduced. As a consequence, fixation durations on word n+1 when it is subsequently fixated are longer than on average. However, if word n is an easy to process word (e.g., a high SND word), then the saccadic system (or the random timer) will not be modulated or inhibited by the properties of the fixated word, which means that there will be more parafoveal processing of word n+1. Therefore, subsequent fixation durations on word n+1 are shorter following a high SND word than a low SND word.

3.1.4 Conclusion

Experiment 1 provides clear evidence that SND characteristics of the currently fixated word can influence lexical identification during normal reading. It was found that increased SND, defined by the distance between a word and all its semantic neighbours falling within a specified threshold (ARC, Shaoul & Westbury, 2010a), facilitated the lexical processing of the target words and the subsequent words. Therefore, these findings provide evidence in support of Shaoul & Westbury's (2010a) conceptualisation of the nature and the influence of semantic representations during lexical processing. These SND effects can be explained by assuming enhanced semantic feedback in the case of high SND words, which facilitates their word identification, compared to words with weaker semantic representations (e.g., low SND words) and, thus, weaker semantic activation. The findings also suggest that target words' SND has a longer lasting effect that it influences the lexical processing of subsequent words. Given this initial study, it was decided to explore how the basic SND effect is modulated by other lexical variables by conducting a second experiment in which target words were not tightly controlled for the variables that influence lexical processing (i.e., a corpus study). The second experiment reported in Chapter 4 would be an exploratory study that would provide the basis for further experimentations that will be reported in the subsequent chapters in this thesis.

Chapter 4: Interactive Effects of Semantic Neighbourhood Density in Normal Reading

The findings of Experiment 1 showed that increased SND plays a facilitatory role in lexical identification in normal reading. This finding gives the motivation to explore whether the basic SND effect interacts with other well-established effects during lexical processing in normal reading. Therefore, Experiment 2 was conducted, as an exploratory corpus-based study, that investigated whether the basic SND effect observed in Experiment 1 would be replicated, and whether this effect was modulated by other variables that are well known to influence lexical processing. In this way, Experiment 2 will reveal the potential variables that interact with SND during lexical identification, hence, will provide a basis for conducting further experimentations on the effects of SND.

4.1 Experiment 2

Experiment 2 was designed to provide an answer to the question of whether target word frequency (and potentially other variables) might modulate the basic effect of target words' SND in lexical identification in normal reading, and whether the basic SND effect found in Experiment 1 would be replicated in this experiment. A corpus-based approach to these questions was adopted (as per Pynte & Kennedy, 2008; Schad, Nuthmann, & Engbert, 2010), constructing passages of text including numerous words that varied in relation to their SND characteristics. In the analyses of the eye movement data obtained in this experiment, reading times for all content words for which an index of SND could be obtained were examined.

Target word frequency was allowed to vary in Experiment 2 because this would provide the opportunity to test whether there would be an interaction between target words' SND and target word frequency. The effect of word frequency has been found to be robust in reading time measures associated with lexical identification (Inhoff & Rayner, 1986; Sereno & Rayner, 2000; Schilling, Rayner, & Chumbley, 1998), with lower frequency words fixated for a longer time than higher frequency words. According to the embellished IA model (Stolz & Besner, 1996), the visual information of a perceived word partially activates its word unit and, to a lesser extent, other orthographically similar word units (i.e., orthographic neighbours) at the word level. The word unit corresponding to a higher frequency word has a higher baseline level of activation compared to a lower frequency word. That is, a high frequency word unit is more active compared to a low frequency word unit at the word level. The activation of the word unit corresponding to a high frequency word will rapidly inhibit the activation of its orthographic neighbours within the word level relative to a lower frequency word. To explain, the degree to which a word representation inhibits other competitor representations at the word level is determined by the degree of its own activation relative to those of the competitors. In other words, a more active word representation will more rapidly inhibit competitors than a less active word representation. Because of this quick inhibition of the orthographic competitors at the word level, activation will quickly feed forward from the word level to the semantic level, activating the sematic representation at the semantic level. That is, the high frequency word will allow an opportunity for its semantic representation to be activated via feedforward activation from the word level to the semantic level and to influence lexical processing. If the high frequency word's semantic representation happens to be rich (e.g., high SND characteristics), the rich semantic representation will be activated and a greater amount of activation will feed back from the semantic level to the word level within the period that the candidate set is being reduced via processes of between-level activation and within-level inhibition. If its semantic representation is weaker (e.g., low SND), then the weak semantic representation will not have a comparable impact on lexical processing due to the nature of its distant semantic neighbours as discussed in the previous chapter.

A low frequency word, on the other hand, has a lower baseline level of activation at the word level and will take longer to inhibit the activation of other orthographically similar words because a word unit corresponding to a low frequency word is less active than the word units corresponding to its orthographic competitors. Because the low frequency word unit takes a longer time to inhibit its orthographic competitors at the word level, there will be very weak activation feeding forward from the word level to the semantic level. Concurrently, the word level will receive activation from the lower levels (the letter level and indirectly the visual feature level) that can contribute to resolving the competition between the orthographic units at the word level. In other words, there will be sufficient visual information provided by the lower levels (letter and visual feature levels) with which the low frequency word can be identified. As such, activating its semantic representation may have a little, if any, benefit in constraining its word identification

Based on the embellished IA model, it was predicted that there would be an interaction between SND and word frequency, and this interaction was predicted to influence lexical identification, as would be evident in the early reading time measures (e.g., single fixation, first fixation and gaze duration).

4.1.1 Method

The analyses in Experiment 2 examined the role of SND in lexical processing while relaxing variables such as word frequency, word length and orthographic neighbourhood size. LME models were also used to analyse the data of this experiment. In addition to the reasons for using LME models mentioned in Section 3.1.1 in Chapter 3, the use of LME models also allows researchers to utilise many items (e.g., target words in the case of this thesis) in their experiments without necessarily controlling for variables that are well known to influence lexical processing, thus, keeping up the statistical power that otherwise would have been reduced if analysis of variance (ANOVA) was used. Thereby, using LME models in Experiment 2 makes it possible to include a large number of target words with a wide range of values of SND.

4.1.1.1 Participants

Forty-nine participants at the University of Southampton, selected according to the same criteria as those for Experiment 1 took part in Experiment 2. Participants were awarded either course credits or given £3.

4.1.1.2 Apparatus

The apparatus was the same as Experiment 1. The text was displayed left aligned and double spaced (the space between two lines was doubled).

4.1.1.3 Materials and Design

Three text passages that contained the target words were constructed. The passages were constructed so that the target words were not predictable by the prior context, and the coherence level of the passages was acceptable to English native speakers in terms of making sense. These passages were passed through three native speakers of English who did not participate in the experiment to check their sensibility and grammaticality; these three native speakers added their comments on how to improve (smooth) these passages so they were more digestible to an English native speaker. The final polished versions of these passages were presented to the participants (the full materials used in this experiment can be found in Appendix B). The experimental passages were presented to the participants in a random order after a practice passage that was presented at the beginning of the experiment. After each passage, there were two comprehension questions to answer by pressing one of the two buttons in in a button box to indicate *'yes'* or *'no'*. The comprehension questions were included to ensure that participants read and understood the passages.

Each passage contained 60 target words that were manipulated for SND (30 words with high SND, 30 words low SND). Each experimental passage was divided in five pages (computer screens); on average there were 9.4 lines per page and 44 lines per passage. The length of each passage was around 600 words (561 words in Passage 1, 647 words in Passage 2, and 532 words in Passage 3). The target words were distributed throughout the passage (some of them appeared towards the end of the sentences, close to the beginning of the sentences, and in the middle of the sentences).

The descriptive statistics for the characteristics of all high SND words and low SND words in Experiment 2 are presented in Table 4.1. The target words' average ARC (the SND metric) was 0.44 in Passages 1 and 3 and was 0.45 in

Passage 2. The target words significantly differed in the SND measured by ARC ($t\,(179) = 17.51$, $p < 0.05$).

Table 4.1 *Means and Standard Deviations (in Parentheses) of the Characteristics of the Target Words in Experiment 2.*

	High SND	Low SND
	M (SD)	M (SD)
SND	0.58 (0.05)	0.33 (0.06)
Log BNC	32.29 (55.94)	3.80 (2.4)
Length	6.33 (1.73)	6.23 (1.31)
ON	2.03 (3.61)	2.13 (2.69)

Note. SND: semantic neighbourhood density; log BNC: log word frequency drawn from British National Corpus; length: number of letters; ON: orthographic neighbourhood size; M: mean; SD: standard deviation.

4.1.1.4 Procedure

The procedure followed in this experiment was the same as Experiment 1, except that the calibration and validation were carried out for 9 points presented horizontally and vertically across the whole screen. At the end of reading each passage, two comprehension questions were presented to the participants. The entire experiment lasted approximately 20-30 minutes.

4.1.2 Results

Prior to data analyses, data trimming was carried out following the same criteria as in Experiment 1. After fixation trimming and removing outliers above and below three standard deviations from the means, 2.37% of data was removed prior to the analyses. The dependent variables were single fixation duration, first fixation duration, gaze duration, regression path duration, total reading time and skipping rate (see Section 3.1.2 in Chapter 3 for a description of the measures). Based on the normal Quantile-Quantile plot (Wilk & Gnanadesikan, 1968), these fixation times were log-transformed as in Experiment 1 to approximate a normal distribution.

As in Experiment 1, the SND effects were estimated in Experiment 2 using LME models. Participants and items were entered as random effects. Then, the effects of visual and lexical variables associated with the currently fixated word n were examined by including their word frequency, word length, and number of orthographic neighbours (cf. Pollatsek, Juhasz, Reichle, Machecek, & Rayner, 2008), one variable at a time. These variables were included in the models due to the fact that these variables were not tightly controlled in this experiment. By including these variables first in the models, their effects will be, to some degree, statistically partialled out before examining the unique contribution of the SND metric. Then, the effects of previous word n-1 on the fixation durations on the word n were tested by including word n-1 frequency and word n-1 length as predictors (Rayner & Duffy, 1986). The effects of the characteristics of the previous words were tested here mainly because this experiment did not control for the sentence frames in which the target words were embedded. Then, the words' SND metric (ARC, Shaoul & Westbury, 2010b) was entered in the model as a fixed effect. After this, the interaction terms (frequency of n * ARC, length of n * ARC, orthographic neighbourhood size * ARC, frequency n-1 * ARC, length of n-1 * ARC) were added to the resulting model.

All fixed effects including the SND variable were entered as continuous variables, and were all centred at the means to minimise collinearity in the analysis of the data (whereby there are very high correlations among predictors). The following statistics will be reported in the results of this experiment: the regression coefficients (*b*), standard errors (*SE*), *t* values (or the *Z* value in the case of the skipping probability) together with *p* values based on Markov Chain Monte Carlo sampling 10,000 samples (Baayen *et al.*, 2008).

The overall mean comprehension rate was 95.8%, indicating that the participants read and understood the passages. The results of Experiment 2 will include the LME analyses pertaining to the effect of the target words' SND on target word fixation times. Spillover effects and parafoveal-on-foveal effects were not examined in this experiment, as the sentential frames in which the pre-target and post-target words were embedded varied greatly. In addition, the pre-target and post-target words varied in a multitude of characteristics such as word frequency, word length, orthographic neighbourhood size and syntactic category.

Table 4.2 indicates that in the baseline model word frequency of the target words significantly accounted for variability in single fixation duration ($b = -0.006$, $SE = 0.003$, $t = -2.09$, $p < 0.05$), first fixation duration ($b = -0.005$, $SE = 0.002$, $t = -2.03$, $p < 0.05$), gaze duration ($b = -0.008$, $SE = 0.003$, $t = -2.28$, $p < 0.05$), and its significance was marginal in total reading time ($b = -0.0003$, $SE = 0.0002$, $t = -1.65$) (see Table 4.3). As can be seen from the negative signs of the *b* coefficients and the *t* statistics, word frequency exerted a facilitatory effect on all reading time measures (i.e., higher word frequency resulted in decreased time spent on reading the target words). Target word length was also a significant predictor in total reading time ($b = 0.02$, $SE = 0.009$, $t = 2.68$, $p < 0.05$) (see Table 4.3), and approached significance in gaze duration ($b = 0.01$, $SE = 0.007$, $t = 1.99$, $p < 0.07$) (see Table 4.2). As the positive signs of the *b* and *t* values indicated, longer words were fixated for a long time compared to shorter words. Note also that this effect was less robust in the first fixation and single fixation measures due to the fact that it is primarily driven by refixations on words. The significant effect of orthographic neighbourhood size appeared somewhat later in reading time measures as evident in regression path duration ($b = -5.327$, $SE = 1.542$, $t = -3.45$, $p < 0.05$) and total reading time ($b = -4.415$, $SE = 1.742$, $t = 2.53$, $p < 0.05$) (see Table 4.3). The influence of orthographic neighbourhood size was a late facilitatory effect, that is, words with a high number of orthographic neighbours were read for a shorter time than words with a low number of orthographic neighbours.

SND did not predict the skipping probability (after running a logistic LME) ($b = 0.069$, $SE = 0.46$, $Z = 1.486$, $p > 0.12$) as well as SND did not predict the refixation probability ($b = -1.48$, $SE = 1.08$, $Z = -1.37$, $p > 0.15$). SND also did not predict any reading time measures (all *t*s < 1.26). Recall that Experiment 2 was conducted to explore whether the SND effect would be modulated by other variables such as word frequency. Word frequency of the target words and other variables were not controlled in this experiment. It is, therefore, reasonable to find no SND effect *per se* in this experiment. It is also possible that the target words in Experiment 1 were selected from a certain frequency range and, thus, word frequency might have produced the pattern of fixation times observed in the previous experiment. This possibility will be handled in the Discussion section of this experiment.

Whether the effects of SND were different for the words with various values of frequency, length and number of orthographic neighbours were tested for statistical interactions. Consistent with the predictions, the interaction between the target words' SND and their word frequencies yielded significance in the early reading time measures of single fixation duration (b = 0.009, SE = 0.005, t = 2.01, p = 0.04) and gaze duration (b = 0.012; SE = 0.005, t = 2.19, p = 0.02), and this interaction was marginally significant in first fixation duration (b = 0.007, SE = 0.003, t = 1.96, p = 0.05) (see Table 4.2).

Table 4.2 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Single Fixation, First Fixation and Gaze Duration Recorded on the Target Words as Dependent Variables (Experiment 2).*

	Log Single Fixation Duration			Log First Fixation Duration			Log Gaze Duration		
	Variance	Std. Dev.		Variance	Std. Dev.		Variance	Std. Dev.	
Random effects									
Item (intercept)	0.013	0.11		0.008	0.088		0.02	0.14	
Subject (intercept)	0.01	0.12		0.01	0.10		0.017	0.13	
Residual	0.09	0.31		0.10	0.33		0.14	0.38	
AIC			2664.5			4642			6977
BIC			2690.2			4690			7032
logLik			-1328.3			-2314			-3481
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects									
Intercept	5.37	0.04	134.15	5.31	0.03	165.07	5.34	0.07	76.31
Freq. n	-0.007	0.003	-2.09*	-0.005	0.002	-2.03*	-0.009	0.004	-2.28*
ON n	-0.659	0.690	-0.95	0.144	0.533	0.27	-1.779	1.205	-1.476
Length n	-0.001	0.007	-0.23	-0.003	0.005	-0.58	0.015	0.007	1.99*
Freq. n-1	9.7e-08	4.9e-07	0.20	0.076	0.061	1.24	2.9e-07	6.05e-07	0.48
Length n-1	0.006	0.005	1.25	3.8e-07	3.9e-07	0.99	0.007	0.006	1.17
ARC	0.09	0.08	1.11	-5.8e-05	3.9e-03	-0.01	0.08	0.09	0.89
Interactions									
Freq. * ARC	0.009	0.005	2.01*	0.007	0.003	1.96*	0.012	0.005	2.19*
Length * ARC	-0.003	0.013	-0.30	-1.80	2.11	-0.85	0.028	0.016	1.76*
ON n * ARC	3.216	4.828	0.666	3.976	3.717	-1.069	11.394	6.716	1.637
Freq. n-1 * ARC	-3.3e-07	9.5e-07	-0.35	4.8e-07	7.5e-07	0.65	3.5e-08	1.16e-06	0.03
Length n-1 * ARC	0.01	0.01	1.39	0.0005	0.007	0.06	0.02	0.01	1.41

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients (1.645 ≤t<1.96); robust significant coefficients (t≥2). No significant coefficients (t<1.645).

Table 4.3 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Regression Path Duration and Total Reading Time Recorded on the Target Words as Dependent Variables (Experiment 2).*

	Log Regression Path Duration			Log Total Reading Time		
	Variance	Std. Dev.		Variance	Std. Dev.	
Random effects						
Item (intercept)	0.033	0.182		0.035	0.187	
Subject (intercept)	0.027	0.166		0.027	0.167	
Residual	0.225	0.474		0.190	0.436	
AIC			9915			8841
BIC			9943			8868
logLik			-4954			-4416
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects						
Intercept	5.66	0.028	200.46	5.63	0.03	198.21
Freq. n	-0.0002	0.0002	-0.99	-0.0003	0.0002	-1.65*
Length n	0.009	0.009	0.99	0.03	0.0009	2.68*
ON n	-5.327	1.542	-3.45*	-4.415	1.742	-2.534*
Freq. n-1	-7.67e-07	7.55e-07	-1.02	-2.75e-07	7.65e-07	-0.36
Length n-1	0.006	0.007	0.81	0.008	0.008	1.05
ARC	0.095	0.118	0.81	-0.003	0.12	-0.02
Interactions						
Freq. n * ARC	0.008	0.007	1.21	0.009	0.007	1.36
Length n * ARC	0.017	0.013	1.30	0.02	0.07	0.216
ON n * ARC	1.687	10.811	0.156	7.950	9.652	0.824
Freq. n-1 * ARC	-5.49e -06	5.23e-06	-1.03	-7.50e-06	5.30e-06	-1.41
Length n-1 * ARC	0.105	0.064	1.63	0.077	0.065	1.17

Note. No significant coefficients ($t < 1.645$); marginally significant coefficients ($1.645 \le t < 1.96$); robust significant coefficients ($t \ge 2$).

Visualisations the significant interactions obtained from the LME models are presented in Figure 4.2, 4.3, and 4.4 (each figure represents one measure of reading time). These graphical figures display the interaction of SND (*arc*) and word frequency in single fixation (*sf*), first fixation (*ff*) and gaze duration (*gd*). Before commencing on discussing the pattern of the interaction of SND and word frequency presented in these figures, it will be necessary to provide details of what information is represented in these figures. The below description will focus on Figure 4.2, but similar description can be applied to all other figures displaying the interactive effects.

For each panel (A, B, C, and D) in Figure 4.2, the little black vertical marks on the *x*-axis indicate individual frequency values of each of the target words in

Experiment 2. The *y*-axis represents single fixation duration. The black straight line is a linear curve fitted with a 95% confidence interval (the grey shaded region around the line) to the mean single fixation values across the full range of these values. The orange bars is a representative of key points in the range of the *arc* (the SND metric) value that were present in the full set of the target words. In this figure, there are four panels (A, B, C, and D). The order in which the panels should be interpreted is from Panel A to Panel C. As is clear from the diagram, the orange bar in Panel A is a representative of low values in the *arc* distribution. In Panel B, the *arc* value (the orange bar) represents values approximately one-third of the way through the *arc* distribution (see Figure 4.1). In a similar way, in Panel C (the orange bar) is a representative of the values at two-third of the way through the *arc* distribution. Finally, the *arc* value in Panel D (the orange bar) is a representative of the highest values in the *arc* distribution. The straight fitted line with a 95% confidence interval shows the effect of each individual frequency of the target word on single fixation duration at a specific point in *arc* (at the 0%, 33.334%, 66.666%, 100% values across the *arc* range) (J. Fox, personal communication, July 1, 2014). In this sense, the figure represents snapshots of the nature of the interactive effect of SND and word frequency. The narrower the 95% confidence interval, the more likely that the effect is systemattic and robust. The broader the 95% confidence interval, the less likely that that the effect is robust. The extent to which the confidence interval will be broad or narrow will be determined by the conssitency of the effect.

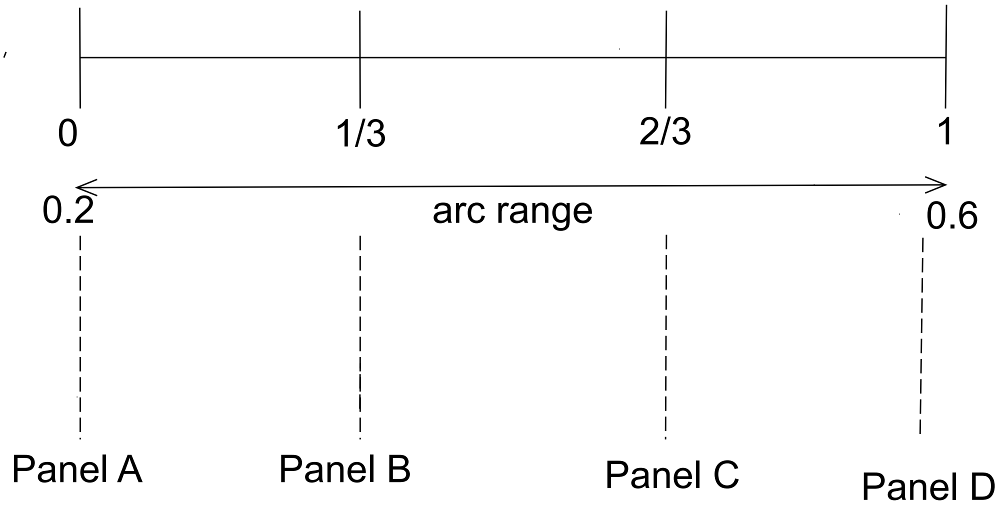

Figure 4.1 The *arc* (the SND metric) distribution in the four panels (A, B, C, and D) presented in the graphical displays of the interaction of SND and word frequency.

Turning to explain the nature of the interaction of target word frequency and SND, the pattern of the interactive effect is almost identical across the reading time measures (single fixation, first fixation and gaze duration) as can be gleaned from Figure 4.2-4.4. In all of these figures, the line in Panel A (representing low SND) is slightly upward with a broad 95% confidence interval. As we move through Panel B to Panel C (representing higher SND) and then to Panel D, it can be seen that the slope of the line becomes downward (and somewhat steeper), and the confidence interval is narrower. The trend seen in these panels can be interpreted as follows. For low SND, the frequency effect is inhibitory (i.e., reading times are shorter for low than high frequency words). As the level of SND increases, however, the frequency effect is facilitatory (i.e. reading times are shorter for high than low frequency words. Though note that the increased confidence intervals at the low SND values indicate this inhibitory effect is less systematic than that observed for higher SND levels. The influence of frequency on fixation times at high SND values is consistent with the theoretical prediction mentioned in the Introduction of this experiment.

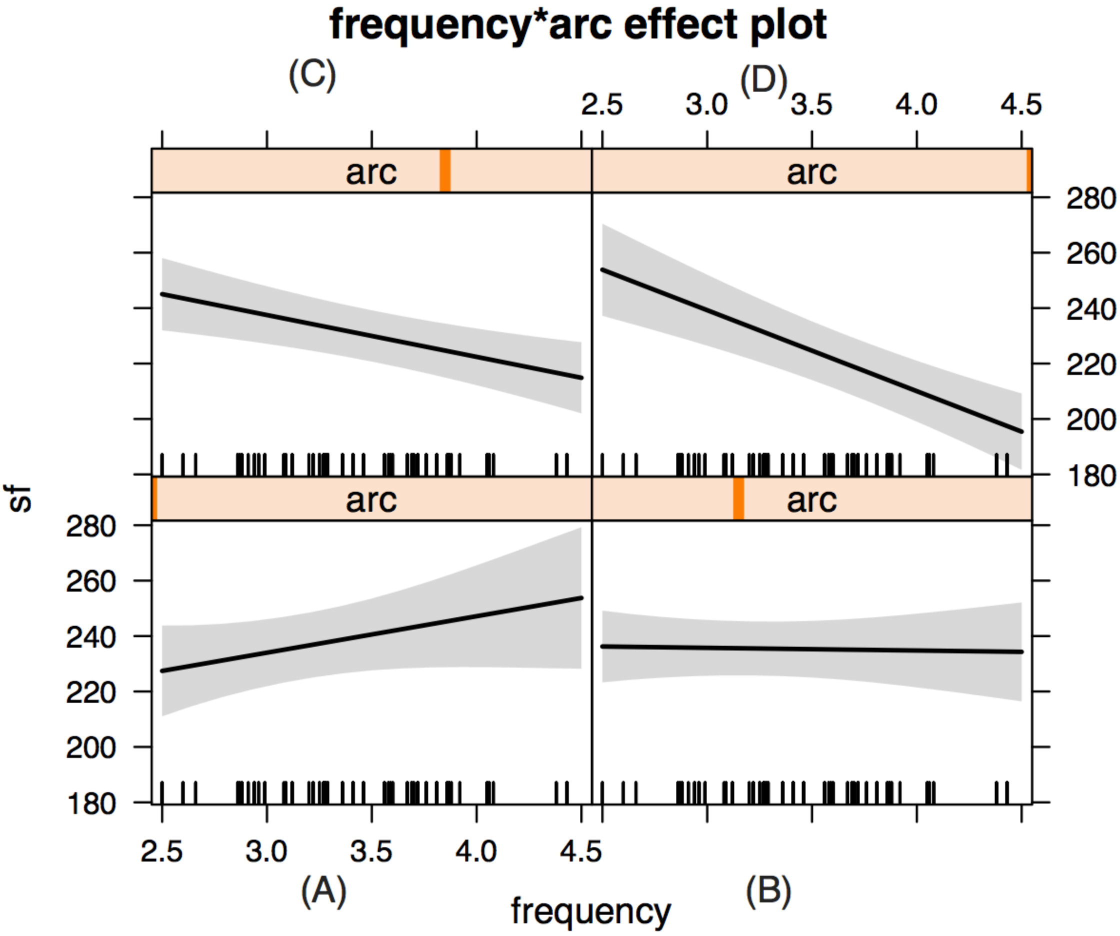

Figure 4.2 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the single fixation duration for the data of the target words (Experiment 2). The vertical axis is labelled on the single fixation duration (*sf*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

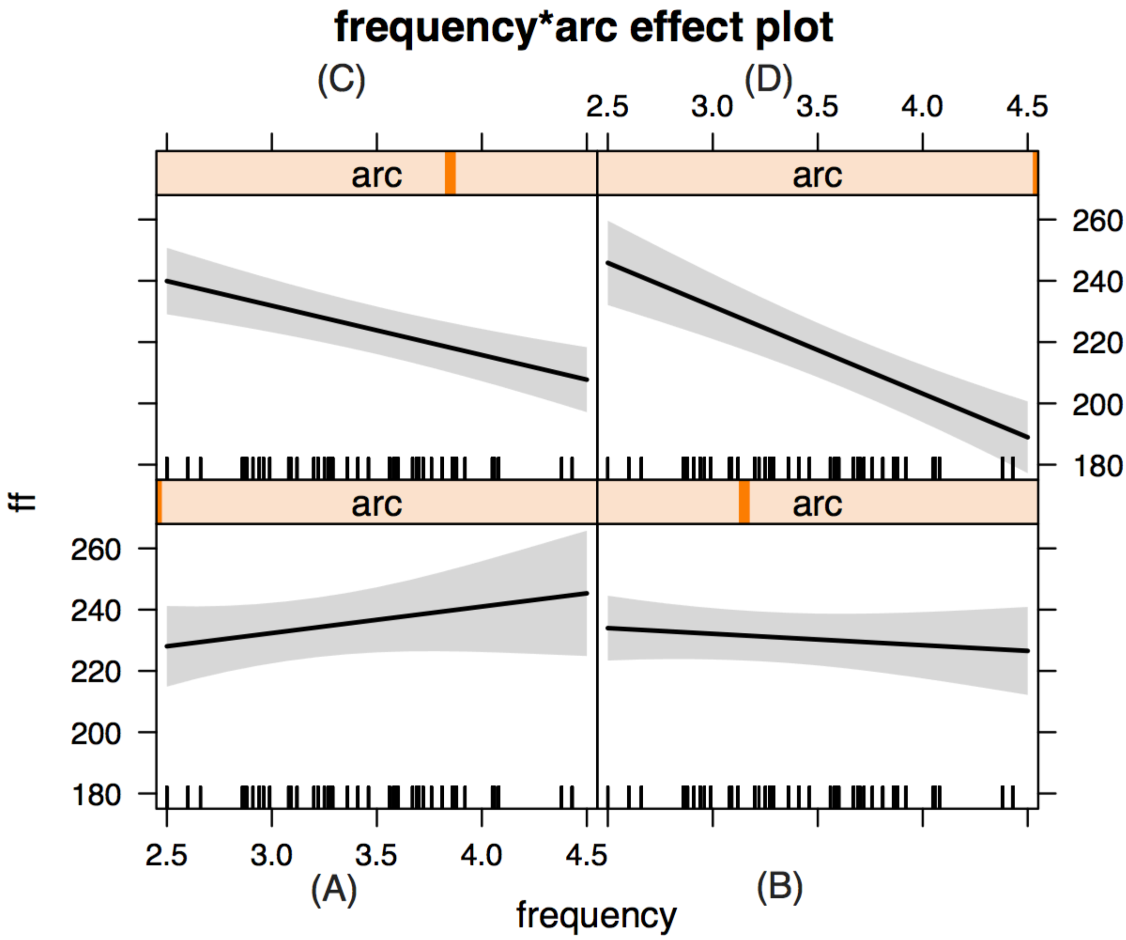

Figure 4.3 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the first fixation duration for the data of the target words (Experiment 2). The vertical axis is labelled on the first fixation duration (*ff*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

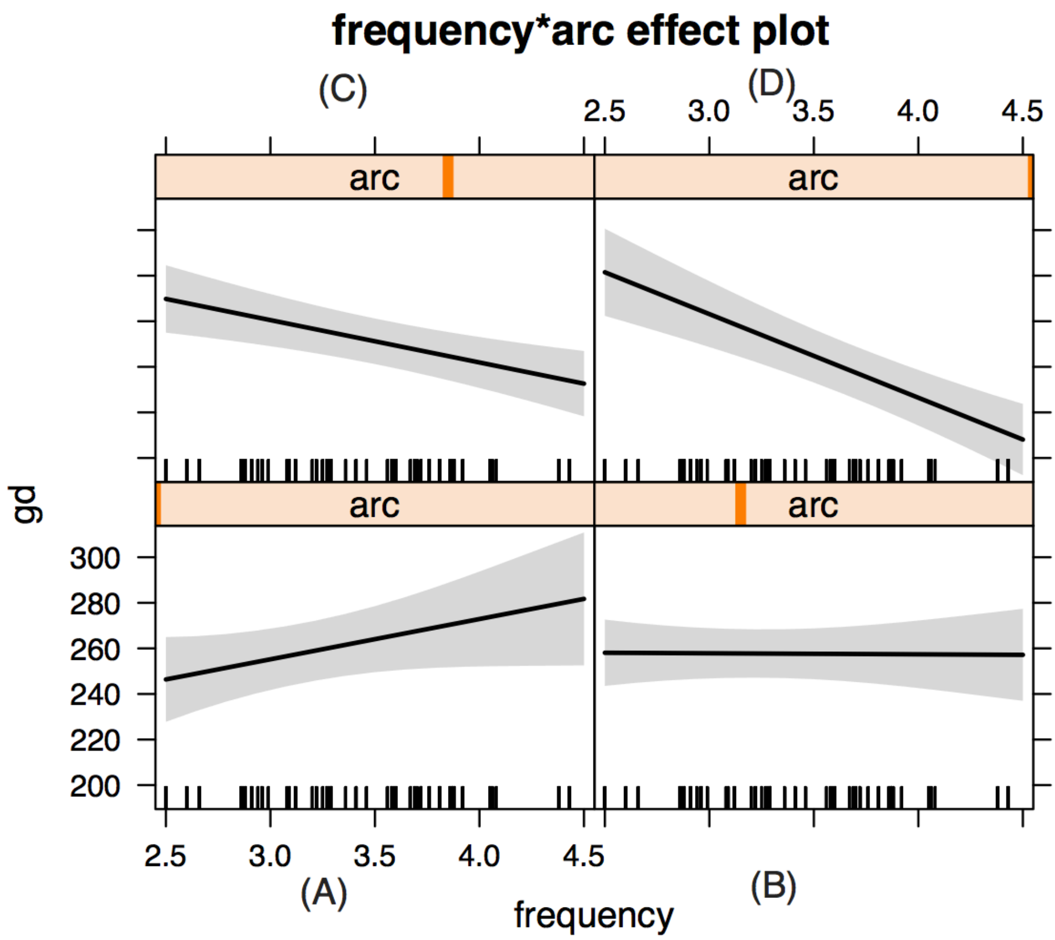

Figure 4.4 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the gaze duration for the data of the target words (Experiment 2). The vertical axis is labelled on the gaze duration (*gd*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

In addition to the interaction between frequency and SND, there was a marginally significant interaction between target words' SND and their word length as evident in gaze duration (b = 0.028, SE = 0.016, t = 1.76); the effect of SND was facilitatory for short words compared to long words (see Figure 4.6). However, the broad 95% confidence intervals (the grey shaded region around the lines, see Figure 4.5) indicate that this interaction was not robust.

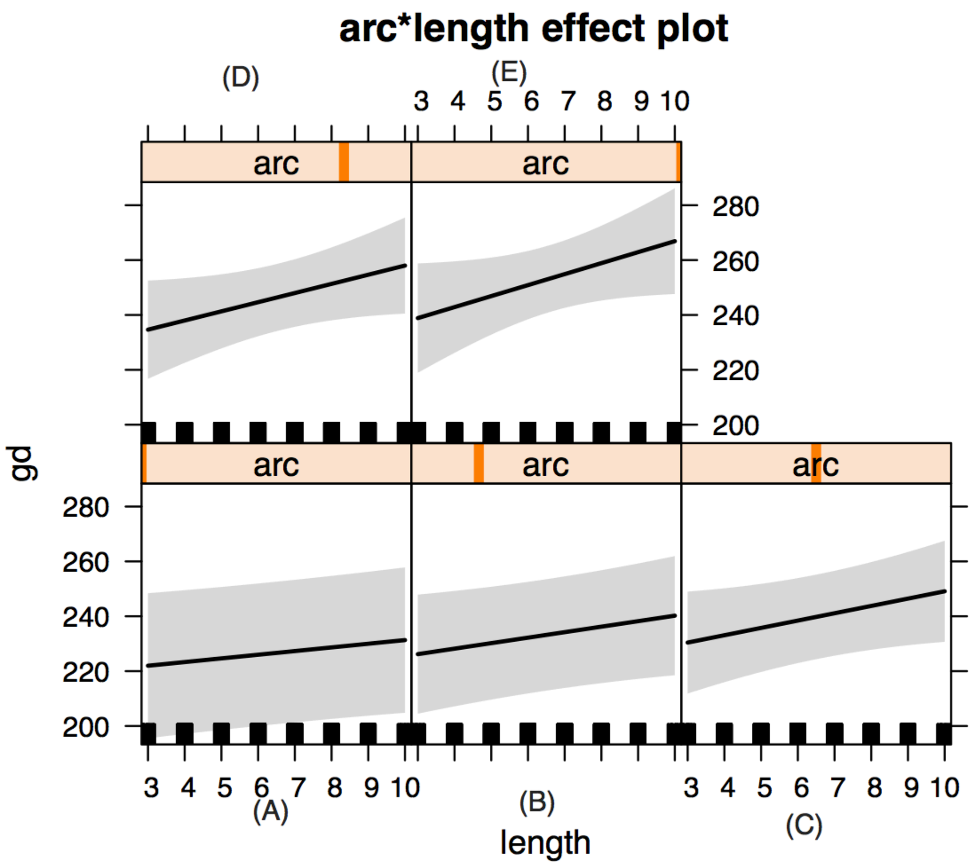

Figure 4.5 Effect display for the marginally significant interaction of target word length and SND (*arc*) in the LME model fit to the gaze duration for the data of the target words (Experiment 2). The vertical axis is labelled on the gaze duration (*gd*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

4.1.3 Discussion

Experiment 2 investigated whether target words' SND effect was modulated by other variables such as word frequency, word length and orthographic neighbourhood size, and whether the basic SND effect found in the previous experiment would be replicated. Before discussing the results pertinent to the SND effects, it is important to discuss the effects of word frequency, word length and orthographic neighbourhood size that appeared in the baseline models of LMEs used to analyse the data. In so doing, one can demonstrate that readers exhibited the well-established effects that one might expect as they read the texts. This serves to indicate that they were reading normally.

Analysing the baseline model of the target words, the results of the Experiment 2 showed that there was a significant facilitatory effect of word frequency, with high frequency words fixated for a shorter time compared to low frequency words. This finding is consistent with the eye movement literature showing that high frequency words are read faster than low frequency words (e.g., Inhoff & Rayner, 1986; Sereno & Rayner, 2000; Schilling, Rayner, & Chumbley, 1998). The effect of word length was also significant, with long words fixated for a longer time compared to shorter words. Again, this finding is in line with the findings from previous eye movement studies (e.g., Drieghe, Brysbaert, Desmet, & De Baecke, 2004). As mentioned in the previous section, word length effect was less robust in the first fixation and single fixation measures due to the fact that it is primarily driven by refixations on words, consistent with the findings of previous eye movement studies (Rayner & McConkie, 1976; Rayner, Sereno, & Raney, 1996).

The findings also showed that a significant effect of orthographic neighbourhood size appeared later in the reading time measures, as evidenced by decreased regression path duration and total reading time for words with increased orthographic neighbourhood size. Thus, the influence of orthographic neighbourhood size was a late facilitatory effect. This is inconsistent with the previous eye movement findings that demonstrated inhibitory effects of orthographic neighbourhood size (e.g., Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). Since this effect was not found in earlier reading time measures such as first fixation and gaze duration on the target words, but only in later measures (i.e., regression path duration that includes any fixations made to earlier parts of the sentences, and total reading time that includes both forward and backward movements), then the observed facilitatory effects of orthographic neighbourhood size in this experiment might have been driven by prior and post-target words that were not controlled in this experiment, and, hence, can be spurious.

Apart from the effect of orthographic neighbourhood size, these findings, thus far, demonstrate well-documented lexical effects that have been shown in previous eye movement studies. Furthermore, they serve to show that the readers were processing the passages normally as they read, and that one might reasonably argue that any SND effects that do occur in the present experiment are not the results of peculiarities associated with either the passages of text used in this experiment, or the participants who were tested.

The main effect of SND found in Experiment 1 was not replicated in the present experiment. This result was not totally unexpected since word frequency and other variables were allowed to vary in the present experiment to explore any modulations to the SND effect. One could speculate that the observed facilitatory SND effect in Experiment 1 might have been driven by a certain range of frequency of the selected target words as mentioned earlier. However, a post hoc analysis indicated that the selected target words of Experiment 1 were evenly spread in terms of (log) frequencies. The log frequency range was 1 (corresponding to a frequency count of less than 100 per million) to 9.6 (corresponding to a frequency count of about 15000 per million). In particular, there were fourteen target words of a log frequency between 1.9 and 2.94, ten words between 3 and 4.58, twelve words between 5.43 and 6.91, and eight words between 7.27 and 9.6. In addition, the correlation between SND and word frequency of the target words in Experiment 1 was weak and non-significant ($r = 0.19$, $n = 44$, $p > 0.216$). Therefore, this speculation can be ruled out from the explanations as to why the main SND effect was not observed in the present experiment. A possible, and more plausible, explanation of this finding is that the variability in word frequency and other variables may have affected the presence of the basic SND effect in this experiment.

As predicted, very robust interactive effects between target word frequency and their SND characteristics were obtained, consistent the finding of Buchanan's *et al.* (2001). However, the interaction found in the present study was to the opposite direction of the interaction found by Buchanan *et al.* (2001). In particular, the present study found that high SND benefited the processing of high frequency words more than low frequency words, while the effect of low SND was almost flat. The present findings are consistent with the predictions made earlier in the introduction of this experiment based on Stolz and Besner's (1996) embellished IA model. To reiterate, a high frequency word will have a higher baseline level of activation than a low frequency word. Thus, it will inhibit the activation of its orthographic neighbours faster than the low frequency word will. As a consequence, the high frequency word will send activation to the semantic level sooner than the low frequency word will. As such, the high frequency word will be quicker in activating its semantic representation than the low frequency word, and will benefit from high SND (via semantic feedback activation) more quickly than the low frequency word will. Words with low SND will receive weaker semantic feedback, and, thus, will not have a strong impact on word identification, compared to high SND.

It was also found that the interactions with target words' SND and target word length was marginally significant in gaze duration. In particular, increasing target word length and SND was associated with longer gaze duration. However, the visualisation of this interaction showed this interaction was not robust as indicated by the broad 95% confidence intervals.

4.1.4 Conclusion

The findings of the SND manipulation in Experiment 2 showed that the SND characteristics of the target words influenced the fixation times on the target words only via an interaction with the target word frequency in a situation where the lexical variables were allowed to vary just like in a natural reading of text found in everyday language (i.e., a non-experimental text). This interactive effect appeared early in reading time measures (single fixation, first fixation and gaze duration), giving credence to the theoretical consideration that a word's SND can impact on the ease with which a word is lexically identified. However, Experiment 2 did not orthogonally manipulate word frequency and SND; therefore, this interaction between SND and word frequency, and the direction of this interaction are still questionable.

Chapter 5: Interaction between Semantic Neighbourhood Density and Word Frequency

The findings of Experiment 2 showed that the target words' SND interacted with their frequency. In particular, it was found that high SND benefited the lexical processing of high compared to low frequency words. However, Experiment 2 passively observed how SND interacted with other lexical variables that are known to influence lexical processing. Hence, many variables including word length, orthographic neighbourhood size and sentential frames in which the target words were embedded were not controlled in the previous experiment. As such, Experiment 3 was conducted as a follow-up experiment to provide a careful validation for the interaction of SND and word frequency observed in Experiment 2. To do so, target word frequency and SND were actively manipulated to directly examine the joint effect of these two variables in word identification during normal reading.

5.1 Experiment 3

Experiment 3 examined the interactive effect of target words' SND and their frequencies in lexical identification during single sentence reading. Word frequency and SND were orthogonally manipulated while controlling for word length and orthographic neighbourhood size and holding the number of high frequency orthographic neighbours constant. The plausibility of the stimuli and the predictability of the target words were also controlled. These variables were controlled because they have been widely reported to influence lexical identification in normal reading as discussed in Chapter 1 (Rayner, 1998, 2009).

Abundant eye movement data, including the data from the previous experiment in this thesis, indicate that the speed with which a word can be identified is influenced by it frequency (Inhoff & Rayner, 1986; Rayner & Duffy, 1986; Sereno & Rayner, 2000; Schilling, Rayner, & Chumbley, 1998). High frequency words are fixated for a shorter time compared to low frequency words. The data from Experiment 2 in this thesis also showed that there was a significant interaction between target word frequency and SND in early reading time measures associated with lexical processing. Because of the nature of Experiment 2 as mentioned earlier, this interaction, however, is still questionable. Therefore, the present experiment was motivated by the findings of Experiment 2 (a corpus-based study); Experiment 3 actively manipulated SND and word frequency, rather than passively observing them, to directly investigate their joint effect.

The predictions for Experiment 3 were also derived from the theoretical account of Stolz and Besner's (1996) embellished IA mode (McClelland & Rumelhart, 1981) that was outlined in the Introduction of Experiment 2 (see Section 4.1 in Chapter 4). To briefly reiterate this account, a high frequency word unit will be more active at the word level than a word unit corresponding to a low frequency word. Therefore, the high frequency word will rapidly inhibit the activation of its orthographic neighbours relative to the low frequency word. This rapid inhibition of the orthographic competitors at the word level will allow activation to be sent quickly to the semantic level, activating the semantic representation of the word. If the high frequency word has a rich semantic representation (e.g., high SND characteristics), there will be a great amount of activation at the semantic level, which will feed back from the semantic level to word level. Thus, this enhanced semantic feedback contributes to the resolution of the competition between the orthographic neighbours at the word level. If the high frequency word has a weaker semantic representation (e.g., low SND), then the weak semantic representation will not have a comparable impact on lexical processing due to the nature of its distant semantic neighbours.

A low frequency word, on the other hand, will take longer to inhibit the activation of its orthographic neighbours because a word unit corresponding to a low frequency word is less active than the word units corresponding to a high frequency word. Because of this slower inhibition of the orthographic competitors at the word level, there will be very weak activation feeding forward from the word level to the semantic level. Concurrently, the competition between the orthographic competitors might be resolved by the activation received from the lower levels (the letter level and indirectly the visual feature level) prior to the influence of semantic feedback (recall that a low frequency word's semantic representation will be activated to a lesser extent, compared to that of a high frequency word, due to having reduced activation sent from the word level to the semantic level). As such, activating its semantic representation will have little, if any, benefit in constraining its word identification.

Based on the theoretical account given above and the findings of the previous experiment in this thesis, it was predicted that high SND would be more pronounced in the lexical processing of high than low frequency words. It was also predicted that high frequency words with high SND would be fixated for the shortest time compared to other target words. This is because a high frequency word with high SND will have two sources of strong activation (at the word level and at the semantic level), reducing the time to identify the word. Low frequency words with low SND, on the other hand, were predicted to be fixated for the longest time due to having reduced activation at both the word level and the semantic level.

Since target word frequency and SND were predicted to influence lexical processing, then these two variables and/ or their interaction should be reflected in the fixation durations on the target words themselves, and potentially, subsequent words in the text if the effect spills over (as per Rayner & Duffy, 1986; Slattery, Pollatsek, & Rayner, 2007). If SND and word frequency manipulations influence only the ease with which currently fixated words are lexically identified, then this interactive effect should be short lived, and should appear in only early fixation times on target words (e.g., single fixation, first fixation, and gaze duration). If the SND and word frequency manipulation has a longer lasting effect, and influences later stages of lexical processing and even produces effects that carry over into post-lexical processing, then this interactive effect might also appear in later reading time measures (e.g., regression path duration and total reading time) on target words. Previous eye movement studies indicate that target word frequency significantly influences the fixation times on subsequent words (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Pollatsek, Juhasz, Machacek, & Rayner, 2008; Rayner & Duffy, 1986; Slattery, Pollatsek, & Rayner, 2007). The spillover effect of SND was established in Experiment 1. Therefore, it was predicted that the effect of target word frequency and SND would spill over onto the next words.

If the claim that a word to the right of fixation can influence the durations on the currently fixated word (i.e., parafoveal-on-foveal effects) is correct, then the characteristics of the target word (e.g., frequency, SND and/ or their joint effect) would influence the fixation durations on the pre-target word, giving support to the parallel processing models of eye movement control during reading (e.g., the SWIFT model). If no parafoveal-on-foveal effects of the target word frequency, SND, or their joint effect are established, then such findings will support the serial processing models such as the E-Z Reader model.

5.1.1 Method

5.1.1.1 Participants

Forty participants took part in Experiment 3. All were students at the University of Southampton, with an age range of 18-30 years, were native English speakers, and all had normal vision or corrected-to-normal vision. None of the participants were dyslexic. Participants were awarded either course credits or given £3 for taking part. None of the participants of this experiment had taken part in the previous two experiments reported in this thesis.

5.1.1.2 Apparatus

The apparatus was the same as Experiment 1.

5.1.1.3 Material and Design

12 sets of stimuli were created; each set contained four target words that were manipulated for SND and word frequency. In total, there were 48 target words (12 high frequency words with high SND (HSND-HF), 12 high frequency words with low SND (LSND-HF), 12 low frequency words with high SND (HSND-LF), and 12 low frequency words with low SND (LSND-LF)). Table 5.1 presents the descriptive statistics of the characteristics of the target words. The target

words in each set were matched on word length and orthographic neighbourhood size (*Fs* < 1). The high and low SND target words significantly differed in SND ($F(3, 44) = 49.33$, $p = 0.0005$) and the high and low frequency words differed in word frequency (Log BNC: $F(3, 44) = 45.72$, $p = 0.0005$). Orthographic neighbourhood frequency was controlled so that the frequencies of the target words were higher than any of their respective orthographic neighbours (i.e., the frequency of the orthographic neighbours of the target words did not exceed the frequency of the target words themselves).

Table 5.1 *Means and Standard Deviations (in Parentheses) of the Characteristics of the Target Words Used in Experiment 3.*

	HSND-HF	HSND-LF	LSND-HF	LSND-LF
	M (SD)	M (SD)	M (SD)	M (SD)
SND	0.57 (0.03)	0.55 (0.02)	0.43 (0.04)	0.36 (0.07)
Log BNC	4.08 (0.35)	3.18 (0.24)	3.73 (0.15)	2.98 (0.25)
Length	6.58 (0.79)	6.58 (0.79)	6.50 (0.79)	6.56 (0.90)
ON	0.42 (0.79)	0.42 (0.79)	0.42 (0.79)	0.42 (0.79)

Note. SND: semantic neighbourhood density; BNC: word frequency from British National Corpus; length: number of letters; ON: orthographic neighbourhood size. HF > 3.50 log BNC, LF < 3.45 log BNC; HSND > 0.49 (ARC); LSND < 0.45 (ARC).

Initially, eight sentences were created for each set such that any of the four target words within a set could fit plausibly in the eight sentence frames. All of the eight sentences for each target word within a set were pre-screened for plausibility and predictability as will be described later in this section. After pre-screening the sentences, the top four sentences in each set that were given the highest plausibility and the lowest predictability were selected to be used in the main reading experiment. In total, there were 48 experimental sentences that were presented to the participants. Four lists of these sentences were created, with each list containing all 48 sentences. For each set of stimuli, the same sentence in each list differed in the target word as can be seen in

Table 5.2. In this way, each participant was presented with all of the 48 sentence frames and all of the target words in the present experiment, maximising the statistical power. A full set of the stimuli used in Experiment 3 can be found in Appendix C.

Table 5.2 A *Sample of the Sentences Containing the Experimental Manipulation in Experiment 3. The Target Words are Presented in Bold.*

	HSND-HF	HSND-LF	LSND-HF	LSN-LF
	Carpet	**Tattoo**	**Blouse**	**Napkin**
List A	She had a blue **carpet** that I liked.	Jenny pointed to the pale green **tattoo** she had just chosen.	Mary had an expensive **blouse** from that shop.	I saw an oriental **napkin** in the magazine.
List B	Jenny pointed to the pale green **carpet** she had just chosen.	Mary had an expensive **tattoo** from that shop.	I saw an oriental **blouse** in the magazine.	She had a blue **napkin** that I liked.
List C	Mary had an expensive **carpet** from that shop.	I saw an oriental **tattoo** in the magazine.	She had a blue **blouse** that I liked.	Jenny pointed to the pale green **napkin** she had just chosen.
List D	I saw an oriental **carpet** in the magazine.	She had a blue **tattoo** that I liked.	Jenny pointed to the pale green **blouse** she had just chosen.	Mary had an expensive **napkin** from that shop.

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HF: high frequency; LF: low frequency

Before conducting Experiment 3, all the sentences were pre-tested for plausibility and predictability using pen and paper questionnaires. In the plausibility ratings, the participants were asked to rate how likely it was that the event in the given sentences would occur. These ratings were made on a 7-point Likert scale (1 = *very implausible*, 7 = *very plausible*) by four participant groups (twelve participants in all of the four groups: three participants assigned to each list). The results showed that the sentences in the four lists were plausible (list A *mean* = 5.27; list B *mean* = 5.38; list C *mean* = 5.29; list D *mean* = 5.33), and the one-way ANOVA indicated that the four lists were not statistically different from each other in terms of plausibility ($F < 1$). A fifth group of twelve participants completed a predictability cloze test in which they saw the beginning of the sentences up to the word preceding the target words and were asked to complete the sentences with the most obvious word that came to mind. The result of the cloze test showed that none of the participants predicted the target words (the total number of predicted target words = 0).

5.1.1.4 Procedure

The procedure followed in this experiment was the same as Experiment 1.

5.1.2 Results

Prior to data analysis, data trimming was carried out along with removing outliers following the same criteria as in Experiment 1, and this resulted in removal of 1.79% of the data prior to the analyses. The dependent variables were single fixation duration, first fixation duration, gaze duration, regression path duration, total reading time and skipping probability (see Section 3.1.1.5 in Chapter 3 for a description of the measures). A normal Quantile-Quantile plot (Wilk & Gnanadesikan, 1968) was obtained to check whether the fixation durations (the dependent variables of this experiment) were normally distributed. The plot indicated that the fixation durations were not normally distributed. Therefore, the fixation durations were log-transformed to approximate a normal distribution.

As in Experiment 1 and 2, all measures were analysed with linear mixed effect (LME) models. Participants and items were entered as random effects and the target words' frequency, word length and orthographic neighbourhood size followed by the SND metric (ARC, Shaoul & Westbury, 2010b) were entered in the models as fixed effects, one variable at a time. These variables were entered in the models before the interaction terms in order to examine the unique contribution of the interaction of SND and word frequency in lexical processing. Finally, interaction terms (frequency * ARC, and length * ARC, orthographic neighbourhood size * ARC) were subsequently added to the resulting models also as fixed effects. All the fixed variables including the SND variable were entered as continuous variables, and were all centred at the means to minimise collinearity in the analysis of the data. To make interpretation of the data easier, word frequency and SND were dichotomised using a median split (HF > 3.50 log; LF < 3.45; HSND > 0.49; LSND < 0.45) when presenting and discussing the findings of the LME models (note, though, as specified earlier, the frequency and SND were entered in the models as continuous variables). The following statistics will be reported in the results of this experiment: the regression coefficients (*b*), standard errors (*SE*), *t* values (or the *Z* value in the case of the skipping probability) together with *p* values based on Markov Chain Monte Carlo sampling 10,000 samples (Baayen *et al.*, 2008).

As in Experiment 1, all sentences were divided into four regions as shown in the examples below.

REGION	1	2	3	4
	She put the/	black/	**knife**/	on the table/.

The particular regions of interest for the analyses of this experiment were Region 2 constituting a pre-target word, Region 3 constituting a target word, and Region 4 constituting post-target words. The criterion for determining the post-target region was identical to the criterion used in Experiment 1.

The overall mean comprehension rate was 96.83% indicating that the participants read and understood the sentences. The reported results will include: 1) the interactive effect of SND and word frequency on the fixation times on the pre-target words (i.e., parafoveal-on-foveal effects), 2) the interaction of SND word frequency on fixation times on the target words, and 3), any interactive effects of SND and word frequency (i.e., spillover effects) on the fixation times on the post-target words.

To investigate whether the characteristics of the target words influenced the fixation durations on the pre-target words, LME analyses were conducted. The reading time measures for the pre-target words were the dependent variables;

participants and items were entered as random effects, and the predictors of target word frequency and SND along with the joint effect of these two variables were entered as fixed effects. There were no reliable effects for these analyses (all *t*s < 0.9) providing no evidence of any parafoveal- on-foveal effects (see Table 5.3 for the mean values associated with these analyses).

Table 5.3 *Means of Fixation Times on the Pre-Target Words Preceding Target Words with the Experimental Manipulation in Experiment 3. Standard Deviations are Given in Parentheses.*

	HSND-HF	HSND-LF	LSND-HF	LSND-LF
Single Fixation Duration	213 (64)	209 (69)	210 (62)	205 (54)
First Fixation Duration	216 (66)	209 (70)	208 (61)	208 (63)
Gaze Duration	232 (93)	233 (105)	234 (92)	235 (106)
Regression Path Duration	305 (185)	304 (202)	294 (188)	392 (175)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HF: high frequency; LF: low frequency.

Next, the interaction of SND and word frequency on the fixation times on the target words was examined. Table 5.4 and 5.5 list the results of the LME analyses carried out on the fixation times on the target words. As can be seen from these two tables, in the baseline models, the effect of word frequency was significant in all early and late reading time measures (single fixation duration: $b = -12.545$, $SE = 4.765$, $t = -2.633$; first fixation duration: $b = -13.443$, $SE = 3.882$, $t = -3.463$, gaze duration: $b = -14.707$, $SE = 5.583$, $t = -2.63$; regression path duration: $b = -21.307$, $SE = 6.241$, $t = -3.414$; total reading time: $b = -16.117$, $SE = 7.907$, $t = -2.038$). As the negative *b* coefficients indicate, high frequency words were read faster than low frequency words. Recall that the length of the target words within a stimulus set was matched in this experiment such that word length was on average similar across the target words in the four conditions. However, since word length varied between stimulus sets (in line with the Latin Square design; e.g., one of the sets may have contained a four-letter target word and another a six-letter target word), this variable was included in the LME models to capture extraneous variance in the data. These data are reported here (and elsewhere in the thesis) for completeness. Word length significantly predicted the fixation times on the target words in all reading time measures (single fixation duration: $b = 6.567$, $SE = 2.689$, $t = 2.442$; first fixation duration: $b = 7.409$, $SE = 2.123$, $t = 3.491$; gaze duration: $b = 11.736$, $SE = 3.058$, $t = 3.838$; regression path duration: $b = 11.572$, $SE = 3.436$, $t = 3.368$; total reading time: $b = 17.294$, $SE = 4.355$, $t = 3.971$). As can be seen from the *b* coefficients, long words were fixated for longer than shorter words. This significant word length effect found in this experiment reflects the variation in word length between stimuli sets (and is therefore of little theoretical interest). The orthographic neighbourhood size was not a significant predictor in any of the early or late reading time measures (all *t*s < 1.09). Given that orthographic neighbourhood effects were controlled across conditions similar to word length effects, it may initially appear surprising that no reliable effects were obtained. However, it is likely that the lack of such effects is due to the fact that orthographic neighbourhood size is a far less influential factor on eye movements during reading than is word length.

Examination of the SND effects indicated that the effects of the target words' SND were significant in all early and late reading time measures (all *t*s > 2) (single fixation duration: $b = -489.44$, $SE = 121.78$, $t = -4.019$; first fixation duration: $b = -368.32$, $SE = 97.07$, $t = -3.79$; gaze duration: $b = -570.39$, $SE = 140.10$, $t = -4.071$; regression path duration: $b = -913.84$, $SE = 160.26$; $t = -5.70$; total reading time: $b = -133.03$, $SE = 203.27$; $t = -6.55$). As the negative coefficients indicate, increased SND resulted in decreased fixation times on the target words. This pattern of the SND effect was predicted, and provides further evidence consistent with the facilitatory SND effect reported in Experiment 1. SND, however, did not predict the skipping probability of the target words ($b = -1.395$, $SE = 1.033$, $Z = -1.35$, $p = 0.17$) as well as SND did not predict the refixation probability (($b = -1.25$, $SE = 0.726$, $Z = -1.72$, $p = 0.08$).

Importantly, an interactive effect of target word frequency and SND emerged in all early reading time measures (single fixation duration: $b = -155.65$, $SE = 36.08$, $t = -4.31$, $p < 0.00005$; first fixation duration: $b = -123.58$, $SE = 28.90$, $t = -4.27$, $p < 0.00005$; gaze duration: $b = -181.31$, $SE = 41.56$, $t = -4.36$, $p <$

0.00005) and the late measures (regression path duration: $b = -269.42$, $SE = 47.53$, $t = -5.66$, $p < 0.00005$; total reading time: $b = -402.64$, $SE = 60.28$, $t = -6.67$, $p < 0.00005$). The interaction of SND and word frequency, however, did not predict the skipping probability of the target words ($b = -2.16$, $SE = 1.642$, $Z = -1.31$, $p = 0.18$). Recall that the decision to skip a word must be made early in processing (when the word is in the parafovea) (Rayner, White, Kambe, Miller, & Liversedge, 2003). As such, it is likely that information about the target word's SND was not obtained parafoveally (before fixating the word), consistent with the findings of parafoveal-on-foveal effects reported earlier in this section.

Table 5.4 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Single Fixation, First Fixation and Gaze Duration Recorded on the Target Words as Dependent Variables (Experiment 3).*

	Log Single Fixation Duration			Log First Fixation Duration			Log Gaze Duration		
	Variance	Std. Dev.		Variance	Std. Dev.		Variance	Std. Dev.	
Random effects									
Item (intercept)	426.08	20.642		375.24	19.371		747.55	27.341	
Subject (intercept)	680.08	26.078		503.32	22.435		760.84	27.583	
Residual	2654.91	51.526		2945.53	54.273		5681.57	75.376	
AIC			10875			17928			19286
BIC			10895			17950			19308
logLik			-5434			-8960			-9639
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects									
Intercept	247.899	17.098	16.078	272.262	13.970	19.489	304.243	19.933	15.263
Freq. n	-12.545	4.765	-2.633*	-13.443	3.882	-3.463*	-14.707	5.583	-2.634*
Length n	6.567	2.689	2.442*	7.409	2.123	3.491*	11.736	3.058	3.838*
ON n	-1.736	2.831	-0.613	-2.464	2.301	-1.071	-1.871	3.272	-0.572
ARC	-489.447	121.784	-4.019*	-368.324	97.070	-3.794*	570.395	140.109	-4.071*
Interactions									
Freq. * ARC	-155.65	36.083	-4.315*	-123.587	28.903	- 4.276*	-181.319	41.565	- 4.362*
Length * ARC	9.761	5.311	1.838*	-35.338	23.586	-1.498	7.714	34.007	0.227
ON * ARC	6.409	34.211	0.187	27.579	28.949	0.953	44.874	39.695	1.130

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients (1.645 ≤t<1.96); robust significant coefficients (t≥2). No significant coefficients (t<1.645);

Table 5.5 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Regression Path Duration and Total Reading Time Recorded on the Target Words as Dependent Variables (Experiment 3).*

	Log Regression Path Duration			Log Total Reading Time		
	Variance	Std. Dev.		Variance	Std. Dev.	
Random effects						
Item (intercept)	723.37	26.896		1752.9	41.868	
Subject (intercept)	886.13	29.768		1785.3	42.252	
Residual	9205.89	95.947		10147.0	100.733	
AIC			19045			18657
BIC			19066			18678
logLik			-9518			-9324
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects						
Intercept	344.234	22.217	15.494	353.011	28.302	12.473
Freq. n	-21.307	6.241	-3.414*	-16.117	7.907	-2.038*
Length n	11.572	3.436	3.368*	17.294	4.355	3.971*
ON n	1.191	3.766	0.316	5.309	4.889	1.086
ARC	-913.845	160.268	-5.702*	1333.03	-203.27	-6.558*
Interactions						
Freq. * ARC	-269.429	47.533	-5.668*	-402.641	60.285	-6.679*
Length * ARC	4.516	6.173	0.732	-19.853	48.794	-0.407
ON* ARC	65.586	44.549	1.472	87.753	56.501	1.553

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients ($1.645 \leq t < 1.96$); robust significant coefficients ($t \geq 2$). No significant coefficients ($t < 1.645$).

To gain a general insight into the nature of the interactive effect of SND and word frequency, the pattern of effects observed in the mean of reading times for each measure was considered (see Table 5.6). Recall that it was predicted that high SND would benefit the lexical identification of high frequency words to a greater extent than low frequency words. As explained earlier, a high frequency word will activate its semantic representation sooner than a low frequency word will due to having a higher baseline level of activation at the word level. As a consequence, the high frequency word will benefit from enhanced semantic feedback activation (provided by high SND characteristics) more quickly than a low frequency word will. As can be seen from Table 5.6, even though high frequency words with high SND were fixated for the shortest time compared to the other conditions, and the low frequency words with low

SND were fixated for the longest time relative to the reading times in the other conditions (both findings consistent with the predictions), the nature of the interactive effects changed from the early to the late measures. For single fixation duration and first fixation duration, there was a larger SND effect for the high than the low frequency words (HF: 19ms, 11ms; LF: 12ms, 7ms, respectively). In contrast, for gaze duration, and regression path duration and total reading time measures, which are all slightly later measures of processing, the SND effect was larger for the low than the high frequency words (LF: 22ms, 23ms, 31ms; HF: 20ms, 13ms, 13ms, respectively).

Table 5.6 *Means of Fixation Times on the Target Words in Experiment 3. Standard Deviations are Given in Parentheses.*

	HSND-HF	HSND-LF	LSND-HF	LSND-LF
Single Fixation Duration	217 (31)	238 (45)	236 (39)	250 (58)
First Fixation Duration	217 (28)	234 (37)	228 (33)	241 (42)
Gaze Duration	242 (33)	264 (49)	262 (43)	286 (62)
Regression Path Duration	291 (85)	303 (72)	304 (65)	326 (85)
Total Reading Time	325 (97)	324 (86)	338 (86)	355 (97)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HF: high frequency; LF: low frequency.

Visualisations of the significant interactions obtained from the LME models are presented in Figure 5.1, 5.2, 5.3, 5.4, and 5.5 (each figure represents one measure of a reading time). These graphical figures display the interaction of SND (*arc*) and word frequency in single fixation (*sf*), first fixation (*ff*), gaze duration (*gd*), regression path duration (*goPast2*) and total reading time (*TOT*). The details of what information is represented in these figures are provided in Section 4.1.2 in Chapter 4 (pp. 126-128). To reiterate, the order in which the panels should be interpreted is from Panel A to Panel D, and the orange bar in each panel represents a value of *arc* (the SND metric) at which the effect of word frequency is plotted. In Panel A in Figure 5.1, the line is almost flat, if anything, slightly upward, with a broad confidence interval. As we move through Panel B to Panel C and then to Panel D, it can be seen that the slope of the line becomes downward (and somewhat steeper), and the confidence interval is reduced. Panel A can be interpreted as showing that for reduced SND, the influence of frequency is inhibitory, that is, reading times are shorter for low than high frequency words at this reduced level of SND. However, as the level of SND increases (through Panels B-D) this trend shifts such that it lies in the opposite direction. That is to say, for higher SND, the influence of frequency is facilitatory, with reduced reading times for high than low frequency words. The latter influence is consistent with the theoretical prediction outlined in the Introduction. Broadly, this pattern is comparable in the figures for each of the other reading time measures; however, there is a gradual shift in the pattern as we consider increasingly late measures. The shift is such that in Panel A of the figures (e.g., for regression path duration and total reading time), the slope of the line has an increasingly upward projection, indicating that for the later measures, for reduced SND, the inhibitory influence of frequency is more pronounced than is the case for the earlier reading time measures. This is a trend that is in the opposite direction to the effects that were predicted, though note that the increased confidence intervals here indicate that this effect is less systematic than that observed for higher SND values. Note also, however, that the facilitatory influence of frequency on fixation times that exists for higher levels of SND is present in both the early and the later reading time measures.

At this point, it should be clear that the joint influence of word frequency and SND on the different reading time measures is complex and changes quite systematically from the early to late measures. To summarise, there are two important points to note. First, from the means in Table 5.6 it is clear that there is a shift in the nature of the interactive effect such that the SND effects are larger for high than low frequency words in the early measures (first and single fixation duration), but are larger for the low than the high frequency words in the somewhat later measures (gaze duration, regression path duration and total reading time). Second, the visualisations of the data in Figures 5.1-5.5 show that this shift occurs due to counteractive influences of frequency at different levels of SND. For low values of SND, effects of frequency are inhibitory, whereas, for higher values of SND, effects of frequency are facilitatory. Thus, aspects of the data offer very clear support for the primary hypothesis that a positive relationship should exist between SND and frequency. However, it is also the case that there are aspects of the data that are inconsistent with this suggestion, specifically, at low levels of SND where the influence of frequency is inhibitory. Thus, overall, the modulatory influence of frequency at different levels of SND changes across the time course of processing of a word during reading.

Finally, for completeness, all other interactions failed to reach significance, except a marginally significant interaction of target word length and SND that only occurred in the single fixation duration measure (b = 9.761, SE = 5.311, t = 1.838).

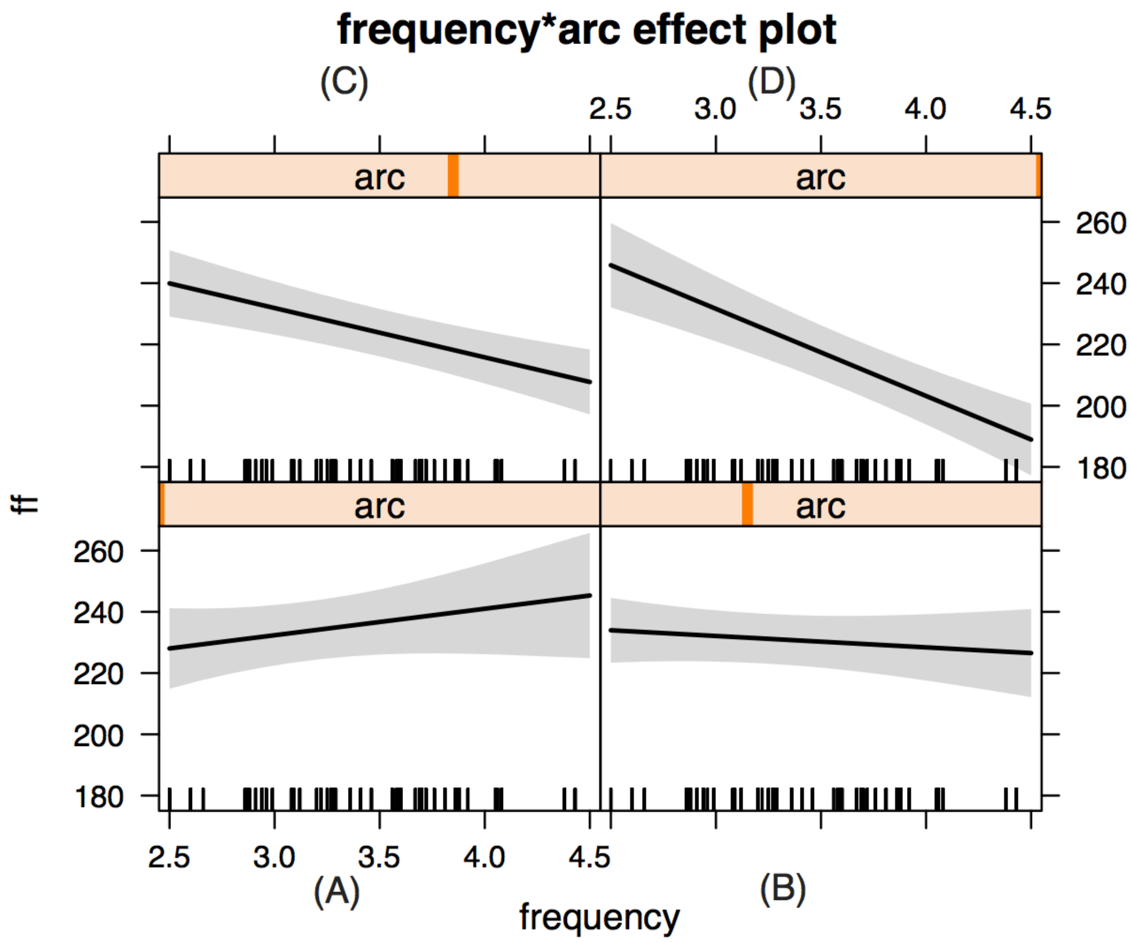

Figure 5.1 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the first fixation duration for the data of the target words (Experiment 3). The vertical axis is labelled on the first fixation duration (*ff*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

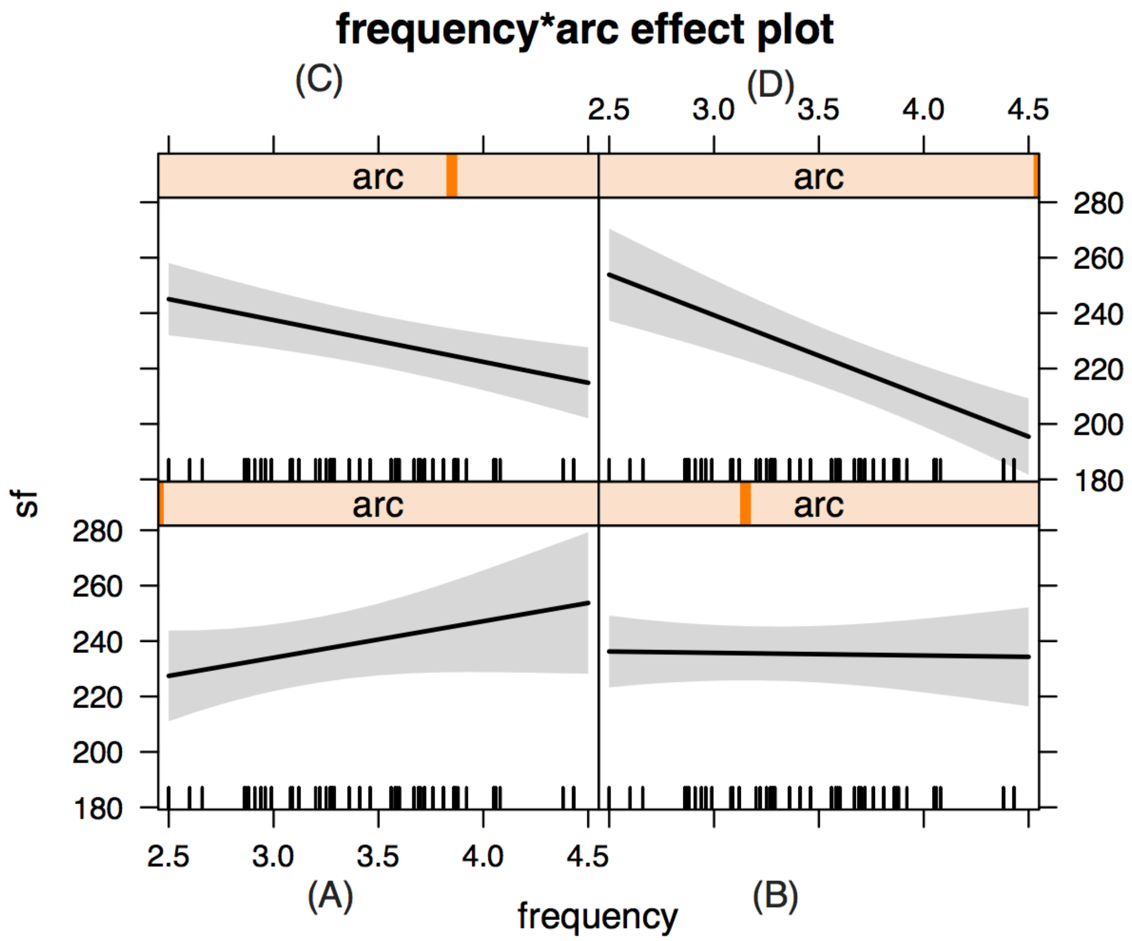

Figure 5.2 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the single fixation duration for data of the target words (Experiment 3). The vertical axis is labelled on the single fixation duration (*sf*) on the target word, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

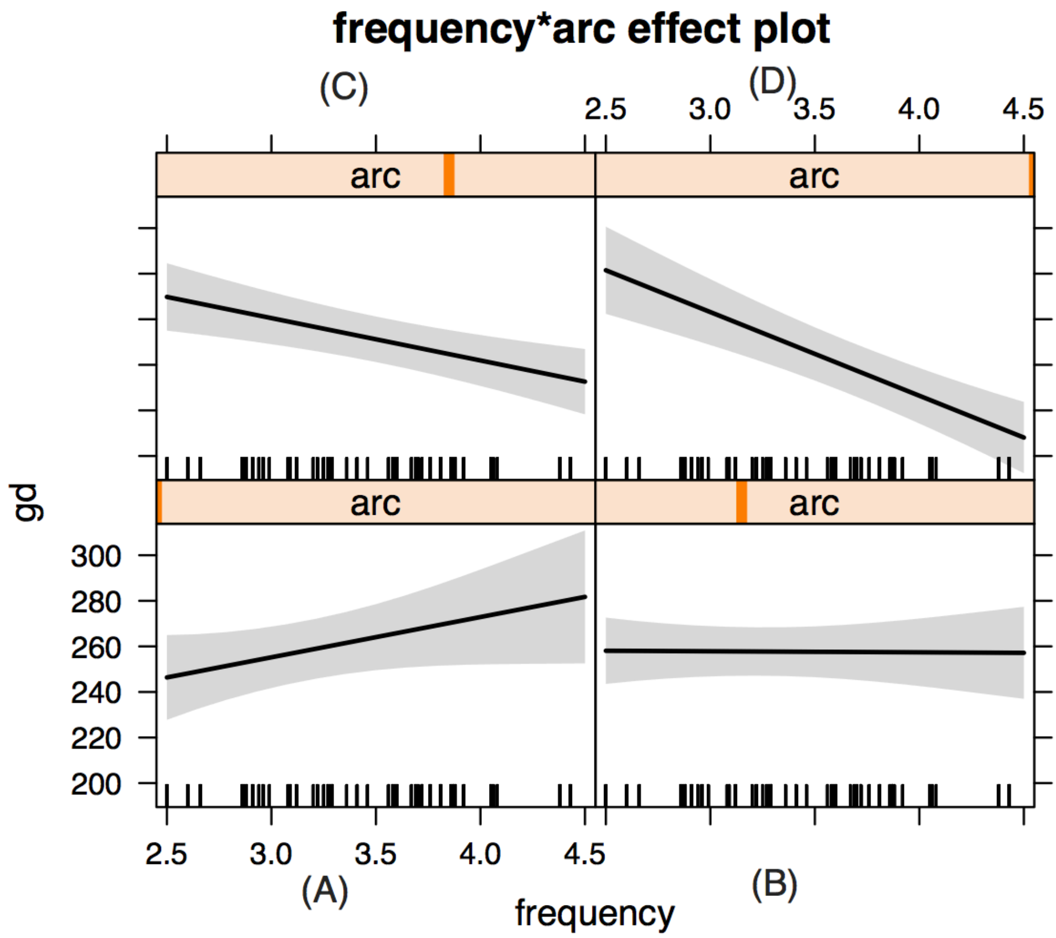

Figure 5.3 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the gaze duration for the data of the target words (Experiment 3). The vertical axis is labelled on the gaze duration (*gd*) on the target word, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

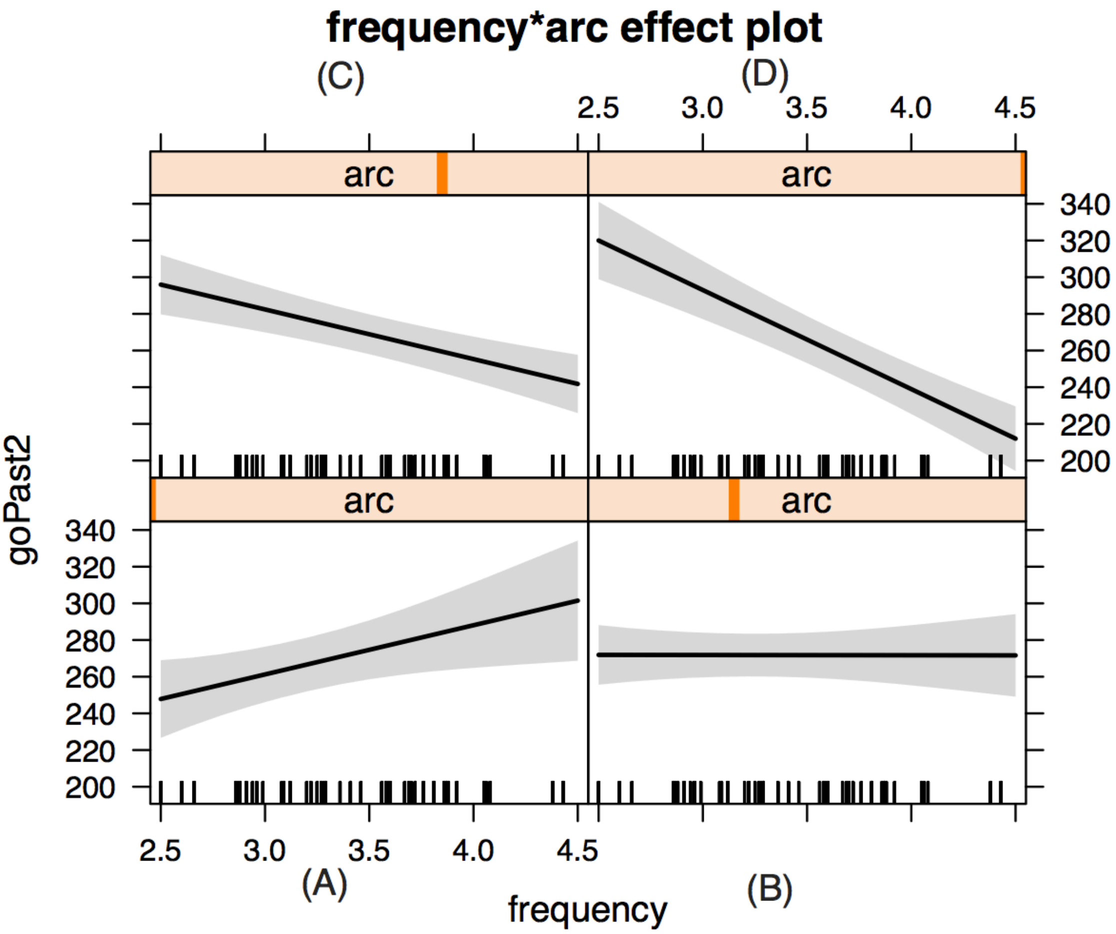

Figure 5.4 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the regression path duration for the data of the target words (Experiment 3). The vertical axis is labelled on the regression path duration (*goPast2*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

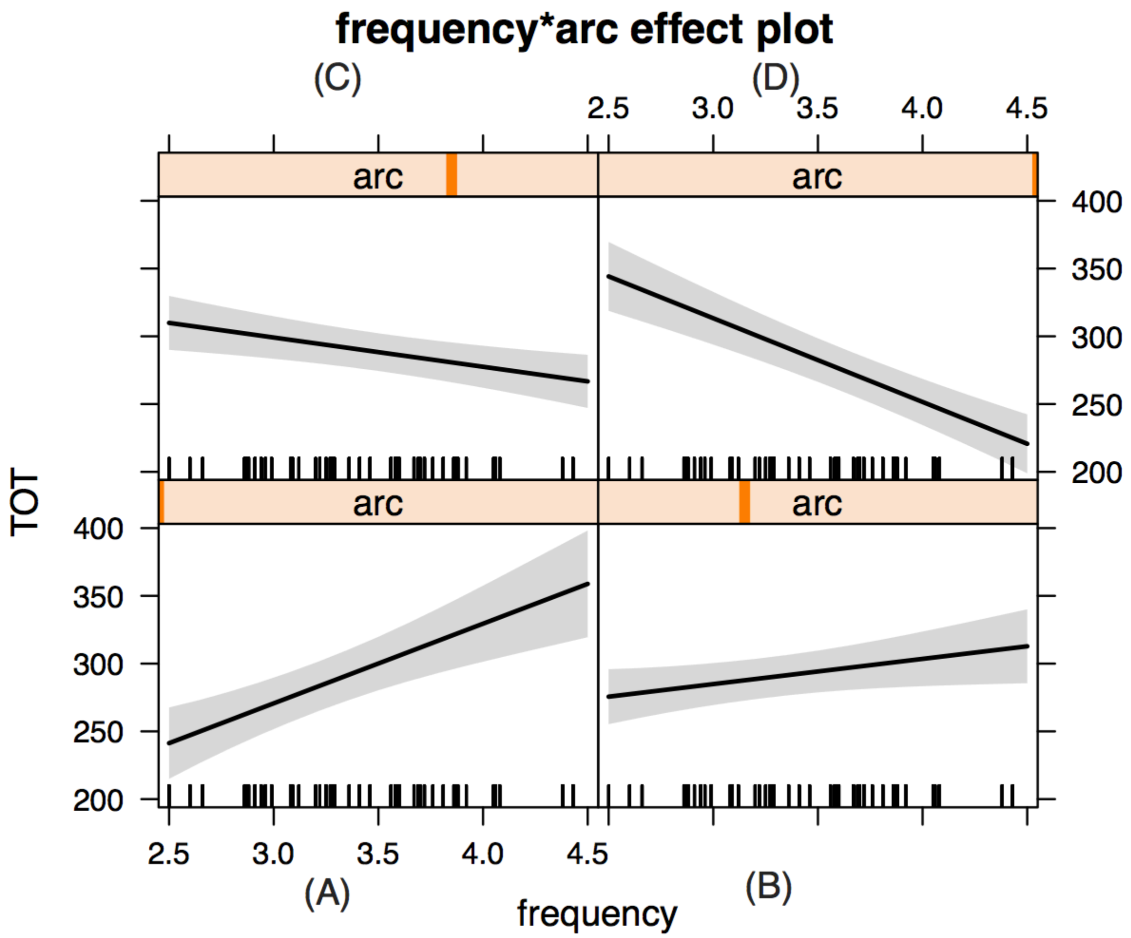

Figure 5.5 Effect display for the significant interaction of target word frequency and SND (*arc*) in the LME model fit to the total reading time for the data of the target words (Experiment 3). The vertical axis is labelled on the total reading time (*TOT*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

SND and Word Frequency

The last aspect of the data that was considered concerned whether effects due to the characteristics of the target words spilled over onto subsequent words. Again, LME models were conducted. The reading times for the post-target words were dependent variables, participants and items were entered as random effects, and the predictors of target word frequency, SND, and the interaction of these variables were entered as fixed effects. The results showed that each of word frequency and SND significantly predicted first fixation duration on the post-target words (word frequency: $b = -14.478$, $SE = 6.944$, $t = -2.085$; SND: $b = -77.89$, $SE = 34.14$, $t = -2.282$), while their effect was marginal in gaze duration (frequency: $b = -17.52$, $SE = 10.74$, $t = -1.632$; SND: $b = -91.44$, $SE = 52.27$, $t = -1.749$). The interactive effect of target word frequency and SND was significant showing spillover effects of the post-target region in the early reading time measures of first fixation duration: $b = -164.47$, $SE = 55.40$, $t = -2.969$; gaze duration: $b = -291.90$, $SE = 84.82$, $t = -3.441$). The interactive effects of target word frequency and SND were not significant in all other measures (all *t*s < 1.50).

The pattern of effects observed in the mean of reading times of the post-target region is presented in Table 5.7. The discussion here will focus on the reading measures of first fixation and gaze duration that showed a significant spillover effect of the interactive effects of target word frequency and SND. As can be seen from Table 5.7, the post-target regions following high frequency words with high SND were fixated for the shortest time compared to other condition, and the post-target regions following low frequency words with low SND were fixated for the longest time relative to the reading times in other conditions. For first fixation and gaze duration, the SND effect was comparable for the high and the low frequency words in the spillover regions (HF: 11ms, 14ms, LF: 12ms, 10ms respectively).

Table 5.7 *Means of Fixation Times on the Post-Target Regions Preceding the Target Words with the Experimental Manipulation in Experiment 3. Standard Deviations are Given in Parentheses.*

	HSND-HF	HSND-LF	LSND-HF	LSND-LF
Single Fixation Duration	261 (124)	272 (132)	276 (149)	279 (109)
First Fixation Duration	246 (112)	252 (120)	257 (121)	264 (137)
Gaze Duration	287 (161)	293 (172)	301 (166)	303 (173)
Regression Path Duration	370 (295)	377 (313)	374 (272)	380 (330)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HF: high frequency; LF: low frequency.

Visualisations of the significant interactive effect of word frequency and SND on the fixation times of the post-target region (obtained from the LME models) are presented in Figure 5.6, and 5.7 (each figure represents one measure of reading time). Broadly, the interactive effects observed on the post-target region as presented in these figures are comparable to those observed on the target words as presented in Figure 5.1-5.5. For the low SND (see Panels A and B in Figure 5.6-5.7), the effects of target word frequency on the fixation times of the post-target region were inhibitory, while the effects of word frequency was facilitatory for high SND (see Panels C and D in both figures). The visualisations of the data in Figure 5.6-5.7 show that the comparable size of SND effect observed in first fixation and gaze duration is driven by counteractive influence of word frequency at different levels of SND, rather than being driven by the high vs. low SND alone.

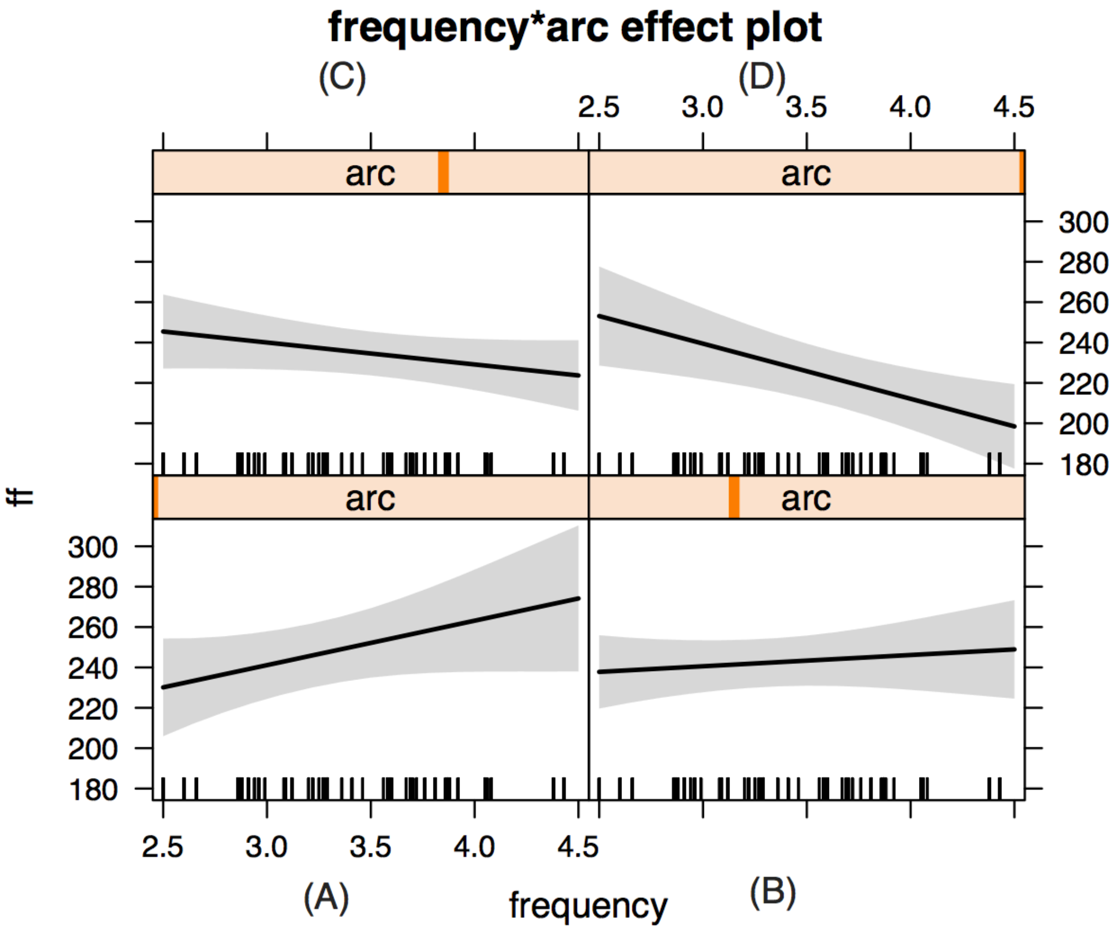

Figure 5.6 Effect display for the effect of frequency by SND (*arc*) in the LME model fit to the first fixation duration for the spillover region data (Experiment 3). The vertical axis is labelled on the first fixation duration (*ff*) on the post-target region, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

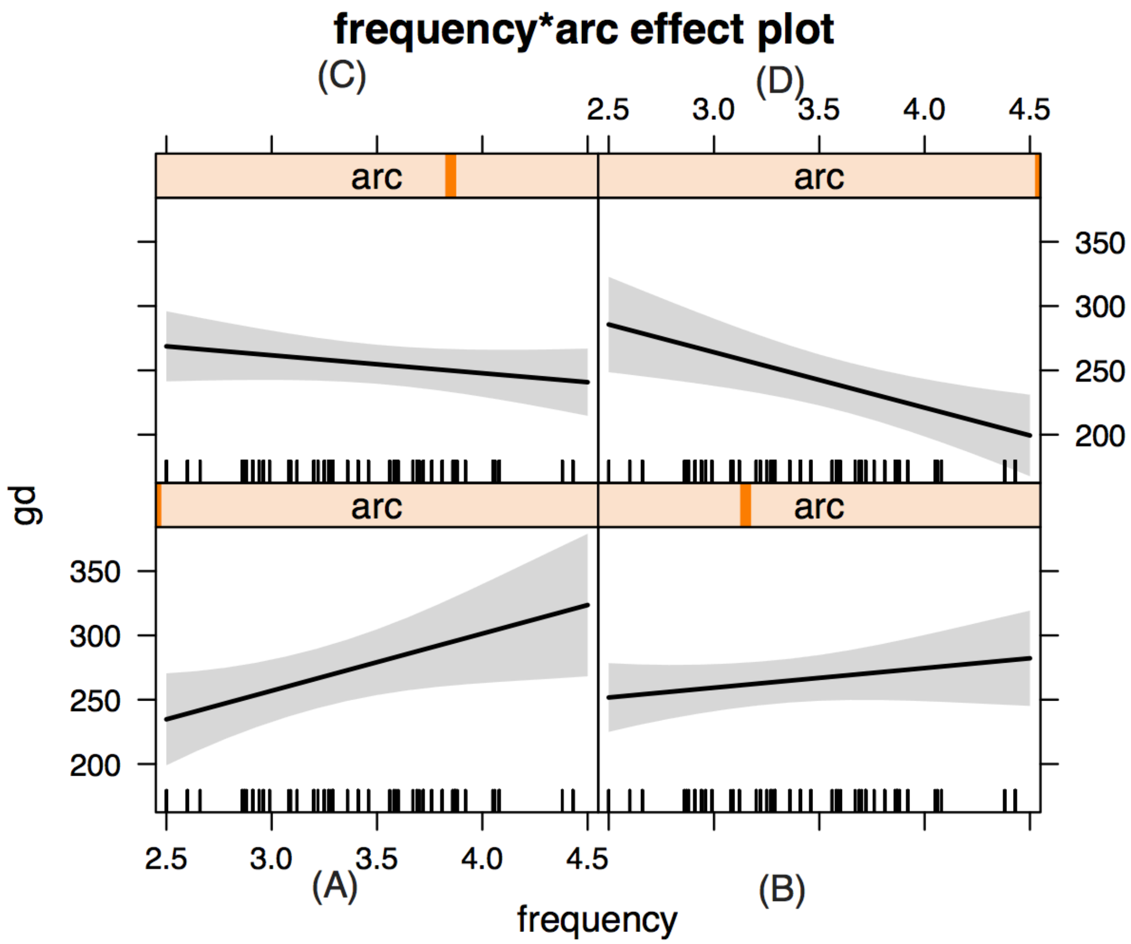

Figure 5.7 Effect display for the significant effect of frequency by SND (*arc*) in the LME model fit to the gaze duration for the spillover region (Experiment 3). The vertical axis is labelled on the gaze duration (*gd*) on the post-target region, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

5.1.3 Discussion

Experiment 3 investigated 1) whether the interaction between target word frequency and SND would influence the fixation times on prior words (i.e., parafoveal-on-foveal effects), 2) whether the interactive effect of the target words' SND and their frequencies would influence the fixation times on target words, and 3) whether this interactive effect would influence the fixation times on the subsequent words (i.e., spillover effects).

5.1.3.1 Parafoveal-on-Foveal Effects of SND and Word Frequency

This part of analysis examined whether the characteristics of a yet-to-be fixated word (a target word in the case of the present analysis) influenced the fixation durations of the fixated word (i.e., a pre-target word). It was found that the fixation durations on a pre-target word were not influenced by the characteristics of the yet-to-be fixated (adjacent) word. In particular, word frequency, SND, and the interaction of SND and word frequency associated with the adjacent word did not predict the fixation times on the pre-target word. This result was consistent with the findings of Experiment 1 in relation to the null parafoveal-on-foveal effects. Taken together, the results of Experiment 1 and 3 provide no evidence for the claim that the processing of a currently fixated word is affected by the lexical or semantic characteristics of the parafoveal word, which is in line with the findings of many eye movement studies (e.g., Altarriba *et al.*, 2001; Rayner, Balota, & Pollatsek, 1986; Rayner & Morris, 1992; Rayner, White, Kambe, Miller & Liversedge, 2003).

As discussed in Section 3.1.3.1 in Chapter 3, the present findings of no parafoveal-on-foveal effects might be considered to be inconsistent with the predictions of the SWIFT model that stipulates that the characteristics of a yet-to-be fixated (adjacent) word in the activation field can influence the fixation durations on the fixated word. As described, the findings of the present experiment showed no evidence that this was the case. In contrast, the findings are consistent with the predictions of the E-Z Reader model that stipulates that parafoveal processing occurs only after lexical processing of the foveal word has been completed and the programming of a saccade to the parafoveal word has been initiated. Therefore, the E-Z Reader model predicts that the lexical or semantic characteristics of the parafoveal word should not influence the fixation durations on the foveal word, a prediction that was met by the findings of Experiment 3 as well as Experiment 1.

5.1.3.2 Immediate Interactive Effects of SND and Word Frequency

The findings of Experiment 3 showed that the interaction between SND and word frequency was significant in all early and late reading measures on the critical word, consistent with the results of Experiment 2 and suggesting an interactive influence of the target word's frequency and its SND on lexical identification in normal reading. As predicted, high frequency words with high SND were fixated for the shortest time, and low frequency words with low SND were fixated for the longest time in all reading time measures. However, there was a shift in the nature of the interactive effect from the early to the later measures. In particular, the SND effect was larger for high than low frequency words in the early measures of single fixation and first fixation duration, while it was larger for low than high frequency in the later measures of gaze duration, regression path duration and total reading time. The visualisations of the interactive effects showed that word frequency effects were inhibitory for low SND, with increasing reading times for high than low frequency words. For high SND, however, word frequency effects were facilitatory, with decreasing reading times for high than low frequency words. Though it should be noted that the effect of word frequency at the lower SND values was less systematic than that observed for the high SND values, as indicated by the confidence intervals.

Based on the visualisations of the interactions, the shift of the SND effect observed in the means can be attributed to the counteractive influences of word frequency at different levels of SND. In particular, the (less systematic) inhibitory effect of frequency at the lower SND values was more pronounced in the later measures than the earlier measures (as indicated by the steeper upward projection in Panel A and B, especially for total reading time).

Before discussing the interactive effects on reading times, it is important to also note that neither the effect of SND, nor the interactive effect of SND and word frequency predicted skipping probability of the target words. Recall that the decision to skip a word must be made early in processing (when the word is in the parafovea) (Rayner, White, Kambe, Miller, & Liversedge, 2003). Accordingly, it is reasonable to assume that information about the word's SND was not obtained parafoveally (before fixating the word). This finding is consistent with the lack of parafoveal-on-foveal effects reported earlier and also consistent with the previous eye movement findings that suggest that mainly visual variables (e.g., word length), rather than lexical variables, are the strongest predictors of skipping probability (Brysbaert, Drieghe & Vitu, 2005, O'Regan, 1990, O'Regan & Levy-Schoen, 1987).

The joint effect of SND and word frequency found in this experiment is not entirely consistent with the interactive effect found by Buchanan *et al.* (2001) who used a lexical decision task. To remind the reader, Buchanan *et al.* found that the effect of SND was larger for low frequency compared to high frequency words. In contrast, based on the mean reading time results of the present experiment, the (high) SND effect was systematically larger for high frequency compared to low frequency words in the early measures. However, the nature of this effect changed and the visualisations of the current data indicate that, in fact, frequency had a differential modulatory influence on reading times at different levels of SND in early compared with later measures.

It is worth consideration of why the differences in results between Buchanan *et al.* and the present study occurred. It seems likely that three possible factors could have contributed to these differences. First, in their analyses, Buchanan and colleagues undertook basic Analyses of Variance to demonstrate their interactive effects. In contrast, Linear Mixed Modelling with advanced visualisation techniques were used to examine the current data set. Linear mixed modelling allows for more sophisticated examination of influences of variables at different levels of another variable, and in relation to different reading time measures reflecting different stages of the reading process. As such, it may well be possible that more complex patterns of influence may lie beneath the effects observed in the basic analyses of variance reported in the Buchanan *et al.* study.

A second point concerns the stimuli that Buchanan *et al.* employed. Buchanan *et al.* did not control for several extraneous variables in their stimuli. In contrast, in the present study, these variables were controlled. Target words in

Buchanan *et al.*'s study were matched in only word length. The target words in the present experiment were controlled for word length, orthographic neighbourhood size and frequency of orthographic neighbours. Orthographic neighbourhood size and frequency of orthographic neighbours have been reported to influence lexical processing in both isolated visual word recognition and normal reading studies (e.g., Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). Because Buchanan and colleagues did not control for orthographic neighbourhoods, the pattern of the interactive effect of SND and word frequency they obtained may have been influenced by them.

Finally, and in addition, Buchanan and colleagues used a single word recognition task while the present study used a normal reading task. This may also have resulted in differences between the results of the present experiment and those of Buchanan *et al*. In lexical decision tasks, participants are asked to make an overt decision about the presented target words to indicate that they have identified them. So response latencies obtained from these tasks also reflect a secondary task of making a decision about the target words. Reading tasks whilst recording eye movements, on the other hand, do not require participant to make such secondary tasks, and eye movement data reflect only the cognitive processes taking place in normal reading.

Indeed, previous eye movement research has demonstrated that the effects found in visual word recognition tasks can be different to the effects found in normal reading tasks. An example of such a discrepancy between results from the two paradigms is that observed for effects of orthographic neighbourhood size. Such effects were found to be facilitatory in lexical decision tasks (Andrews, 1997; Forster & Shen, 1996; Sears, Hino, & Lupker, 1995) and inhibitory in normal reading tasks (e.g., Pollatsek, Perea, & Binder, 1999). Arguably, then, the patterns of effects obtained in the current study are more relevant to those with an interest in normal reading since fixations in reading reflect lexical processing as it happens naturally (Kuperman *et al*., 2013).

Although the current analyses offer more sophisticated insight into the data sets than those offered by Buchanan *et al*., it remains necessary to provide an explanation of the changing modulatory influence of SND on frequency effects observed in the present study. To do this, it is necessary to once again consider the predictions that were formed in the Introduction of this experiment in relation to the current analyses. Recall that an interaction of SND and word frequency was predicted to influence word identification. Recall that this this derives from the fact that activation influences from the semantic level impact on the process of word identification before a word is fully identified.

As mentioned in Section 1.1 in Chapter 1, on average it takes about 250ms to identify a word. Therefore, the earliest two reading time measures that one should focus on in relation to these predictions are first fixation and single fixation duration. Let us consider the pattern observed in these two measures in relation to the prediction. Entirely consistent with the predictions, it was found that high frequency words with high SND were fixated for the shortest time compared to other conditions and that low frequency words with low SND were fixated for the longest time in these two measures. It was also found that the SND effect was larger for high frequency than low frequency words in first and single fixation duration. In addition, the findings showed that the (inhibitory) effect of word frequency was almost flat at the low SND levels compared to the robust (facilitatory) effect of word frequency at the higher SND levels.

The above-mentioned findings occurred for first and single fixation duration and met the predictions based on Stolz and Besner's (1996) embellished IA model (McClelland & Rumelhart, 1981). To reiterate, a high frequency word will rapidly inhibit the activation of its orthographic neighbours at the word level. This rapid inhibition within the word level allows activation to quickly feed forward to the semantic level, activating the semantic representation of the word. If the word's semantic representation is rich (e.g., closely packed semantic neighbours at the sematic level, i.e., high SND), then there will be strong activation at the semantic level. This strong activation will quickly feed back to the word level, contributing to resolving the competition between the orthographic competitors at the word level. That is, a high frequency word with high SND will have two strong sources of activation, reducing the time required to identify a word. If a high frequency word has a weaker semantic representation (e.g., distant semantic neighbours at the semantic level, i.e., low SND), then there will be weaker activation at the semantic level and, consequently, weaker feedback to the word level. As such, low SND will not have a comparable impact on the lexical processing of high frequency words due to the reduced semantic activation.

A low frequency word will take a longer time to inhibit the activation of its orthographic neighbours, compared to a high frequency word. Because of this longer inhibition at the word level, there will be very weak activation feeding forward from the word level to the semantic level. If the low frequency word's semantic representation is rich (e.g., high SND), then there will be strong activation at the semantic level and, as a result, enhanced feedback activation to the word level. Thus, the low frequency word's weaker activation at the word level can benefit from the enhanced semantic activation. If the low frequency word has a weaker semantic representation (i.e., low SND), then there will be weaker feedback semantic activation that will have a little, if any, impact on its lexical identification due to receiving activation from the lower levels (letter and visual feature levels) with which the low frequency word with low SND can be identified. That is, the low frequency word with low SND will have weaker activation at both the word level and the semantic level, and consequently, will take the longest time to be identified, as opposed to other conditions.

Now, let us consider the slightly later measure of gaze duration. The pattern of the interactive effects observed in gaze duration is similar to that observed in single and first fixation duration. However, remember that additional fixations on the word are included in the measure of gaze duration. As such, gaze duration, a measure that does not quite reflect processing as early as first fixation or single fixation, is still a measure that largely reflects word identification. In the present data, the pattern of effect for this measure starts to shift slightly. The SND effect was slightly larger for low than high frequency words in gaze duration. High frequency words tend to be identified in first fixation and single fixation duration compared to low frequency words that are more likely to receive additional fixations before leaving the words (i.e., gaze duration) (Inhoff & Rayner, 1986; Rayner, Sereno, & Raney, 1996) and, thus to be identified in gaze duration. Accordingly, it is reasonable to find that the SND effect was larger for the high frequency words in single and first fixation duration since they are more likely to be identified in these two measures (recall that SND influences occur before the word is uniquely identified). In the same way, it is also reasonable to find that the SND effect was larger for the low frequency words in the later measure of gaze duration since these words are more likely to be identified in gaze duration.

Finally, let us consider the later measures of regression path duration and total reading time. These measures reflect lexical and post-lexical processes associated with the integration of the words meaning into the meaning of the sentence, along with the construction of a coherent discourse representation. These effects are late. For this reason, we may well see patterns of the influence that differ in relation to the hypotheses that were generated in the basis of Stolz and Besner's (1996) embellished IA model. The findings showed that the inhibitory effects of frequency became more pronounced in these two late measures compared to these effects in the earlier measures (though these effects were less systematic). As described in Section 1.1 in Chapter 1, regression path duration involves making backward saccades to earlier parts of the sentence while total reading time involves making backward and forward saccades. Since a low frequency word with low SND is a difficult to process word with a weaker semantic representation (distant semantic neighbours), it is more likely that the readers found the meaning of this word difficult to process, and, thus, they immediately made use of the earlier and later parts of the sentence (by making backward and forward saccades) to successfully process its meaning and integrate it into the sentential context. On the other hand, a high frequency word with low SND is a relatively easy-to-process word with a weaker semantic representation. Accordingly, it seems that the reader found this word relatively easy to process, and thus, they did not need to make such immediate use of the earlier parts and the later parts of the sentence to process and integrate its meaning into their interpretation. Therefore, the effects of frequency at the lower SND levels appeared inhibitory in the later measures.

To sum up, the results of the present experiment showed that joint effect of word frequency and SND was significant in all early and late reading time measures. The nature of this interactive effect of SND and word frequency on different reading time measures was found to be complex. The SND effects were larger for high than low frequency words in earlier measures of single fixation and first fixation duration, but was larger or low than high frequency words in later measures of regression path duration and total reading time. The visualisation of the interactions showed that this shift occurred was due to the counteractive influences of word frequency at different levels of SND. In particular, the effects of word frequency for low SND values were inhibitory, whereas that effects for high SND values were facilitatory. These findings were taken to provide evidence for the assumption that SND (i.e., a semantic influence) could constrain unique word identification in normal reading via semantic feedback assumed in Stolz and Besner's (1996) embellished IA model (McClelland & Rumelhart, 1981).

The significant interactive effect of word frequency and SND found in the present experiment can be explained by both the E-Z Reader model (e.g., Reichle, Pollatsek, & Rayner, 2006) (a serial-attention-shift model) and SWIFT (Engbert *et al.*, 2002, 2005) (a guidance-by-attentional-gradient model). Since the interactive effect of target word frequency and SND was found in single fixation and first fixation duration, then this interaction should occur in the L1 stage of lexical processing during which the currently fixated word (n) is checked for its familiarity. Recall that the L1 stage of lexical processing of word n typically starts before word n is fixated while it is in the parafoveal vision. Then, the L1 processing on the parafoveal word continues when it is subsequently fixated. As such, orthographic and phonological processing of word n starts when the word n is in the parafovea (Inhoff & Topolski, 1994; Sereno & Rayner, 2000), allowing sufficient time for the rich semantic representation of word n to be activated during the L1 stage of lexical processing when word n is subsequently fixated. Accordingly, a greater amount of orthographic and phonological processing of a high frequency word is carried out when it is in the parafovea, compared to the amount of the parafoveal processing of a low frequency word. Given this, a rich semantic representation (e.g., high SND) can be activated during the L1 processing on the fixated high frequency word, and can be used to assess the orthographic familiarity of this word. Recall also that Reichle and colleagues developed the E-Z Reader model in a way that the L1 stage on word n is influenced by word frequency among other variables (e.g., word predictability). Taken together, both word frequency and SND can influence the same stage of lexical processing, namely the L1 lexical processing on the currently fixated word, as evident in the early reading time measures. If the high frequency word has a weak semantic representation (e.g., low SND), then the weak semantic representation will have a relatively little, if any, effect on the L1 stage when the word is fixated. Less orthographic and phonological information about a low frequency word, on the other hand, is processed in the parafovea, which does not allow sufficient time for the (high or low) SND characteristics of the low frequency word to be used during L1 stage to assess its familiarity. Based on this explanation, high frequency words with high SND are expected to be fixated for a shorter time than high frequency words with low SND and low frequency words with high (or low) SND, consistent with pattern of the means values of the reading times of the target words in the current experiment. Thus, the E-Z Reader model can provide a good explanation of the results of the present experiment.

The present findings of the immediate effect of the interaction of SND and word frequency can also be accounted by the SWIFT model (Engbert *et al.*, 2002, 2005). According to this model, as mentioned in Section 3.1.3.2 in Chapter 3, the target of a saccade is selected from an activation field, which evolves over time depending on linguistic and visual processing. A word with the highest activation (e.g., a difficult-to-process word) in the activation field is selected as the next saccade target. If a foveal word is highly activated at the time of saccade target selection, then it will inhibit the random timer from executing a forward saccade. As such, the foveal word receives a refixation. If the activation of a foveal word is low (i.e., an easy-to-process word), on the other hand, the random timer will not be inhibited by the foveal word, and will execute a saccade to the next word with the highest level of activation. Based on these assumptions of the SWIFT model, if the currently fixated word is a high frequency word with high SND (i.e., an easy-to-process word), the activation of this word will be lower than the case of a high frequency word with low SND or a low frequency word with either high or low SND (i.e., difficult-to-process words). As a result, the random timer will initiate a saccade to the subsequent word that has the highest activation in the activation field. Accordingly, a high frequency word with high SND should be fixated for less time than a high frequency word with low SND or a low frequency word with low or high SND.

It should be noted that the findings that neither SND nor the interaction of SND and word frequency influenced the skipping probability of the target words are more in line with the predictions of the E-Z Reader model than the SWIFT model. As described in different parts of this thesis, the SWIFT model assumes that multiple words are lexically processed at the same time, and, thus, semantic and lexical information about the word to the right of fixation is obtained parafoveally. As such, SND and word frequency of the yet-to-be fixated (adjacent) word could influence the skipping of the adjacent word under this model. The findings of the present experiment showed no evidence that this was the case. In contrast, the findings are consistent with the predictions of the E-Z Reader model that acknowledges processing of the parafoveal (adjacent) word at only orthographic and phonological levels. As such, SND information (i.e., semantic information) would not be obtained parafoveally under this model, and, therefore, would not influence the skipping probability of the adjacent words. Given this, it seems that the E-Z Reader model, rather than the SWIFT model, provides a better account of the skipping data.

5.1.3.3 Spillover Effects of SND and Word Frequency

In this part of the analysis, whether the interactive effect of target word frequency and SND spilled over onto next words was examined. The results showed that the spillover of this interactive effect was significant in first fixation and gaze duration on subsequent words. In particular, it was found that first fixation and gaze duration spent on next words were shortest following high frequency words with high SND than the other conditions. It should be noted that spillover effects occur frequently in reading (e.g., Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Pollatsek, Juhasz, Machacek, & Rayner, 2008; Rayner & Duffy, 1986; Slattery, Pollatsek, & Rayner, 2007), and the current finding are not particularly surprising.

The spillover of the interactive effect found in this experiment can be explained by the E-Z Reader model (e.g., Reichle, & Rayner, 2006; and Pollatsek *et al.*, 2008) as follows. Since the interactive effect was found to influence the lexical processing of the upcoming word (i.e., spillover effects) in this experiment, then it can be argued that the target words' SND and their frequencies also affected the L2 lexical processing on the target words (as per Pollatsek, Juhasz, Reichle, Machacek, & Rayner, 2008; Reingold & Rayner, 2006). If a currently fixated word n is a high frequency word with high SND (i.e., both characteristics of high frequency and high SND make the word easy to process), then the L2 processing on this word will be completed faster than if the word is a high frequency word with low SND or a low frequency word with high or low SND. This is because the enhanced semantic and lexical activation associated with the high frequency word with high SND helps to rapidly complete the L1 and L2 lexical processing. Therefore, the quick completion of the L2 lexical processing associated with high frequency words with high SND will allow more parafoveal preview of the subsequent word (n+1) (i.e., the time between attention has shifted to word n+1 and before the eyes start to move away from word n) compared to high frequency words with low SND or low frequency words with high or low SND. During this parafoveal preview, the familiarity check on word n+1 is carried out while still fixating the target word. Thus, the post-target word will have a head start when it is subsequently fixated following a high frequency word with high SND, as a great amount of its familiarity check will have already been carried out while fixating the previous word.

The SWIFT model of eye movement control (Engbert *et al.*, 2002, 2005) can also explain the spillover effects of the interaction of target word frequency and SND. If the currently fixated word is a difficult-to-process word, it is more likely that this word will be re-fixated by means of the foveal inhibition mechanism as mentioned in the previous section. Because of this foveal inhibition associated with processing a difficult fixated word, parafoveal processing of word n+1 will be reduced. As a consequence, fixation durations on word n+1 when it is subsequently fixated are longer than on average. However, if word n is an easy to process word (e.g., a high frequency with high SND), then the saccadic system (or the random timer) will not be inhibited by the properties of the fixated word, and there will be more parafoveal processing of word n+1. Therefore, subsequent fixation durations on word n+1 are shorter following a high frequency word with high SND than a high frequency word with low SND or a low frequency word with high or low SND.

5.1.4 Conclusion

Experiment 3 showed that the interaction of word frequency and SND of currently fixated words influenced the lexical processing of the fixated words and subsequent words, as evident in all early and late reading time measures. The effects of word frequency at the high SND levels were systematically facilitatory in all early and late measures, with decreasing reading times observed for high than low frequency words. The effects of word frequency at the lower SND levels were almost flat in the earlier measures of single, first fixation and gaze duration. These effects at the lower SND levels became more pronounced and inhibitory (though less systematically compared to these at high SND levels) in the later measures of regression path duration and total reading time. The effect of SND was larger for high than low frequency words in the early measures associated with lexical processing, while the opposite pattern was observed in the later measures associated with post-lexical processing.

These findings were consistent with the results of Experiment 2. As such, Experiment 3 and the previous two experiments consistently suggest that semantic neighbourhood density (SND) plays a role in word identification. The joint effects of word frequency and SND in the early measures reflecting lexical processing are predicted by Stolz and Besner's (1996) embellished IA model (McClelland & Rumelhart, 1981). In addition, the findings of the current experiment provided further evidence in support of Shaoul and Westbury's (2010a) claim about the nature of semantic representations during lexical processing.

Overall, the findings of the three experiments carried out so far support the assumption that SND can constrain the unique identification of a word's orthographic form. In order to provide a further validation of this assumption, the next experiment will test the interaction between SND and another word-level variable, namely, orthographic neighbourhood size. Therefore, the next experiment will determine whether the theoretical account provided, thus far, to explain when and how SND influences lexical identification is plausible.

Chapter 6: Interaction between Semantic Neighbourhood Density and Orthographic Neighbourhood Size

Experiment 3 suggested that there was a significant interaction between SND and word frequency. This significant interactive effect occurred in all early reading time measures associated with lexical processing as well as later measures associated with post-lexical processing. Thereby, Experiment 3 was concluded to provide support for the claim that SND could constrain unique word identification. To provide further evidence for this claim, Experiment 4 was carried out to examine the joint effect of SND and orthographic neighbourhood size (a word-level lexical variable). As such, the results of Experiment 4 will, potentially, offer further support for the theoretical account provided in the previous chapters to explain when and how SND influences lexical processing.

6.1 Experiment 4

Experiment 4 examined the interactive effect of the target words' SND and their orthographic neighbourhood size in lexical identification during reading of single sentences. Orthographic neighbourhood size was defined as the number of same-length words that could be generated by changing a single letter within a word, e.g., *mint*, *pint, hunt, hind, dint, hilt, lint,* and *tint* are neighbours of *hint* (Coltheart *et al.*, 1977*)*. The number of orthographic neighbours and SND were orthogonally manipulated while controlling for word frequency and word length. In addition, the number of high frequency orthographic neighbours was held constant; the frequencies of the target words were higher than the frequencies of their respective orthographic neighbours. The plausibility of the stimuli and the predictability of the target words were also controlled. These variables were controlled because they are known to influence lexical identification in normal reading as discussed in Chapter 1.

SND and Orthographic Neighbourhood Size

Previous studies showed that high orthographic neighbour words whose frequencies were higher than the frequencies of their respective orthographic neighbours were processed faster than matched low orthographic neighbour words (i.e., whose frequencies were higher than the frequencies of their orthographic neighbours) (Andrews, 1997; Perea & Rosa, 2000; Pollatsek, Perea, & Binder, 1999). The findings of Experiment 3 (along with the findings of the first and the second experiment in this thesis) suggest that SND can constrain unique word identification in normal reading. To provide further evidence for the conclusion made in Experiment 3, Experiment 4 was conducted to directly examine the joint effect of SND and orthographic neighbourhood size (as another word-level lexical variable).

Stolz and Besner's (1996) embellished IA model (McClelland & Rumelhart, 1981) was also used to generate the predictions of Experiment 4. Consider that we have two target words (*W1*, *W2*) that are matched on their word frequency (and word length). Both words have higher frequencies compared to the frequencies of their respective orthographic neighbours. *W1* has a higher number of orthographic neighbours while *W2* has a lower number of orthographic neighbours. A word with many orthographically similar neighbours (e.g., *W1*) has a more familiar orthographic pattern than a matched word with few orthographically similar neighbours (e.g., *W2*) (Ziegler, Muneaux, & Grainger, 2003). According to the IA model, both *W1* and *W2* will have higher baseline levels of activation at the word level compared to their respective orthographic competitors. However, the activation of *W1* will be, to some degree, higher than the activation of *W2* due to the fact that a high orthographic neighbour word (*W1*) is more orthographically familiar than a matched low orthographic neighbour word (*W2*). As such, the activation of a word unit corresponding to *W1* will more rapidly inhibit the activation of its orthographic competitors, compared to the activation of the word unit corresponding to *W2*. Because of this rapid within-word level inhibition, there will be an opportunity for the semantic representation of *W1* to be activated (via feedforward activation) and to influence lexical processing. If the semantic representation of *W1* is rich (e.g., high SND), then the rich semantic representation can feed back to the word level, constraining unique word identification as mentioned in the previous chapters. Similarly, the semantic representation of *W2* can be also activated and can impact on lexical processing; however, the activation of its semantic representation will be relatively later in time compared to a matched word with an increased orthographic neighbourhood (e.g., *W1*). Based on this theoretical account, it was predicted that high orthographic neighbour words with high SND would be fixated for less time than low orthographic neighbour words with high SND.

If *W1* and *W2* have weak semantic representations (e.g., low SND), on the other hand, then the weak semantic representations will provide weak activation feeding back from the semantic level to the word level. This is due to the nature of their distant semantic neighbours as discussed in Section 3.1.3.2 in Chapter 3. Therefore, the weak semantic representations (low SND characteristics) will not have a comparable impact on the lexical processing of *W1* and *W2*, compared to enhanced semantic representations (high SND).

Since the interactive effect of SND and orthographic neighbourhood size was predicted to influence lexical processing, then this joint effect should be mirrored in the fixation durations spent on target words themselves, and potentially, subsequent words in text if the effect spills over (as per the Kennison & Clifton, 1995; Pollatsek, Juhasz, Machacek, & Rayner, 2008; Rayner & Duffy, 1986). If the SND and orthographic neighbourhood size manipulation influences only the lexical processing of the target words, then this effect should appear in only early reading times measures such as single fixation, first fixation and gaze duration. If the present manipulation has a longer lasting effect, and influences later lexical processing and post-lexical processing, then this joint effect should also appear in later measures of reading times on target words, and also may spill over onto the words following target words.

If the characteristics of the word to the right of fixation affect the fixation durations on the currently fixated word (i.e., parafoveal-on-foveal effects), then it is expected that the investigated joint effect would influence the fixation durations on the prior word. Such findings will challenge the serial processing models of eye movement control during reading such as the E-Z Reader model that assumes that the parafoveal processing occurs only after the lexical processing of the foveal word has been completed. Such findings, on the other hand, would be more consistent with the assumptions of the parallel processing models such as the SWIFT model that assumes that multiple words are processed at a time, and predicts that the words to the right of fixation can influence the fixation times on the prior words.

6.1.1 Method

6.1.1.1 Participants

Forty-four students at the University of Southampton, selected according to the same criteria as those for Experiment 3 took part in Experiment 4. Participants were awarded either course credits or given £2.

6.1.1.2 Apparatus

The apparatus was the same as Experiment 1.

6.1.1.3 Materials and Design

10 sets of stimuli were created; each set contained four target words that were manipulated for SND and the number of orthographic neighbours. In total, there were 40 target words (10 high orthographic neighbour words with high SND (HSND-HON), 10 high orthographic neighbour words with low SND (LSND-HON), 10 low orthographic neighbour words with high SND (HSND-LON), and 10 low orthographic neighbour words with low SND (LSND-LON)). Table 6.1 presents the descriptive statistics of the characteristics of the target words. The target words in each set were matched on word length and word frequency (Fs < 1). The high and low SND target words significantly differed in SND ($F(3, 36) = 26.32$, $p = 0.0005$) and high and low orthographic neighbour words differed in orthographic neighbourhood size ($F(3, 36) = 7.13$, $p = 0.001$). Orthographic neighbourhood frequency was controlled so that the frequencies of the target words were higher than any of their respective orthographic neighbours (i.e., the frequency of the orthographic neighbours of the target words did not exceed the frequency of the target words themselves).

Table 6.1 *Means and Standard Deviations (in Parentheses) of the Characteristics of the Target Words Used in Experiment 4.*

	HSND-HON		HSND-LON		LSND-HON		LSND-LON	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
SND	0.55	(0.03)	0.55	(0.03)	0.43	(0.07)	0.36	(0.07)
Log BNC	3.68	(0.21)	3.71	(0.09)	3.55	(0.35)	3.57	(0.17)
Length	5.70	(0.67)	5.90	(0.87)	5.90	(0.87)	5.90	(0.87)
ON	3.50	(2.32)	1.10	(0.73)	2.90	(1.28)	1.10	(0.99)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HON: high orthographic neighbours; LON: low orthographic neighbours; BNC: word frequency from the British National Corpus; length: number of letters; ON: orthographic neighbourhood size.

Initially, eight sentences were created for each set such that any of the four target words within a set could fit plausibly in the eight sentence frames. All the eight sentences for each target word within a set were pre-screened for plausibility and predictability as will be described later in this section. After pre-screening the sentences, the top four sentences in each set that were given the highest plausibility and the lowest predictability were selected to be used in the experiment. In total, there were 40 experimental sentences that were presented to the participants. Four lists of these sentences were created, with each list containing all 40 sentences. For each set of stimuli, the same sentence in each list differed in the target word as can be seen in Table 6.2. In this way, each participant was presented with all of the 40 sentence frames and all the target words in the present experiment, maximising the statistical power. A full set of the stimuli used in Experiment 4 can be found in Appendix D.

Table 6.2 *A Sample of the Sentences Containing the Experimental Manipulation in Experiment 4. The Target Words are Presented in Bold.*

	HSND-HON	HSND-LON	LSND-HON	LSN-LON
	Spoon	**Flute**	**Buggy**	**Duvet**
List A	Jane bought that expensive **spoon** from this shop.	She looked at the little **flute** in the catalogue.	I bought this large **buggy** at a discounted price.	She threw out the white **duvet** that she had owned for many years.
List B	She looked at the little **spoon** in the catalogue.	I bought this large **flute** at a discounted price.	She threw out the white **buggy** that she had owned for many years.	Jane bought that expensive **duvet** from this shop.
List C	I bought this large **spoon** at a discounted price.	She threw out the white **flute** that she had owned for many years.	Jane bought that expensive **buggy** from this shop.	She looked at the little **duvet** in the catalogue.
List D	She threw out the white **spoon** that she had owned for many years.	Jane bought that expensive **flute** from this shop.	She looked at the little **buggy** in the catalogue.	I bought this large **duvet** at a discounted price.

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HON: high orthographic neighbours; LON: low orthographic neighbours.

Before conducting Experiment 4, all the sentences were pre-tested for plausibility and predictability using pen and paper questionnaires. In the plausibility ratings, the participants were asked to rate how likely it was that the event in the given sentences would occur. These ratings were made on a 7-point Likert scale (1 = *very implausible*, 7 = *very plausible*) by four participant groups (twelve participants in all of the four groups: three participants assigned to each list). The results showed that the sentences in the four lists were plausible (list A *mean* = 4.85; list B *mean* = 4.93; list C *mean* = 4.78; list D *mean* = 4.88), and the one-way ANOVA indicated that the four lists were not statistically different from each other in terms of plausibility (*F* < 1). A fifth group of twelve participants completed a predictability cloze test in which they saw the beginning of the sentences up to the word preceding the target words and were asked to complete the sentences with the most obvious word that came to mind. The result of the cloze test showed that none of the participants predicted the target words (the total number of predicted target words = 0).

6.1.1.4 Procedure

The procedure followed in this experiment was the same as Experiment 1.

6.1.2 Results

Prior to data analysis, data trimming was carried out along with removing outliers following the same criteria as in Experiment 1, and this resulted in removal of 0.48% of the data prior to the analyses. The dependent variables were single fixation duration, first fixation duration, gaze duration, regression path duration, total reading time and skipping probability (see Section 3.1.1.5 in Chapter 3 for a description of the measures). A normal Quantile-Quantile plot (Wilk & Gnanadesikan, 1968) was obtained to check whether the fixation durations (the dependent variables of this experiment) were normally distributed. The plot indicated that the fixation durations were not normally distributed. Therefore, the fixation durations were log-transformed to approximate a normal distribution.

As in the previous experiments reported in this thesis, all measures were analysed with linear mixed effect (LME) models. Participants and items were entered as random effects and the target word frequency, word length and orthographic neighbourhood size followed by the SND metric (ARC, Shaoul & Westbury, 2010b) were all entered in the models as fixed effects, one variable at a time. These variables were entered in the models before the interaction terms in order to examine the unique contribution of the interaction of SND and orthographic neighbourhood size in lexical processing. Finally, interaction terms (frequency * ARC, and length * ARC, orthographic neighbourhood size * ARC) were subsequently added to the resulting model also as fixed effects. All the fixed variables including SND were entered as continuous variables, and were all centred at the means to minimise collinearity in the analysis of the data. To make interpretation of the data easier, orthographic neighbourhood size and SND were dichotomised using a median split (HON > 2; LON ≤1; HSND > 0.49; LSND < 0.45) when presenting and discussing the findings of the LME models (note, though, as specified earlier, the frequency and SND were entered in the models as continuous variables). The following statistics will be reported in the results of this experiment: the regression coefficients (*b*), standard errors (*SE*), *t* values (or the *Z* value in the case of the skipping probability) together with *p* values based on Markov Chain Monte Carlo sampling 10,000 samples (Baayen *et al*., 2008).

As in Experiment 1 and 3, all sentences were divided into four regions as shown in the example below.

REGION	1	2	3	4
	She looked at the/	little/	**buggy**/	in the catalogue/.

The particular regions of interest for the analyses of this experiment were Region 2 constituting a pre-target word, Region 3 constituting a target word, and Region 4 constituting post-target words. The criterion for determining the post-target region was identical to the criterion used in Experiment 1.

The overall mean comprehension rate was 94.79% indicating that the participants read and understood the sentences. The reported results will include: 1) the joint effect of SND and orthographic neighbourhood size on the fixation times on the pre-target words (i.e., parafoveal-on-foveal effects), 2) the interaction of SND orthographic neighbourhood size on fixation times on the target words, and 3) the joint effects of SND and orthographic neighbourhood size on the fixation times on the post-target words (i.e., spillover effects).

To investigate whether the characteristics of the target words influenced the fixation durations on the pre-target words, LME analyses were conducted. The reading time measures for the pre-target words were the dependent variables; participants and items were entered as random effects, and the predictors of target word orthographic neighbourhood size and SND along with the joint effect of these two variables were entered as fixed effects. There were no reliable effects for these analyses (all *t*s < 0.9) providing no evidence of any parafoveal on foveal effects (see Table 6.3 for the mean values associated with these analyses).

Table 6.3 *Means of Fixation Times on the Pre-Target Words Preceding Target Words with the Experimental Manipulation in Experiment 4. Standard Deviations are Given in Parentheses.*

	HSND-HON	HSND-LON	LSND-HON	LSND-LON
Single Fixation Duration	214 (30)	217 (73)	218 (71)	217 (80)
First Fixation Duration	217 (74)	218 (69)	217 (68)	218 (86)
Gaze Duration	243 (108)	249 (111)	246 (119)	244 (129)
Regression Path Duration	294 (187)	291 (176)	291 (175)	297 (210)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HON: high orthographic neighbour; LON: low orthographic neighbours.

Next, the interaction of SND and orthographic neighbourhood size on the fixation times on the target words was examined. Table 6.4 and 6.5 list the results of the LME analyses carried out on the fixation times on the target words. As can be seen from these two tables, in the baseline models, word frequency was not a significant predictor of any reading time measures (all *t*s < 2), which is not surprising as the target words of this experiment did not vary in terms of word frequency as well as other variables such as word length. Also, the effect of word length was not significant in all reading time measures (*t*s < 2), except marginally in total reading time (b = 63.20, SE = 35.86, t = 1.762). Of primary theoretical importance was the interactive influence of a word's neighbourhood and SND. The main effect of orthographic neighbourhood size was not significant in any measures (*t*s < 2), except gaze duration (b = 3.468, SE = 1.550, t = -2.24, p = 0.03). Increasing the number of orthographic neighbours led to decreased gaze duration on the target words. Recall that the frequencies of the target words in this experiment were higher than the frequencies of their respective orthographic neighbours. This finding is consistent with the prediction made based on the previous studies that indicated that high orthographic neighbour words whose frequencies were higher than the frequencies of the their respective orthographic neighbours were processed faster than matched low orthographic neighbour words (Andrews, 1997; Perea & Rosa, 2000; Pollatsek, Perea, & Binder, 1999). It may be striking that the main effect of orthographic neighbourhood size was found significant in only one measure of reading time. However, as mentioned elsewhere in this thesis, the effect orthographic neighbourhood size is not as influential as other variables, and we might therefore expect less robust effects of this variable.

SND also did not predict the skipping probability (after running a logistic LME) (b = -0.254, SE = 0.992, Z = 0.256, p > 0.79), consistent with the parafoveal-on-foveal findings in that SND information was not obtained parafoveally. The main SND effect was not significant in any reading time measures in this experiment (all ts < 2). Of more theoretical interest here, however, is the interactive influence of SND and orthographic neighbourhood size. SND also did not predict the refixation probability (b = 0.661,SE = 0.644, Z = 1.025, p > 0.30).

The interaction between the target words' SND and orthographic neighbourhood size appeared in the early measures of first fixation duration: b = -23.78, SE = 11.80, t = -2.04, p = 0.04 and gaze duration: b = -41.108, SE = 17.42, t = -2.36, p = 0.01, as well as the comparatively late measure of regression path duration (b = -60.39, SE = 24.10, t = -2.50, p = 0.01). There was no interaction between SND and orthographic neighbourhood size for target word skipping probability (b = 0.425, SE = 0.644, Z = -1.31, p > 0.50).

Table 6.4 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Single Fixation, First Fixation, and Gaze Duration Recorded on the Target Words as Dependent Variables (Experiment 4).*

	Log Single Fixation Duration			Log First Fixation Duration			Log Gaze Duration		
	Variance	Std. Dev.		Variance	Std. Dev.		Variance	Std. Dev.	
Random effects									
Item (intercept)	232.2	15.24		176.92	13.301		625.39	25.008	
Subject (intercept)	948.8	30.80		657.53	25.642		926.23	30.434	
Residual	2869.6	53.57		3165.35	56.261		5104.48	71.446	
AIC			10652.89			16644			16968
BIC			10706.56			16665			16989
logLik			-5315.44			-8318			-8480
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects									
Intercept	459.951	214.786	2.141	221.2404	4.8117	45.98	246.048	6.227	39.51
Freq. n	-90.181	59.261	-1.522	-5.999	7.960	-0.754	-12.77	12.01	-1.064
Length n	15.970	15.706	1.017	0.3397	2.2528	0.15	4.878	3.397	1.44
ON n	-4.449	8.344	-0.533	-0.7694	1.0256	-0.90	3.468	1.550	-2.24*
ARC	-443.843	502.843	-0.883	21.7655	25.9868	-0.86	-35.132	27.286	-1.29
Interactions									
Freq. * ARC	184.275	133.243	1.383	27.1937	93.0655	0.29	103.238	138.111	0.75
Length * ARC	-38.873	33.047	-1.176	-0.1524	1.0194	-0.15	-4.789	37.426	-0.13
ON * ARC	-11.312	17.020	-0.665	-23.7812	11.8024	-2.04*	-41.108	17.426	-2.36*

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients (1.645 ≤t<1.96); robust significant coefficients (t≥2). No significant coefficients (t <1.645);

Table 6.5 *Results from Linear Mixed Effects Models Fit by Restricted Maximum Likelihood (REML); Regression Coefficients with Associated Standard Errors and t-Values of Fixed Effects on Regression Path Duration and Total Reading Time Recorded on the Target Words as Dependent Variables (Experiment 4).*

	Log Regression Path Duration			Log Total Reading Time		
	Variance	Std. Dev.		Variance	Std. Dev.	
Random effects						
Item (intercept)	958.68	30.962		2101	45.84	
Subject (intercept)	2009.05	44.822		2794	52.86	
Residual	12167.1	110.305		15496	124.48	
AIC			18728			18592.17
BIC			18750			18650.40
logLik			-9360			-9285.08
	Estimate	Std. Error	*t*-value	Estimate	Std. Error	*t*-value
Fixed effects						
Intercept	349.56	60.20	5.806	566.90	496.74	1.141
Freq. n	-21.44	16.45	-1.304	-90.05	135.09	-0.667
Length n	6.326	4.700	1.35	63.20	35.86	1.762*
ON n	2.306	2.134	1.081	-2.328	2.858	-0.815
ARC	20.462	38.491	0.53	-21.357	50.540	-0.423
Interactions						
Freq. * ARC	-7.689	4.802	-1.601	307.580	303.323	1.014
Length * ARC	68.081	52.254	1.303	89.614	64.815	1.383
ON* ARC	-60.394	24.106	-2.505*	-52.036	32.864	-1.583

Note. Asterisks correspond to significant effects as follows: marginally significant coefficients (1.645 ≤t<1.96); robust significant coefficients (t≥2). No significant coefficients (t<1.645).

To gain a general insight into the nature of the interactive effect of SND and orthographic neighbourhood size, the pattern of effects observed in the mean reading times for each measure was considered (see Table 6.6). Recall that it was predicted that high SND would benefit the lexical identification of high orthographic neighbour words to a greater extent than low orthographic neighbour words (recall that the frequencies of the target words in this experiment were higher than the frequencies of their respective orthographic neighbours). This prediction was made based on the previous studies that indicated that high orthographic neighbour words whose frequencies were higher than the frequencies of the their respective orthographic neighbours were processed faster than matched low orthographic neighbour words (Andrews, 1997; Perea & Rosa, 2000; Pollatsek, Perea, & Binder, 1999). Given this, a high orthographic neighbour word will activate its semantic representation sooner than a low orthographic neighbour word will due to being more active at the word level. As a consequence, the high orthographic neighbour word will benefit from enhanced semantic feedback activation (provided by high SND characteristics) more quickly than the low orthographic neighbour word will.

The patterns of effects observed in the mean of reading times of target words are presented in Table 6.6. The discussion here will focus on the reading time measures of first fixation, gaze duration and regression path duration that showed significant interactive effects of SND and orthographic neighbourhood size. As can be seen from Table 6.6, high orthographic neighbour words with high SND were fixated for the shortest time compared to other conditions in these measures. The other conditions had quite comparable reading times in earlier measures of first fixation duration (HSND-LON: 223ms, LSND-HON: 224ms, LSND-LON: 220ms) and gaze duration (HSND-LON: 245ms, LSND-HON: 247ms, LSND-LON: 247ms). However, their reading times were quite different in the later measure of regression path duration (HSND-LON: 279ms, LSND-HON: 292ms, LSND-LON: 257ms). In these three measures (first fixation, gaze duration and regression path duration), the SND effect was larger for high than low orthographic neighbour words (HON: 16ms, 35ms, 39m; LON: 3ms, 2ms, 22ms, respectively).

Table 6.6 *Means of Fixation Times on the Target Words in Experiment 4. Standard Deviations are Given in Parentheses.*

	HSND-HON	HSND-LON	LSND-HON	LSND-LON
Single Fixation Duration	214 (30)	223 (45)	228 (39)	230 (47)
First Fixation Duration	208 (25)	223 (40)	224 (31)	220 (35)
Gaze Duration	212 (28)	245 (43)	247 (43)	247 (46)
Regression Path Duration	253 (47)	279 (72)	292 (76)	257 (47)
Total Reading Time	242 (38)	299 (79)	312 (47)	280 (62)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HON: high orthographic neighbour; LON: low orthographic neighbours.

Visualisations of the significant interactions obtained from the LME models are presented in Figure 6.1, 6.2, and 6.3 (each figure represents one measure of reading time). These graphical figures display the interaction of SND (*arc*) and orthographic neighbourhood size (*on*) in first fixation (*ff*), gaze duration (*gd*), and regression path duration (*goPast2*). The details of what information is represented in these figures are provided in Section 4.1.2 presented in Chapter 4 (pp. 126-128). To reiterate, the order in which the panels should be interpreted is from Panel A to Panel E, and the orange bar in each panel represents a value of *arc* (the SND metric) at which the effect of orthographic neighbourhood size is plotted. In Panel A, the in Figure 6.1 (first fixation duration), the line is upward, with a broad confidence interval. As we move though Panel B to Panel C, the line slope of this upward projection becomes flatter. Then, as we move through Panel D to Panel E, it can be seen that the slope of the line becomes downward (and gradually steeper), and the confidence interval is reduced. Panels A-C can be interpreted as showing that for reduced SND, the influence of orthographic neighbourhood size is inhibitory, that is, reading times are shorter for low than high orthographic neighbour words at this reduced level of SND. However, this inhibitory effect gradually diminishes with the increase of the SND level (through Panels B-C representing mid-low and medium SND levels). As the level of SND increases (through Panels D-E) this trend shifts such that it lies in the opposite direction. That is to say, in first fixation duration, for higher SND, the influence of orthographic neighbourhood size is facilitatory, with reduced reading times for high than low orthographic neighbour words. The latter influence is consistent with the theoretical prediction outlined in the Introduction of this experiment. Broadly, this pattern is comparable with that in Figure 6.2 for gaze duration and Figure 6.3 for regression path duration.

To summarise, the interactive effect of SND and orthographic neighbourhood size was found to be significant in first fixation, gaze duration and regression path duration. From the means in Table 6.6, it is clear that the effect of SND was larger for high than low orthographic neighbour words in these three reading measures. High orthographic neighbour words with high SND were read for the shortest time compared to the other conditions, and the reading times for other conditions were quite comparable in the earlier measures of first fixation and gaze duration. However, the reading times for these other conditions were quite different in regression path duration. The visualisations of the data in Figure 6.1-6.3 show that effects of orthographic neighbourhood size are inhibitory for low SND values, whereas the effects of orthographic neighbourhood size are facilitatory for high SND values.

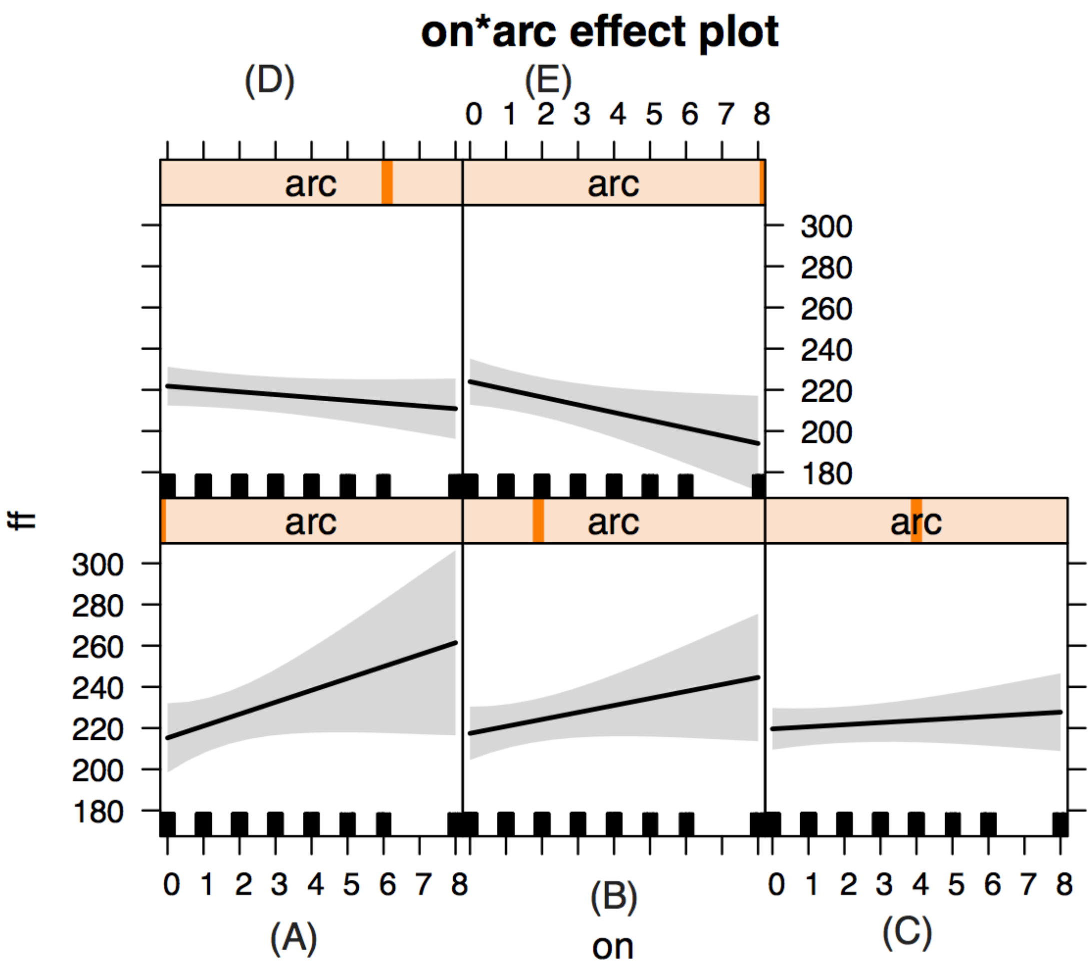

Figure 6.1 Effect display for the significant interaction of orthographic neighbourhood size (*on*) and SND (*arc*) in the LME model fit to the first fixation duration for the data of the target words (Experiment 4). The vertical axis is labelled on the first fixation duration (*ff*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

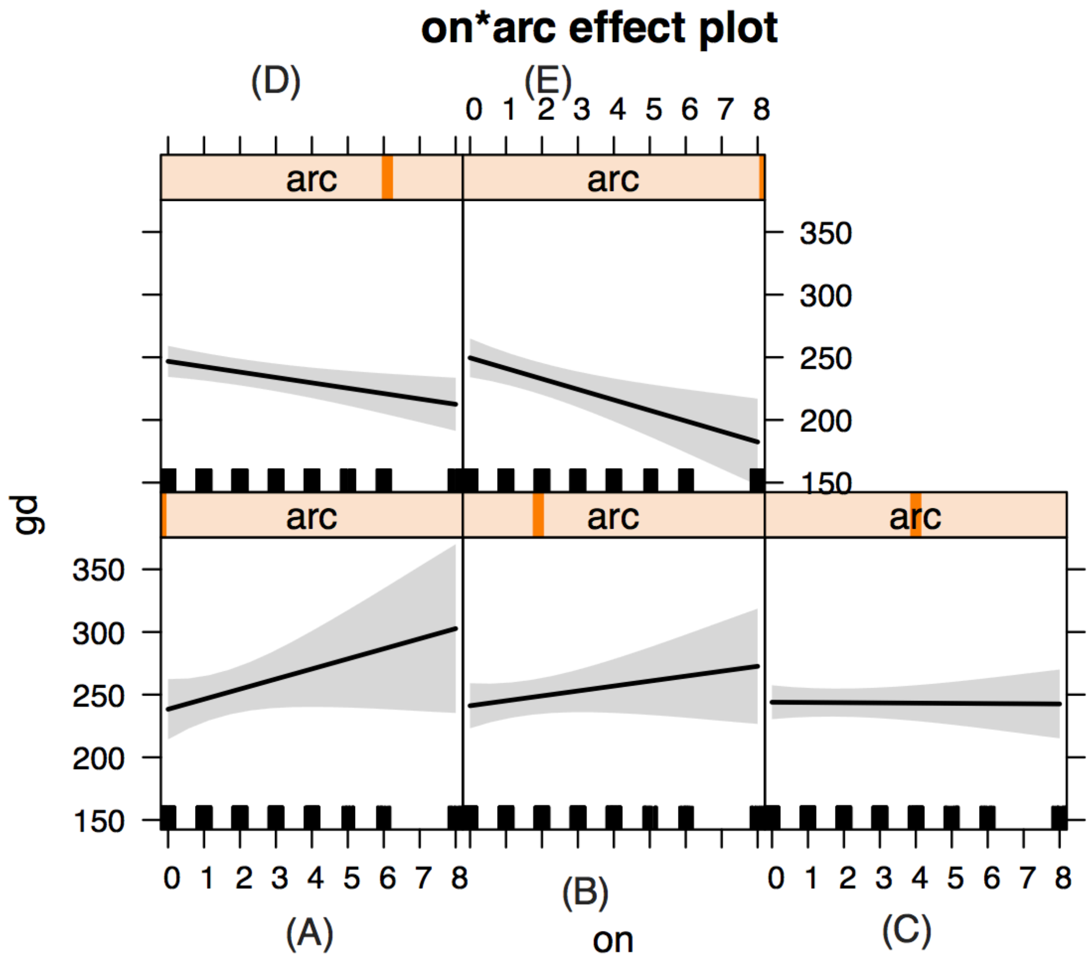

Figure 6.2 Effect display for the significant interaction of orthographic neighbourhood size (*on*) and SND (*arc*) in the LME model fit to the gaze duration for the data of the target words (Experiment 4). The vertical axis is labelled on the gaze duration (*gd*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

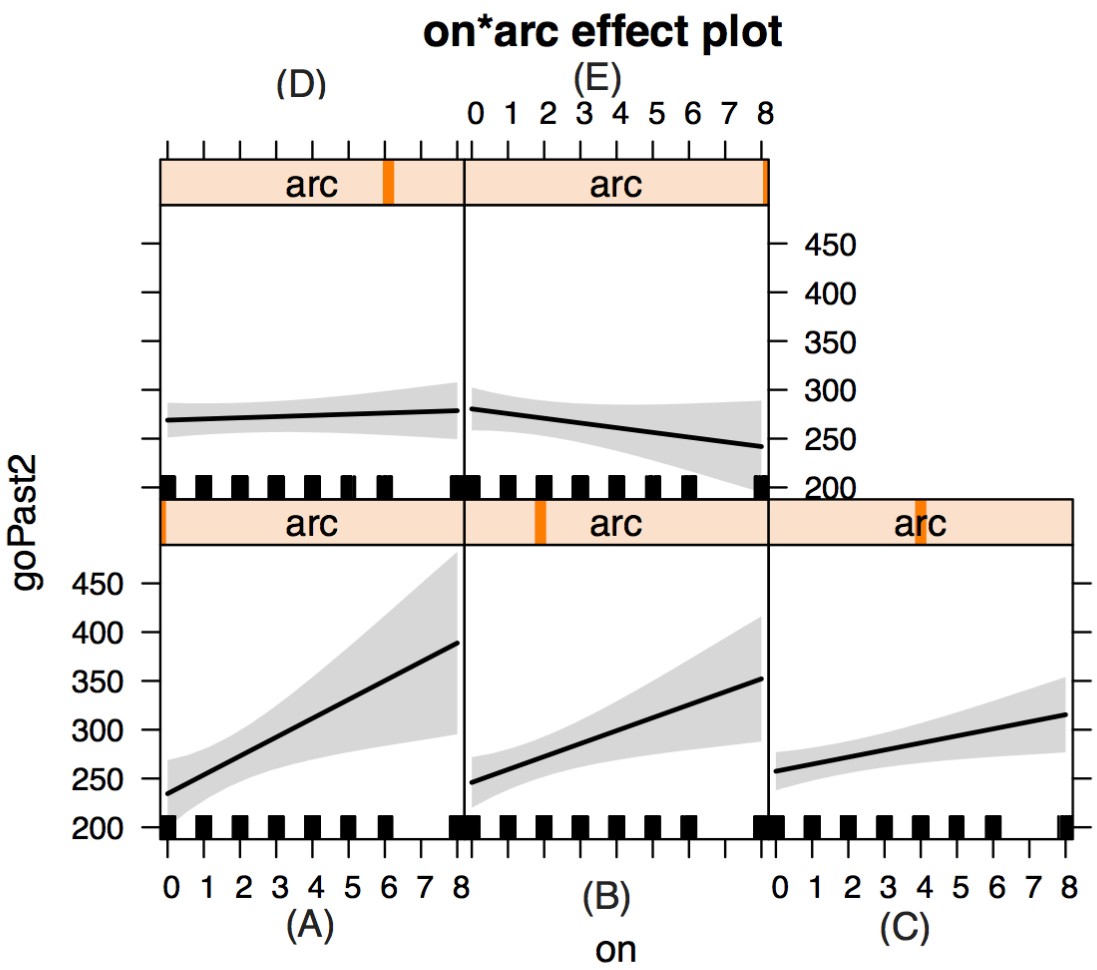

Figure 6.3 Effect display for the significant interaction of orthographic neighbourhood size (*on*) and SND (*arc*) in the LME model fit to the regression path duration for the data of the target words (Experiment 4). The vertical axis is labelled on the regression path duration (*goPast2*) on the target words, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

The last aspect of the data that was considered concerned whether effects due to the characteristics of the target words spilled over onto subsequent words. Again, LME models were conducted. The reading times for the post-target words were dependent variables, participants and items were entered as random effects, and the predictors of target word orthographic neighbourhood size, SND, and the interaction of these variables were entered as fixed effects. The results showed that target words' orthographic neighbourhood size was a significant predictor of early fixation times on the post-target words (single fixation duration: $b = -8.484$, $SE = 2.817$, $t = -3.012$; first fixation duration: $b = -8.051$, $SE = 2.323$, $t = -3.47$; gaze duration: $b = -7.195$, $SE = 2.534$, $t = -2.829$). As the negative *b* coefficients indicate, fixation times on the post-target words were shorter following high orthographic neighbour words than low orthographic neighbour words. The target words' SND also significantly predicted regression path duration on the post-target region ($b = -235.13$, $SE = 87.85$, $t = -2.677$). The joint effect of target words' SND and orthographic neighbourhood size significantly predicted first fixation duration on the post-target region ($b = -56.83$, $SE = 26.36$, $t = -2.156$), and this joint effect was marginal in single fixation duration ($b = -35.57$, $SE = 31.54$, $t = -1.762$). The interactive effects of orthographic neighbourhood size and SND were not significant in all other measures (all *t*s < 1.20).

The pattern of effects observed in the mean of reading times of post-target region is presented in Table 6.7. The discussion here will focus on the reading measures of single fixation and first fixation duration that showed significant spillover effects of the interactive effects of target word orthographic neighbourhood size and SND. As can be seen from Table 6.7, the post-target region following high orthographic neighbour words with high SND was fixated for the shortest time compared to other conditions, and the post-target region following low orthographic neighbour words with low SND was fixated for the longest time relative to the reading times in other conditions. For first and single fixation duration, the SND effect was slightly larger for high orthographic neighbour words than low orthographic neighbour words in the spillover region (HON: 10ms, 11ms; LON: 6ms, 3ms, respectively).

Table 6.7 *Means of Fixation Times on the Pre-Target Words Preceding Target Words with the Experimental Manipulation in Experiment 4. Standard Deviations are Given in Parentheses.*

	HSND-HON	HSND-LON	LSND-HON	LSND-LON
Single Fixation Duration	234 (99)	251 (104)	244 (98)	257 (120)
First Fixation Duration	227 (97)	248 (105)	238 (100)	251 (114)
Gaze Duration	249 (141)	274 (139)	261 (144)	278 (153)
Regression Path Duration	339 (236)	354 (282)	360 (302)	370 (343)

Note. HSND: high semantic neighbourhood density; LSND: low semantic neighbourhood density; HON: high orthographic neighbour; LON: low orthographic neighbours.

Visualisations of the significant interactive effect of orthographic neighbourhood size and SND on the fixation times of the post-target region (obtained from the LME models) are presented in Figure 6.4, and 6.5 (each figure represents one measure of reading time). Broadly, the interactive effects observed on the post-target region as presented in these figures are comparable to those observed in the early measures of first fixation and gaze duration of the target words as presented in Figure 6.1 and 6.2. For low SND (see Panels A-C in Figure 6.4-6.5), the effects of target word orthographic neighbourhood size on the fixation times of the post-target region were inhibitory, while the effects of orthographic neighbourhood size were facilitatory for high SND (see Panels D-E in both figures).

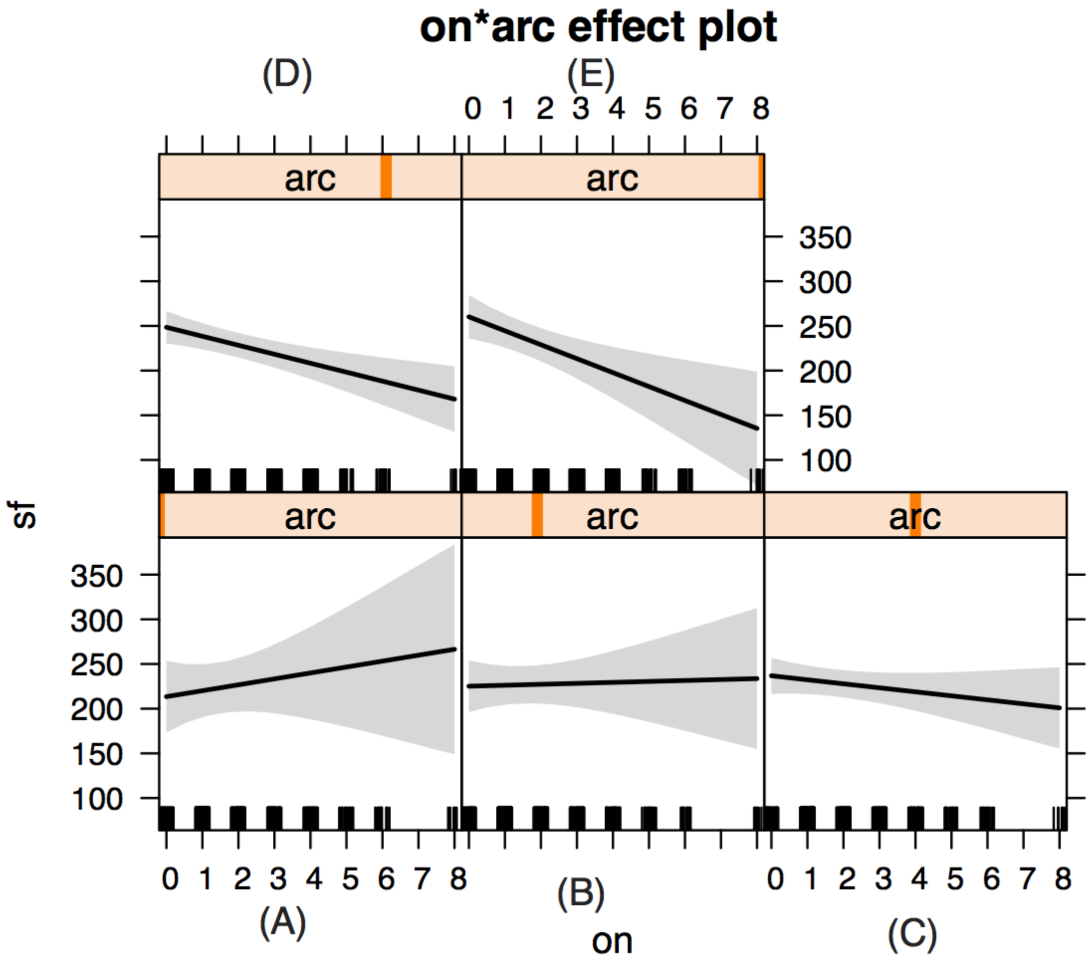

Figure 6.4 Effect display for the effect of frequency by SND (*arc*) in the LME model fit to the single fixation duration for the spillover region data (Experiment 4). The vertical axis is labelled on the single fixation duration (*sf*) on the post-target region, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

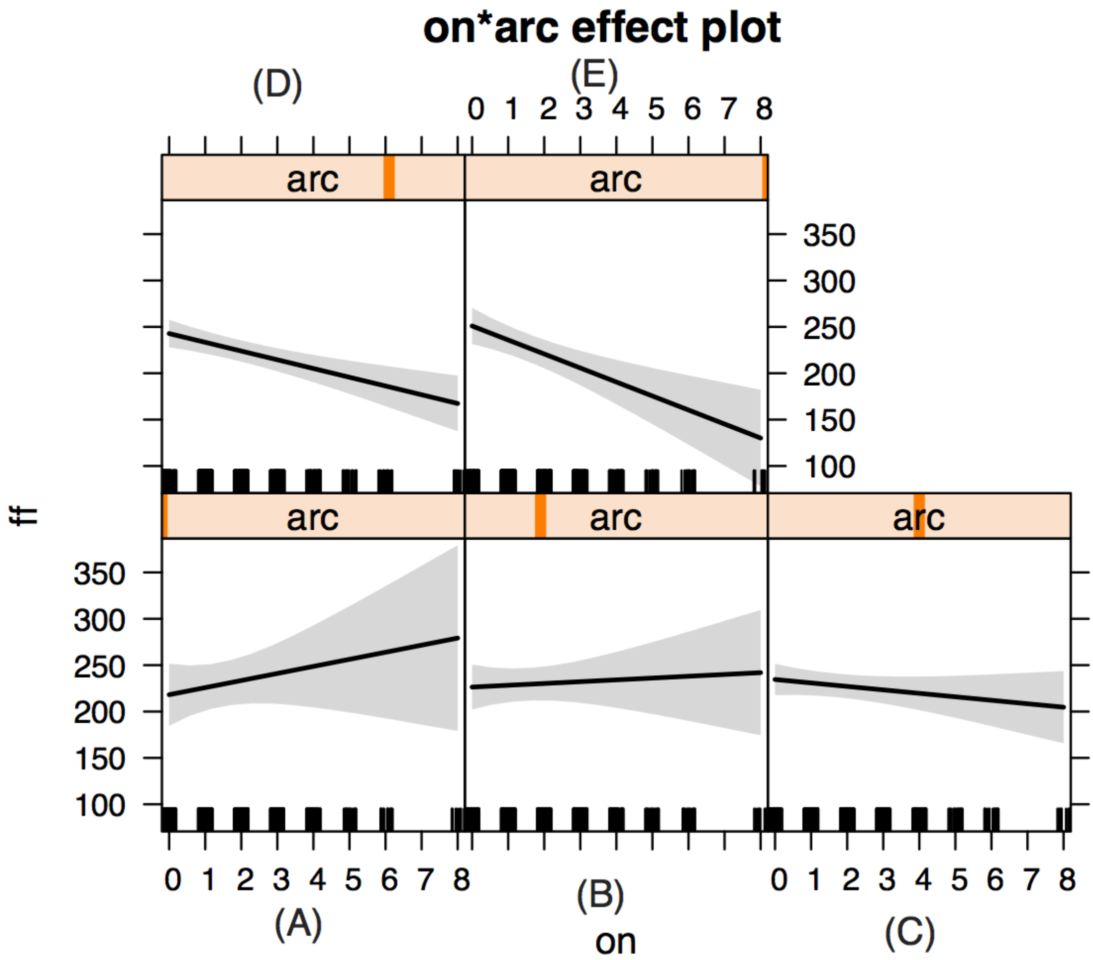

Figure 6.5 Effect display for the effect of frequency by SND (*arc*) in the LME model fit to the first fixation duration for the spillover region data (Experiment 4). The vertical axis is labelled on the first fixation duration (*ff*) on the post-target region, and a 95-percent pointwise confidence interval (the grey shaded region) drawn around the estimated effect.

6.1.3 Discussion

Experiment 4 investigated 1) whether the interaction between target word orthographic neighbourhood size and SND would influence the fixation times on prior words (i.e., parafoveal-on-foveal effects), 2) whether the interactive effect of the target words' SND and their orthographic neighbourhood size would influence the fixation times on target words, and 3) whether this interactive effect would influence the fixation times on the subsequent words (i.e., spillover effects).

6.1.3.1 Parafoveal-on-Foveal Effects of SND and Orthographic Neighbourhood Size

The findings of Experiment 4 showed that orthographic neighbourhood size, SND, and the joint effect of these two variables associated with the yet-to-be fixated word (i.e., the target word) did not predict the fixation times on the currently fixated word (i.e., the pre-target word). This result is consistent with the findings of Experiment 1 and 3 in relation to parafoveal-on-foveal effects, and consistent with the findings of many eye movement studies with respect to null parafoveal-on-foveal effects (e.g., Altarriba *et al.*, 2001; Rayner, Balota, & Pollatsek, 1986; Rayner & Morris, 1992; Rayner, White, Kambe, Miller & Liversedge, 2003). As such, the results of Experiment 1, 3 and 4 might be considered to be inconsistent with the assumption of parallel processing models of eye movement control (e.g., the SWIFT model), as there was no evidence in the results of these experiments for the claim that the processing of a currently fixated word can be affected by the lexical or semantic characteristics of the parafoveal word. Instead, the findings are congruent with the assumptions of the serial processing models of eye movement control (e.g., the E-Z Reader model) that postulates that parafoveal processing occurs only after the lexical processing of the foveal word has been completed.

6.1.3.2 Immediate Joint Effects of SND and Orthographic Neighbourhood Size

The findings of Experiment 4 showed that the interaction between SND and orthographic neighbourhood size was significant in the earlier reading time measures of first fixation and gaze duration as well as the later measure of regression path duration. As predicted, high orthographic neighbours with high SND were fixated for the shortest time in these reading time measures, compared to other conditions. In addition, the effect of SND was larger for high than low orthographic neighbour words. The visualisations of the interactive effects of orthographic neighbourhood size and SND provided a sophisticated insight into the changing modulatory influence of SND on the effects of orthographic neigbourhood size observed in the present study. In all these three reading time measures, the effects of orthographic neighbourhood size were inhibitory for the low SND values, with increasing reading times for high than low orthographic neighbour words. However, this inhibitory effect was reduced in its systematicity as indicated by the broad confidence intervals, and gradually diminished with the increase of the SND level (mid-low and medium SND levels). For the higher SND values, on the other hand, the effects of orthographic neighbourhood size were systematic and facilitatory, with decreasing reading times for high than low orthographic neighbour words.

Before offering an explanation of the above-described modulatory influence, it is important to note the following. The finding of facilitatory effects of orthographic neighbourhood size in the case of high SND may seem initially incongruent with the findings of previous eye movement studies that demonstrated a late inhibitory effect of orthographic neighbourhood size (e.g., Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). However, the present findings are not totally inconsistent with previous eye movement work. Indeed, Pollatsek, Perea and Binder (1999) matched high and low orthographic neighbour words on the number of high frequency orthographic neighbours, and manipulated the number of low frequency orthographic neighbours (in their third experiment). They showed that increasing the number of low frequency orthographic neighbours increased the likelihood of skipping the target words. Therefore, Pollatsek *et al.*'s findings suggest that the effect of increasing the number of orthographic neighbours can be initially facilitatory in lexical identification, particularly in the case that the frequencies of the target words are higher than the frequencies of their orthographic neighbours, similar to the manipulation of the preset experiment.

To explain of the modulatory influence described earlier, the predictions that were formed in the Introduction of this experiment will be considered next in relation to the current analyses. It was predicted that an interaction of SND and orthographic neighbourhood size would influence lexical processing via semantic feedback that impacts on word identification before a word is fully identified.

As mentioned in in the previous chapter, the two earliest measures that are associated with lexical identification and that one should focus on in relation to these predictions are first fixation and single fixation duration. The interactive effect was not found to be significant in single fixation duration. This is presumably due to the fact that single fixation duration data are a subset of first fixation duration data, and, thus, the effects observed in single fixation duration tend to be weaker than those observed in first fixation duration. However, next, let us consider the pattern observed in first fixation duration in relation to the prediction. It was found that high orthographic neighbour words with high SND were fixated for the shortest time compared to the other conditions. The reading times for the other conditions were comparable in first fixation duration. The findings also indicated the effects of orthographic neighbourhood size at the low SND levels were less systematic and inhibitory, and these inhibitory effects gradually diminished with increased SND. In contrast, these effects at the higher SND levels were systematic and facilitatory, with decreasing fixation times on high than low orthographic neighbour words.

The inhibitory effects of orthographic neighbourhood size at the low SND levels can be considered to be less robust, compared to the facilitatory effects of orthographic neighbourhood size at the high SND levels, for the following reason. First, the inhibitory effects of orthographic neighbourhood size at the lowest SND levels (Panel A) were noisier (compared to these effects at the high SND levels; Panels D-E) and gradually attenuated to be almost flat at the lower SND levels (Panels B-C). Second, the mean of fixation times indicates that the reading times of the target words with the low SND were quite comparable in first fixation duration (as well as gaze duration).

These findings that occurred for first fixation duration were similar to those that occurred for gaze duration. Recall that gaze duration, also, largely reflects word identification. As is obvious by now, the pattern of interactive effects observed in the present experiment is similar to that observed in Experiment 3. The nature of the interactive effect in both experiments was quite similar in earlier reading time measures and then the pattern of effects began to change in the later measures. That is, SND (as a semantic influence) interacted with word-level variables such as word frequency (Experiment 3) and orthographic neighbourhood size (Experiment 4) in a similar way. The results of both experiments can be explained by Stolz and Besner's (1996) embellished IA model. Under this model, semantic influences can impact on word identification via semantic feedback that is sent to the word level, contributing to resolving the competition between orthographic neighbours, as described in detail in the Introduction of this experiment as well as in previous chapters of this thesis. As such the findings of the present experiment along with those of the previous experiment suggest that word meaning can be activated prior to the perceived word's orthographic representation being uniquely identified.

The comparable reading times for the conditions in first fixation duration can be explained as follows. Though the semantic representation of a high orthographic neighbour word with low SND will be quickly activated, its semantic representation will be weak and, thus, the influence of its semantic representation will be weak in constraining its unique word identification. For a low orthographic neighbour word with high SND, the activation of its semantic representation will be somewhat later compared to a high orthographic neighbour word as discussed in the Introduction of this experiment. As such, the influence of its high SND characteristics will be somewhat later in time. For a low orthographic neighbour word with low SND, the weak semantic representation will have a weak influence on word identification. Given this, it is not surprising that the reading times of these words were found to be comparable.

Finally, let us consider the later measure of regression path duration. The pattern of the interactive effect observed in this regression path duration is quite similar to that observed for first fixation and gaze duration. Just like in first fixation and gaze duration, the reading times for high orthographic neighbour words with high SND were fixated for the shortest time compared to other conditions. However, the reading times for the other conditions were quite different, compared to the earlier measures in which the reading times for these words were found to be comparable. Recall that regression path duration largely reflects post-lexical processes associated with the integration of word meaning into the sentential meaning, along with the construction of a coherent discourse representation. As such, the differences in the reading times that occur for regression path duration can be attributed to the post-lexical influences associated with regression path duration, rather than lexical influences.

The joint effect of target words' SND and their orthographic neighbourhood size can be accommodated by both the E-Z Reader model (e.g., Reichle, Pollatsek, & Rayner, 2006) and the SWIFT model (Engbert *et al*., 2002, 2005). Since this joint effect was found in first fixation duration, then this join effect should occur in the preliminary stage of lexical processing (L1) in which a word is checked for its familiarity. In all the versions of the E-Z Reader model, the durations of L1 and L2 are assumed to be functions of, but not limited to, a word's frequency and its predictability from prior context in text (Reichle, 2011). However, the results of the current experiments reported in this thesis suggest that this assumption is an oversimplification, as it was also noted by Williams, Perea, Pollatsek, and Rayner (2006). To make the E-Z Reader model flexible, Reichle, Tokowicz, Liu, and Perfetti (2011) proposed that the familiarity check stage involves any kind of information that contributes to assessing the familiarity of the fixated word in a way that makes the familiarity check mechanism rapid enough to allow sufficient time for the saccadic programming (125-150ms).

Also, Reichle and Laurent (2006) suggest that the familiarity check process is developed over the course of many years of education and practice of reading. Accordingly, skilled readers have learned to use different types of lexical information that would allow their eye movement system to rapidly or efficiently decide when a saccade will move the eyes from one word to another (see also Reichle & Perfetti, 2003; Reichle, Tokowicz, Liu, & Perfetti, 2011). Having said that, the speed of the familiarity check can be influenced by different lexical variables such as word frequency, orthographic processing of transforming visual features into abstract letter identities, phonological processing, and possibly semantic processing (Vanyukov, Tokowicz, Reichle, & Perfetti, 2011). The findings of the current experiment showed that a rich semantic representation associated with a target word can influence word identification, as evident in first fixation and gaze duration. Thus, the results of this experiment along with the results of the previous experiments reported in this thesis are consistent with the notion that familiarity check is a "heuristic that allows the reader to move his or her eyes efficiently, so as to maximise the overall reading rate while maintaining some minimal level of comprehension" (Reichle, Tokowicz, Liu, & Perfetti, 2011, p. 995). In particular, it may be the case that lexical processing in skilled readers has developed to make use of rich semantic information (e.g., high SND) in lexical processing along with other types of lexical information (e.g., orthographic neighbourhood size, lexical frequency, etc.) to allow them to read efficiently and thereby rapidly move their eyes from one word to another. Although the E-Z Reader model can provide possible explanations for the results of the experiments reported in this thesis, namely, effects of frequency, orthographic neighbourhood size and SND, it is also important to note that the mechanisms underlying such effects are not well specified in any versions of this model. It is also important to note, though, that the E-Z Reader model is a model of eye movement control during reading, not a model of language comprehension.

The present findings of the immediate joint effect SND and orthographic neighbourhood size can also be explained by the SWIFT model (Engbert *et al.*, 2002, 2005). The assumptions of the SWIFT model are mentioned in different parts of this thesis (e.g., Section 3.1.3.2 in Chapter 3, pp. 113-114 and Section 5.1.3.2 in Chapter 5 pp. 169-170). Based on these assumptions, when a word is easy to process, the activation rate associated with it is low. When a word is hard to process, its activation rate is high. Furthermore, an increased activation rate associated with a word within the attentional window serves to inhibit the speed with which a saccade to leave the word occurs. Hence, if the currently fixated word is a high orthographic neighbour word that has a higher frequency than its orthographic neighbours with high SND (i.e., an easy-to-process word), the activation of this word will be lower than the case of a matched low orthographic neighbour word with high SND or a high orthographic neighbour word with low SND (i.e., a difficult-to-process word). Thus, saccades to leave the word will occur more rapidly for high orthographic neighbour, high SND words than for counterpart words with higher activation levels. That is, the random timer will program a saccade to the subsequent word that has the highest activation in the activation field. Accordingly, a high orthographic neighbour word with high SND should be fixated for less time than a low orthographic neighbour word with high SND or a high orthographic neighbour word with low SND.

Finally, the findings that neither SND nor the interaction of SND and orthographic neighbourhood size influenced the skipping probability of the target words is more consistent with the assumptions of the E-Z Reader model rather than the SWIFT model, as discussed in Section 5.1.3.2 in Chapter 5.

Based on the assumption of parallel processing of multiple words at time, the SWIFT model predicts that SND and/ or orthographic neighbourhood size of a yet-to-fixated (adjacent) word could influence the skipping of the adjacent word. There was no evidence in the present experiment that this was case. In contrast, the E-Z Reader model assumes that semantic information about the about the adjacent is not obtained parafoveally, and thus would not influence skipping probability of the adjacent word, meeting the findings of the present experiment in relation to skipping probability.

6.1.3.3 Spillover Effects of SND and Orthographic Neighbourhood Size

This part of analysis examined whether the joint effect of target words' SND and their orthographic neighbourhood size spilled over onto subsequent words. The results showed that the spillover of this joint effect was significant in single fixation and first fixation duration on subsequent words. In particular, it was found that the single fixation and first fixation duration of the subsequent words were shorter following high orthographic neighbour words with high SND than low orthographic neighbour words with high SND.

The spillover of this joint effect found in this experiment can be explained by the E-Z Reader model (e.g., Reichle, & Rayner, 2006; and Pollatsek *et al.*, 2008) and the SWIFT model (Engbert *et al.*, 2002, 2005) in a similar way as the spillover effect of word frequency and SND explained in Section 5.1.3.3 in Chapter 5. A brief explanation of the spillover effects found in this experiment will be considered next (for a detailed account of these effects, the reader can refer to Section 5.1.3.3, pp. 171-172). According to the E-Z Reader model, the joint effects of SND and orthographic neighbourhood size influenced the L2 lexical processing on the target words, given that the joint effect affected the lexical processing of the subsequent words. The L2 processing of a high orthographic neighbour word with high SND will be completed faster than that of the other conditions since this word is considered easier to process compared to the other conditions. This quick completion of the L2 lexical processing of this word will allow more parafoveal preview of the subsequent word (n+1) than the other conditions will. As such, the post-target word will have a head start when it is subsequently fixated following a high orthographic neighbour word with high SND, as a significant amount of processing associated with its familiarity check will have already been carried out while the previous word was fixated.

According to the SWIFT model, if the (fixated) target word is a high orthographic neighbour word with high SND (i.e., an easy-to-process word), then the random timer will not be inhibited by the properties of the fixated (target) word. This is in contrast to the target words that are more difficult to process (e.g., high orthographic neighbour words with low SND, low orthographic neighbour words with either high or low SND) that may inhibit the random timer from executing a forward saccade to the subsequent word. Accordingly, more parafoveal processing of the subsequent word will be carried out while fixating a target word that has high orthographic neighbours and high SND relative to the other conditions. Therefore, subsequent fixation durations on word n+1 would be shorter following a high orthographic neighbour word with high SND than the other conditions.

6.1.4 Conclusion

Experiment 4 demonstrated that the interaction of SND and orthographic neighbourhood size of currently fixated words influenced the lexical processing of the fixated words and subsequent words, as evident in first fixation, gaze duration and regression path duration. The effects of orthographic neighbourhood size at high SND levels were systematically facilitatory in all early and late measures, with decreasing reading times observed for high than low orthographic neighbour words. In contrast, the effects of orthographic neighbourhood size at lower SND levels were less systematic and inhibitory, and this inhibitory effect gradually attenuated with increasing SND levels. In addition, the effect of SND was larger for high than low orthographic neighbour words in these reading measures.

The results of Experiment 4 provide further support for the conclusion formed in the previous chapter that information about a word's SND (i.e., semantic information) can be accessed before full identification of the word, and, thus, can constrain the unique identification of a word's orthographic form. This conclusion is consistent with the idea of enhanced semantic feedback activation assumed in Stolz and Besner's (1996) embellished interactive-activation model (McClelland & Rumelhart, 1981). In addition, the findings of Experiment 4 provided further evidence in support of Shaoul and Westbury's (2010a) conceptualisation of the nature semantic representations.

Chapter 7: Conclusion

The primary goal of this thesis was to address the question of how and when semantic influences emerged in lexical processing during normal silent reading in attempt to inform the development of a comprehensive model of word identification during reading. In this thesis, four experiments have been reported which examined the effects of SND, as defined by the degree of semantic similarity between a word and all its semantic neighbours falling within a specified threshold (Shaoul & Westbury, 2010a). This chapter will discuss the conclusions that can be drawn from the findings of these four experiments. First, the main findings of the experiments will be summarised, and then the theoretical implications of these findings will be discussed. Some directions for future experimental investigations will be suggested before stating a final conclusion at the end of this chapter.

7.1 Summary of the Findings

The first experiment reported in this thesis examined the basic SND effect in lexical processing during reading single sentences. The results showed that increasing SND played a facilitatory role in lexical identification, and the SND effect spilled over onto subsequent words. To explore any modulatory influences to the SND effects, Experiment 2 was conducted as a corpus-based study that passively examined the interaction of SND and the lexical variables that are known to influence lexical processing. The findings of this experiment indicated that there was a significant interaction between SND and word frequency in the early reading measures of single fixation, first fixation and gaze duration.

To actively test the joint effect of SND and word frequency found in the corpus-based study, Experiment 3 was carried out to orthogonally manipulate these two variables. The results showed that the interactive effect of word frequency and SND was significant in all early and late reading time measures. The nature of the joint effect of word frequency and SND on different reading time measures was found to be complex. The SND effects were larger for high than low frequency words in earlier measures of first and single fixation duration, but were larger for low than high frequency words in later measures of regression path duration and total reading time. The visualisations of these interactions showed that this shift occurred due to the counteractive influences of frequency at different levels of SND. In particular, the effects of word frequency for low SND values were inhibitory, whereas these effects for high SND values were facilitatory. The joint effect of word frequency and SND was also found to influence the fixation times on subsequent words. Based on the findings of Experiment 3, it was concluded that SND (i.e., a semantic variable) could constrain unique word identification in normal reading via semantic feedback (assumed in Stolz and Besner's (1996) embellished interactive-activation (IA) model).

To provide further evidence for this conclusion, Experiment 4 was carried out to examine the joint effect of SND and another word-level variable, namely, orthographic neighbourhood size. The results showed that the interactive effect of SND and orthographic neighbourhood size was significant in the early measures of first fixation and gaze duration and the later measure of regression path duration. The effect of SND was larger for high than low orthographic neighbour words in these three reading measures. High orthographic neighbour words with high SND were read for the shortest time compared to the other conditions, and the reading times for the other conditions were quite comparable in the earlier measures of first fixation and gaze duration. However, the reading times for these other conditions were quite different in regression path duration. The visualisations of the interactions showed that effects of orthographic neighbourhood size were inhibitory for low SND values, whereas the effects of orthographic neighbourhood size were facilitatory for high SND values. This joint effect also spilled over onto subsequent words.

Taking the findings of Experiment 3 and 4 together, it is clear that the pattern of the joint effects found in these two experiments were quite comparable. In both experiments, the pattern of the interactive effect was quite similar in earlier reading time measures and then the pattern began to change in the later measures. Such comparison between the findings of the two experiments indicates that SND (i.e., a semantic influence) interacted with word-level variables such as word frequency and orthographic neighbourhood size in a similar way. Thus, both experiments (along with the previous two experiments) consistently suggest that the semantic characteristics of a word (e.g., SND characteristics) can be activated and can influence lexical processing before the completion of unique word identification in normal reading.

Before leaving this section to discuss the theoretical implications of these findings, it is worth mentioning that, in the experiments reported in this thesis, the characteristics of a parafoveal word did not influence the fixation durations on the currently fixated word, providing no evidence for parafoveal-on-foveal effects.

7.2 Theoretical Implications

This section will handle two important theoretical implications of the present findings for current theories of word identification and models of word meaning.

7.2.1 Can Word Meaning be Accessed before the Completion of Unique Word Identification? Or, When Does an Influence of Word Meaning First Emerge in Normal Reading?

Many eye movement studies that investigated semantic plausibility and lexical ambiguity suggest that word meaning can influence lexical identification. However, these studies relied on (sentence) contextual information to arrive at this conclusion. Therefore, it is not clear from these studies whether word meaning can be accessed before the completion of unique word identification. A more direct way to provide evidence for an influence of word meaning in lexical identification is to examine the effects of words' semantic characteristics on their lexical processing. To date, there are only three eye movement studies that paid attention to the effect of the semantic characteristics of a word in lexical processing during normal reading (number of semantic associates: Duñabeitia, Avilés, & Carreiras, 2008; number of semantic features: Cook, Colbert-Getz, & Kircher, 2013; contextual diversity:

Plummer, Perea, & Rayner, 2014). These studies indicate that richer semantic representations (e.g., high number of semantic associates, high number of semantic features, high number of the contexts in which a word appears) facilitate lexical processing in reading.

The present work extends this (modest) line of research by providing evidence for the claim that word meaning, defined by the degree of semantic similarity between a given word and all its semantic neighbours, can constrain unique word identification. Particularly, the present findings indicate that a word that has semantically similar neighbours is processed faster than a word that has semantically less similar (distant) neighbours, as evident in early and late reading time measures. This finding is parallel to the results of Plummer *et al.*'s (2014) study that found that increasing contextual diversity plays a facilitatory role in lexical processing in reading, as evident in early and late reading time measures.

The present findings also offer a characterisation of semantic involvement in lexical identification. The findings suggest that a word's semantic neighbours are more, or less, active during lexical identification in normal reading. As such, these findings theoretically imply that the assumptions about semantic processing are similar to the assumptions about orthographic and phonological processing during lexical identification in reading. To explain, orthographically, phonologically, and semantically similar candidates are activated along with the perceived target word during lexical identification. In the case of this thesis, the findings suggest that a word's orthographic representation (e.g., *MOVIE*) can activate multiple semantic neighbours (e.g., *game*, *theatre*, *cinema*, etc.) within the semantic system during word identification in normal reading. The activation of semantic neighbours feeds back to the word level within the period that the candidate set is being reduced via processes of between-level activation and within-level inhibition. Therefore, it appears that the SND characteristics (i.e., semantic information) of a word can be accessed before the full identification of the word, and, thus, can constrain the unique identification of a word's orthographic form.

The findings of the present work are inconsistent with word identification models that assume that word meaning does not influence lexical processing. For example, serial search models of word recognition (e.g., Forster, 1976) assume that the meaning of a word can be accessed only after the processing of its orthographic (or phonological) representation has been completed, and a unique word has been identified and selected for further processing, as mentioned in Section 1.3 in Chapter 1. To briefly reiterate, under these models, when a target word is perceived, the word is searched for in the orthographic file until a word in the orthographic file matches the perceptual properties of the perceived word. At this point, the search is terminated and a pointer in the orthographic file is used to check the properties of the word in the orthographic file against those in the master file. If the properties of the word in the orthographic file match the properties of the word in the master file, the word is successfully identified. Once a unique word has been identified, semantic information associated with the identified word can then be retrieved.

Thus, serial search models assume that the meaning of a word cannot influence the processing of its orthographic representation. That is, these models predict that SND will influence a stage that is independent of the stage influenced by the word-level variables (e.g., orthographic neighbourhood size and word frequency) in lexical processing (as per Sternberg, 1969). However, there was no evidence in the present findings that this was the case. The findings showed that SND interacted with word-level variables, suggesting that SND and word-level variables influenced a common processing stage.

The findings, on the other hand, are consistent with the interactive models of word identification (e.g., McClelland & Rumelhart, 1981) that assume that activation feeds forward from lower levels to higher levels shortly after processing at the lower levels has begun. Activation also feeds back from the higher levels to the lower levels. According to Stolz and Besner's (1996) embellished interactive-activation (IA) model (McClelland & Rumelhart, 1981), the visual information of a perceived word partially activates the word unit corresponding to the target word and, to a lesser extent, the word units corresponding to orthographically similar words. Shortly after a word's orthographic representation is activated at the word level, activation from the word level feeds forward to the semantic level, activating its semantic representation. The activation of the word's semantic representation feeds back to the word level, contributing to resolving the competition between orthographic competitors (those of the actual word and its orthographic neighbours). Therefore, the embellished IA model predicts that SND and word-level variables can influence a common stage of lexical processing. That is, a word's semantic representation can be accessed and can influence lexical processing before processing at the word level has been completed. This prediction was met by the findings of the significant interactive effects of SND and each of word frequency and orthographic neighbourhood size in the early reading time measures associated with lexical processing.

The current findings of the facilitatory effects of closely packed semantic neighbours (i.e., high SND) is in contrast to the inhibitory effects of close semantic neighbours predicted by the attractor dynamic model as explained in Chapter 2. According to the attractor dynamic model, distant neighbours are far away from the target word, creating a gravitational gradient for faster settling into attractor basins, whilst near neighbours slowed the settling process because their basins of attraction are closer to the target word's basin of attraction. Therefore, this model predicts that inhibitory effects would be observed in the case of words with high SND characteristics (i.e., words more semantically similar neighbours). However, no evidence in the present findings showed that this was the case. In contrast, the present findings consistently showed facilitatory effects of increasing SND, a finding that can be accounted by the embellished IA model. According to the embellished IA model, the closely packed semantic neighbours will provide a greater amount of activation at the semantic level. Therefore, a great amount of activation will feed back from the semantic level to the word level, contributing to resolving the competition between the orthographic neighbours at the word level. As such, high SND can facilitate word identification.

As such, the findings of the present work clearly support a word identification model such as the embellished IA model that assumes that orthographic processing and word meaning processing are cascaded in time. That is, these findings ultimately imply that in order that a lexical identification model provides a comprehensive account of word identification, the model needs to integrate a mechanism by which it explains the influence of the semantic characteristics of words before the completion of unique word identification. This is exactly what the embellished IA model was developed for.

In addition, the present findings indicated that the semantic characteristics (e.g., SND) of a foveal word not only influenced its lexical processing, but also influenced parafoveal processing. Increasing a foveal word's SND was found to provide a greater amount of parafoveal processing of the subsequent words. This finding can be explained by both the serial-attention-shift models of eye movement control such as the E-Z Reader model and the guidance-by-attentional-gradient models such as the SWIFT model. To explain, in terms of the E-Z Reader model, the findings clearly suggest that the fixated word's SND characteristics (i.e., semantic characteristics) can influence both stages of lexical processing: the early stage of lexical processing (L1) in which the fixated word is assessed for its familiarity and the later stage of lexical processing (L2) in in which the meaning of the fixated word is processed and integrated into the sentential context. In terms of the SWIFT model, the findings indicate the SND characteristics can impact on executing a forward saccade by influencing the random timer. However, there was no evidence in the results of the experiments reported in this thesis indicating that the SND (i.e., semantic information) of a parafoveal word could influence the lexical processing of the foveal word. This finding may be considered to be inconsistent with the assumptions of the parallel processing models (e.g., the SWIFT model) that argue that multiple words are processed at a time. In contrast, this finding is in line with the serial processing models (e.g., the E-Z Reader model) that assume that only one word is processed at a time and that parafoveal processing is carried out after the lexical processing of the foveal word has been completed.

The findings of the present study also have important implications for the computational framework of eye movement control. The models such as the E-Z Reader model and the SWIFT model can account for some basic cognitive influences on eye movements during normal reading, and they can explain the influences found in the present study as outlined in the previous experimental chapters in this thesis. These models specify some mathematical equations that determine some basic effects such as word frequency and predictability effects in lexical processing. However, our increasing understanding of word identification places increasing challenges on these models to provide a more sophisticated account or some equations that determine the influence of other variables that have been shown to impact on lexical processing. It is acknowledged here that these models, after all, were developed to account for eye movement control during reading, rather than to specify the nature of lexical processing and language comprehension.

To sum up, the findings of the experiments reported in this thesis clearly indicate that word meaning can be accessed and can influence lexical processing before the completion of unique word identification. These findings are inconsistent with word identification models that assume that the meaning of a perceived word is activated only after its form has been uniquely identified. In contrast, the present findings are consistent with the assumptions of interactive models of word identification that assume that semantic feedback can contribute to resolving the competition between orthographic competitors, thus, can constrain unique word identification. Eye movement control models can potentially explain the influence of the semantic characteristics of a foveal word in its lexical identification; however, there are not yet any mathematical equations that can determine such influences.

7.2.2 To What Extent Can Distributional Semantic Models Capture Informative Aspects of Meaning?

Another fundamental question that remains to be addressed here concerns the credibility of the present findings in terms of informing our understanding of the influence of word meaning in lexical identification during normal reading. Recall that semantic representations in the present thesis were defined in terms of semantic distributional models. To remind the reader, distributional semantic models are theoretically based on the distributional hypothesis that postulates that two words that occur in similar contexts are considered semantically similar (as discussed in Chapter 2). Distributional semantic models have been used to model the meanings of words by analysing and comparing their distributional profiles in large-scale corpora of text. The statistical distributions of words in text delineate some important aspects of their meaning. This section will give a general discussion on the capability of distributional semantic models to capture informative aspects of word meaning.

Distributional semantic models can potentially extract semantic information from linguistic contexts (i.e., text data). Humans also learn word meaning from extralinguistic contexts based on their perception and interactions with the objects in the world (de Vega, Graesser, & Glenberg, 2008). In other words, many words are learned from past perceptual experiences (or circumstances) in which the words were uttered (McRae *et al.*, 2005). Therefore, both linguistic information and sensorimotor information constitute important aspects of words' semantic representations. Indeed, the results of many studies suggest that language captures and encodes much perceptual information and that linguistic contexts can be used to extract referential word meaning (Connell & Ramscar, 2001; Durda & Buchanan, 2008; Kintsch, 2007, 2008; Louwerse & Zwaan, 2009; Riordan & Jones, 2010). For example, Riordan and Jones found that though distributional and featural models tended to emphasise different aspects of word meaning, these two types of semantic models encoded much redundant information about word meaning. Their findings also suggest that children rely on perceptual cues about the referents of the words at the early stages of learning their first language. As they gain more perceptual information, children tend to rely on statistical cues (i.e., the distribution of words in the language) to develop and refine semantic similarity relations between words. Thus, such findings clearly indicate that co-occurrence vectors obtained from distributional semantic models contain an amount of information about the words they represent, including perceptually grounded information (Shaoul & Westbury, 2012).

The notion of semantic similarity used in distributional semantic models, though it may seem broad, has been shown to be psychologically plausible. To explain, human subjects appeared to understand the concept of semantic similarity when they are instructed to make judgments about the semantic similarity of word pairs (e.g., Miller & Charles, 1991). Moreover, many researchers have demonstrated that the participants' agreement (inter-subject agreement) about the semantic similarity of word pairs is very high (e.g., Rubenstein & Goodenough, 1965; Miller & Charles, 1991). Given that researchers in the area of word identification are interested in investigating the psychological phenomena occurring in word identification, the notion of semantic similarity in its broad meaning does not need to be further specified in terms of conventional semantic relations (e.g., synonyms, antonyms, hyponyms, etc.).

Recall that most distributional semantic models such as HiDEx (Shaoul & Westbury, 2010a) were built to account for the meaning of individual words. The interest of the present study was on the effect of the semantic characteristics of individual words in lexical identification in normal reading. As such, operationalising semantic representations in terms of such distributional semantic models is sufficient for the purpose of the present thesis. At this point, it fair to introduce an area that has recently received a great deal of attention in distributional semantics, namely, modelling the meaning of whole sentences. There have been some recent attempts to combine the vectors representing the meaning of individual words in a sentence together so there is a single vector representing the meaning of the whole sentence. To determine the meaning of a sentence, distributional semantic models should account for the meaning of the sentence's parts (i.e., individual words) as well as how these parts are combined (i.e., syntactic structure) (Partee, ter Meulrn, & Well, 1990).

Encoding the syntactic structure of linguistic expressions into distributional representations has been one of the recent interests of the distributional semantic models. For instance, Landauer and Dumais (1997) summed the vectors produced by their model to arrive at the compositional meaning of simple sentences. However, their model did not perform satisfactorily as it could not differentiate between *who* did what to *whom* in a sentence like '*the man bit the dog'*. Kintsch (2001) also used summed vectors in a way that this summing was sensitive to the sentential context, and their model was successful in distinguishing between literal meaning (e.g., *this fish is a shark*) and metaphorical meaning (e.g., *this lawyer is a shark*). Other researchers also attempted to account for semantic composition by multiplying word vectors (e.g., Mitchell & Lapata, 2008) so that their model could distinguish between different word (polysemous) senses (e.g., *mouse* as an animal or as a device). Other researchers (e.g., Jones & Mewhart, 2007) managed to develop distributional semantic models that could track the order of words in sentences of their corpus. Although vector addition or multiplication cannot yet capture full aspects of semantic compositionality, implementing other complex compositionality functions (e.g., those used by Kintsch and Jones and

Mewhart) shows that distributional semantic models can potentially account for polysemous senses of words (e.g., *mouse*) and can differentiate word order. Such attempts are encouraging in that it seems likely that distributional semantics will be able to account for compositionality to a sufficient degree in the future.

To sum up, distributional semantic models were found to successfully handle a variety of semantic tasks, which highlights the importance of considering distributional data in modelling word meaning and the capability of this research field. With more sophisticated computations and algorithms that take syntactic structure of sentences into account, distributional semantic models are expected to evolve over time to efficiently account for the compositional meaning of sentences. Of more theoretical interest to this thesis, distributional semantic models can potentially capture informative aspects of word meaning (both linguistic and referential meanings), and the broad notion of semantic similarity adopted in the these models has been shown to be psychologically plausible. Accordingly, the present findings reported in this thesis can be considered credible in that the metric of semantic similarity used in the present thesis can capture informative aspects of word meaning. Thus, the results based on the metric in this study may well inform our understanding of the involvement of semantic representations in lexical identification in normal reading.

7.3 Recommendations for Future Research

The four experiments in this thesis observed reduced fixation times on target words with high SND characteristics and reduced fixation times on subsequent words following words with high SND. These findings were taken as evidence for the assumption that rich semantic representations of words can potentially contribute to the ease with which words are lexically identified. However, more data are required to provide further support for this assumption. Future studies are recommended to further examine the interplay between semantics and orthography, and/or phonology in lexical processing in normal reading. For example, future studies can investigate the interaction between SND and other word-level variables or the interaction between SND and phonological regularity or homophony. It will be also interesting to examine the interaction between SND and lexical ambiguity (balanced vs. unbalanced ambiguity). Based on the conclusion derived from the findings of the present experiment, it is predicted that high SND would contribute to the ease with which phonologically irregular words and balanced ambiguous words are identified.

Recall that the target words of the experiments reported in this thesis were concrete nouns. As such, the findings are true to the concrete nouns, and the SND effects in the lexical processing of other types of words are not yet known. Given this, future research is recommended to be carried out to examine the influence of SND in processing other types of words such as verbs and adjectives. In addition, examining the influence of the characteristics of a parafoveal word in the lexical processing of the currently fixated word indicated that the SND characteristics of a word are not extracted parafoveally. However, this finding was based on the current experiments that used normal reading of sentences. To provide strong evidence for whether information about SND (i.e., semantic characteristics) can be extracted parafoveally, future research is recommended to be carried out using the boundary technique. In this technique, the target word that appears in the parafoveal region changes to another word as the reader moves the eyes to fixate it. The target words in this proposed research could be high and low SND words.

The results of such proposed future experimental investigations would assess whether the conclusion made based on the findings of the present experiments in this thesis is plausible. Thus, the results of such future studies along with the results of the present study will be a step forward towards developing a comprehensive model of word identification.

7.4 Conclusion

The findings of the experiments in this thesis demonstrate that SND (i.e., a semantic variable) plays a role in the lexical processing of the fixated word and subsequent words, as evident in early reading time measures as well as late measures. The findings can be simply explained by the notion of semantic feedback assumed in Stolz and Besner's (1996) embellished interactive-activation model of lexical identification (McClelland & Rumelhart, 1981). Both the E-Z Reader model and the SWIFT model of eye movement control during reading can account for the present findings of the SND effects on the fixation times of the target words and subsequent words. In addition, the findings also provide evidence for the psychological validity of the corpus-based distributional semantic similarity measure (Average Radius of Co-occurrence; ARC) developed by Shaoul and Westbury (2010a) as capturing some informative aspects of word meaning. Based on the current investigation, it is clear that the semantic characteristics of a currently fixated word can be activated and can influence lexical processing before the completion of unique word identification. Accordingly, a comprehensive model of word identification (and models of eye movement control during reading) should consider providing a mechanism by which it explains how the meaning of a word can influence the unique identification of its orthographic (or phonological) form.

Appendices

Appendix A

Experiment 1 materials (high semantic neighbour words and low semantic neighbours are underlined)

The girl bought the green peppers/fleece from the market.
She lost the new socket/coupon she bought yesterday.
She put her pink badge/scarf on the desk.
She commented on the new puzzle/drawers he bought.
He saw a green snail/melon in the garden.
He said the injured turtle/robber was recovering very well.
She found a brown snake/towel near her bed.
They saw the nice teens/flask in the kitchen.
He pointed at the small dove/tart in the magazine.
They saw a thin worm/vet in the veterinary clinic.
It was the grey wolf/yarn that she was interested in.
She looked at the white pearl/pillow in the shop.
She changed the old tyres/fridge three days ago.
She had a good snack/pouch inside her bag.
She inspected the small scar/buds with her eyes.
She showed him the colourful tattoo/pebble she had.
They threw the unwanted plum/rind in the kitchen pin.
She took the yellow peach/wallet that was on the table.
He read about the escaped crocodile /hooligan in the newspaper.
He cleaned the old tomb/stove with a piece of cloth.
She dropped the orange soda/ twigs she was holding.
She held the green lizard/napkin up tightly.

Appendix B

Experiment 2 materials (High semantic neighbour words and low semantic neighbour words are underlined)

Passage 1

Jane, a waitress and a media studies student, suffered from zinc deficiency that affected some of her glands, and which was manifested in her suffering from obesity. In her hectic lifestyle, she was busy writing critiques on some plays for the module she was taking, and didn't have the time to lose the extra kilos she had gained and remained inactive in this regard. As part of her course, she received a voucher to see a film that tells a convoluted story about a falconer who had to testify against his cadet friend who had got out of an abusive relationship. In the film there was a lawsuit over a crime the cadet had committed when on vacation in a village near a dam in the North.

Jane hated leaving her baby daughter with her mother because she was in the process of weaning her. However, the film was important. She ironed her blouse that matched her knitted skirt and set off to the cinema. She liked the shark fin that the cinema had as its logo; she pressed the zoom function on her phone to take a picture of it. The cinema usher came over to verify her ticket and show her to her seat. She waited in the gloom holding a leaflet about the upcoming film, and enjoyed the homely feel of the cinema theatre. She looked at the painting featuring an oasis and a raft in pastel in a cedar frame that was hung on the wall nearby. She also noticed the emergency exit in case there was a need to evacuate the cinema.

On the floor there were some scraps of paper left by children earlier. Jane liked the neon red varnish on her nails, and the jewels on her wrist caught her eye. Suddenly, it clicked in her mind that she had forgotten about the dessert she planned to make. She quickly texted her mother to ask her to prepare the yeast mixture and to thaw the butter that she'd need for the nutty apple crumble and the cinnamon paste she would need. She jerked her chin up and yawned loudly, forgetting that this behaviour was taboo in public company.

Thinking about her course, Jane became rattled by the poor mark she had received for her course work. The strict tutor deducted several marks because of her tenuous arguments and her repeated use of an inappropriate idiom. Jane was frustrated that she didn't reap the rewards she felt she was due. The more she thought about this, the more she felt unwell, and this worsened as she sat waiting. Suddenly, the adverts before the film started. The first starred a very obedient dog and a not so obedient cat that gave his owner a nip on the left ankle. Bizarrely, it was for car insurance. The second featured a sluggish black beetle that could speak and spent its time moaning about things. Jane wondered where the advertising agencies came up with these ideas.

After the adverts, the film came on. The opening sequence involved a medical scene in which a doctor used a needle to lance a swollen abscess on the cadet's leg. He also applied a bandage to bruising on the cadet's arm to prevent him scratching it on a whim. Afterwards, the cadet gave a salute and the doctor clapped him on the shoulder.

Appendix B

Passage 2

Sitting on a large boulder on the beach with a book and lantern next to him, James looked at the placid water. He was yearning for the good old days with his son who was abducted five years ago. The news of the abduction was wired on TV. He started reading the book that was about using fibre optics to characterise pigment mixtures in the paints used in arts. However, he could not concentrate and quit reading the book. He stared at the lantern, and memories came flashing back. He recalled rambling with his son through the woods at night. Their favourite game was hiding behind the shrubs and using their lanterns to send signals to each other of where they were.

Then, James lay down on the beach and closed his eyes. He recalled how his son wept when he was given his first vaccine, and how lovely their yearly visit to the zoo was. James smiled remembering the penalty kick that his son dispatched on his first football match. Then, he recalled the good times they spent together changing their bike pedals after the breakage to make them usable again. Their adventures in the countryside were also unforgettable. They mingled and giggled with the local people, and then jogged out on cold winter nights until their hands went numb. His son hurt his pelvic bone after falling off the bike in one of their adventures. They used to stop at a gas station on their way back home; his son fetched soda for him and used a vending machine to dispense some chocolate bars.

Food also evoked some unforgettable memories. His son liked to have roasted chicken with zest of lemon and cucumber pickle every time they dined out. On a lazy evening, they would be satisfied with anything edible for their dinner, such as heaps of crisps and some canned foods. Their attempt to make a cake batter was always a failure; sometimes the batter came with lumps and other times it was too runny. Fridays and Saturdays used to be different with his son. His son liked to wear overalls on Friday nights; he looked so graceful in such outfits. James loved their Friday nights with their kidding and idle talks! They used to have a science quiz with timed answers on Saturday evenings. James sometimes teased his son with some questions and his son would not produce a single utterance afterwards.

Suddenly, James shrugged when he remembered his cheating wife. He hated how she used to be emphatic about being tidy. After their divorce, James quarrelled with his son not to keep her mother's things in the house. He remembered how his son wept when James threw out all his mother's favourite things such as a jasmine and lilac-scented candles, the white quilt, the big fridge magnet, and a silver pendant that she left in the house. His son hid his mother's harp and kept it secretly in the attic. James's divorce made him neglect many aspects of his life. He did not bother to pay his car insurance, and as a result, his car was clamped many times. He also made himself so drunk; he became immobile and had to vomit afterwards. However, he managed to get over it. He abstained from alcohol to avoid being ineligible for the custody of his son.

Having all these fond memories of his son, James did not give up. He was determined to reunite with his son. He left the beach, and jogged to his house to use a scanner to scan the only picture of his son he had. After scanning the picture, he filtered it using image software so that the picture looked clearer after the scan. He, then, logged into his Facebook account to navigate through his son's Facebook page and upload his picture and an advert about his abducted son.

Passage 3

Dan and his girlfriend, Jennifer, had an esoteric interest in sky and comet watching. On a rainy day, Dan took her to the desert that was within 100 miles radius of their place to watch the upcoming lunar eclipse as a belated birthday gift. This lightened her heart and cheered her up. She was also happy that her college agreed to waive her tuition fees for the first year. So, Jennifer decided to bake some cookies and pies in the afternoon before their trip.

Using maize starch, Jennifer prepared some oozing cherry pies, bubbling, yummy mince pies and tasty orange and ginger biscuits. Dan asked her to drizzle some caramel over his mince pies. He was so impressed that she could bake all these goodies unaided. She always preferred baking to buying packaged desserts that were full of additives. Before going out, she pampered herself, and rubbed some insect repellent on her arms against any insect bites. She wore a purple cardigan, a green blazer and gaudy neon sandals for the trip, while he wore a thermal coat.

Jennifer hated Dan's almost obsolete car and its chassis that needed to be replaced. However, she liked that he knew how to keep the car clean. He learned how to utilize a nozzle to clean the dust that would obstruct the ventilation duct of the car. On their way to the car, Jennifer looked at Dan and commented that he looked tired and had some wrinkles around his eyes. Indeed, Dan wanted to tell Jennifer about his decision to terminate his work contract. Dan turned on the radio to listen to the news. The first thing they heard was about a man suffering from sickle cell anaemia who was beheaded in his backyard, and another man who nursed a toad back to health after a car accident.

Jennifer found the news disgusting and turned off the radio. She told him that she bought a cactus for herself. After five minutes of silence, Jennifer talked about the increased affluence of the working class and how wage growth was hampered by inflation. Then, she told Dan about the fidelity of her righteous grandfather, the benefit of brushing teeth with coconut oil and the side effects to sniffing baking soda. She, then, offered him some tips on how to wrestle the feeling of inferiority among his colleagues at work. Dan felt that Jennifer was lecturing him and he started humming to try to tell her that he was fed up with her talk. As time elapsed, he became even more annoyed that she was so talkative.

Despite this, Dan tried to adhere to one thing: always reassuring Jennifer of his love. After two hours of her continuous talk, Jennifer felt feeble and thirsty. She asked Dan to stop to get some water from the nearest gas station. After having some water, she felt vibrant and ignited with enthusiasm for the eclipse they were going to witness that night. When they arrived, they found their sculptor friend and other people in the desert gathered to watch the event and they teamed up with them. The clamour of the crowd when the eclipse took place made Jennifer very pleased.

Appendix C

Experiment 3 materials (SND * word frequency)

HSND-HF-------HSND-LF-------LSND-HF------LSND-LF

Carpet----------**tattoo**------------**blouse**----------**napkin**

She had a blue ------------ that I liked.

Jenny pointed to the pale green----------- she had just chosen.

Mary had an expensive---------- from that shop.

I saw an oriental ----------- in the magazine.

Knife-------**banjo**-------------**apron**------**anvil**

She put the black------------ on the table.

I purchased an expensive ------------- from the shop.

Jenny threw away the damaged ------------ after using it for years.

I liked the expensive---------- I bought yesterday.

Camera ----------**shrimp**------------- **coupon**------ **tulip.**

I remembered seeing the large--------- in the kitchen.

Nancy bought the pink ----- from the local store.

Peter threw out the large---------- in the rubbish.

I forgot to pick up the bright pink ------------- from the shop.

Cabinet--------- **furnace**--------- **drawers**---------**griddle**

The children threw stones at the unwanted ------- in my backyard.

My mother finally found the heavy ---------- she'd been looking for.

Appendix C

She liked the expensive ----------- she saw in the shop.

It was the large --------------that required cleaning.

Balloon-------- **gorilla** --------- **cushion**------**spatula**

I liked the enormous ------- very much.

It was the black ------------ that I liked the most.

She looked at the large---------------- in the magazine.

James thought the black ------- was great.

Blanket-------**lantern**-----------**oranges**-----------**avocado**

I liked the large ------------ that I bought yesterday.

The child liked the pale green------------ he saw in the shop.

Peter added the small -------- to his shopping list.

It was the enormous ----------- that I wanted to buy.

Penguin -------**sparrow**-----------**cabbage**--------**cheetah**

I saw a huge-----------on TV yesterday.

They were surprised to see a wild ------------ in their garden.

Peter liked the little ------------ he saw in the book.

I finally found information about the African--------- on the Internet.

Peacock------**panther**---------**lettuce**------**rhubarb**

I read about the Australian ------------- on the Internet.

The child drew a picture of the pink------------ in his sketchbook.

I found some information about the Australian --------- in this book.

The child painted a beautiful --------------- in pink.

Umbrella-------- **trombone**-------- **armchair**------**catapult**

He wanted to buy a brown----------- from the shop.

My father gave his favourite ----------- to me.

I no longer need this brightly coloured -------------- any more.

She threw out the broken -------------- last week.

hammer--------**turtle**--------**carrot**-------**radish**

She took a picture of a brightly coloured ----------- I saw yesterday.

I bought this small---------- from our local shop.

There was a huge ---------- just over there.

I couldn't help but notice the large--------- in the kitchen.

Saddle-----**muzzle**--------**pillow**-----**funnel**

The man had owned the old brown------ for many years.

James kept the filthy ------------- in the garage.

I don't need the old green ------------ any longer.

She noticed the large-------------- almost immediately.

Whistle------**scooter**-------**shelves**------**toaster**

Grace purchased the yellow ---------- last weekend.

I bought the expensive ------------ from the designer store last week.

I liked her large------------- very much.

I found this yellow ----------------- in the garage of our new house.

Appendix D

Experiment 4 materials (SND * orthographic neighbours)

HSND-HON---------HSND-LON---------LSND-HON---------LSND-LON

Spoon------flute-------------buggy ---------duvet

Jane bought that expensive---------- from this shop.

She looked at the little --------------- in the catalogue.

I bought this large ------------at a discounted price.

She threw out the white-------------- she had owned for many years.

Bolts--------badge--------scarf---------------beret

He chose the large---------- he saw in the shop.

He put the silver ------------ on the table.

She left the small ------------ in the cupboard.

I mistakenly left the silver---------- at my friend's house.

Flyer------pearl-----------cuffs----------sieve

He took a picture of the white ----------- he saw in the shop.

I saw the large---------- you mentioned to me yesterday.

She thought the greyish white---------- looked dirty.

I liked the large---------- I saw in the magazine.

Porch----pizza------globe-----------melon

I was satisfied with the small----------- that I had.

She thought the large---------- looked ugly.

He collected some pictures of a small ----------- from the Internet.

In the end, I settled for a small---------- after all.

Barrel------ pistol----------- liners--------pebble

I admired the reddish-brown --------- I saw.

She thought that the white--------- looked terrific.

The child pointed at the enormous ----------- in the magazine.

James posted a picture of a black---------- on Facebook.

Hoover------cereal-----------wallet---------blazer

Fiona helped me to pick the right----- while we were shopping.

I didn't know which ----------- to buy.

He bought this lovely------------ last weekend.

I didn't know which------------- she would like.

Puzzles--------- scanner----------- sweater----------- freezer

I donated my expensive ------------ that I bought three years ago.

I bought the large-------------- five years ago.

Michael added the grey----------- to his shopping list.

I didn't find the small----------- that I was looking for.

Witches-------tractor--------sticker----------plumber

I noticed the unattractive ------------ while reading the novel.

Jane ripped the page with the big ------- in the middle of it.

I finally found the nice -------- I was looking for.

I saw a picture of the unattractive ----------on my sister's phone.

Folder----------ribbon--------camper----------brooch

Lisa liked the large---------- very much.

She got rid of the old red-----------she'd had for several years.

I liked my sister's blue-------- that she bought recently.

Jenny has a little--------------- that looks so cute.

Bullet------- mustard--------platter-----avocado

The man put the odd-looking -------- on the table.

She put the extraordinary----------- in the kitchen cupboard.

Ben bought the cheap ----------- from that shop.

She saw the strange --------------- in the magazine.

Bibliography

Ahn, W. K., Marsh, J. K., Luhmann, C. C., & Lee, K. (2002). Effect of theory-based feature correlations on typicality judgments. *Memory & Cognition*, *30*, 107-118.

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Pschophysics, 63*, 875-890.

Altarriba, J., Kroll, J. F., Scholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye-fixations and reading times. *Memory & Cognition, 24,* 477-492.

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

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 234-254.

Andrews, S. (1997). The effects of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychological Bulletin & Review, 4*, 439-461.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16,* 285-311.

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n+2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1210–1220.

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition, 16,* 697–707.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 416- 424.

Baayen, R. H. (2001). *Word frequency distributions*. Boston, MA: Kluwer Academic.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. F. d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 9-32). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Balota, D. A., Ferraro, F. R., & Connor. L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187-222). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal information in reading. *Cognitive Psychology, 17*, 364-390.

Balota, D. A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 198-232). Lawrence Erlbaum Associates, Hillsdale, NJ.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–284). Oxford, UK: Oxford University Press.

Bates, D., & Sakar, D. (2008). lme4: Linear mixed-effect models using S4 classes, (Version 0.999375-15).

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from reading short and long Finnish compounds. *Journal of Memory and Language, 48*, 615-634.

Besner, D., Coltheart, M., & Davelaar, E. (1984). Basic Processes in reading: Computation of abstract letter identities. *Canadian Journal of Psychology, 38,* 126-134.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience, 17*, 905-917.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49*, 1583–1591.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53-77). Oxford: Oxford University Press.

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.

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review, 8*(3), 531-544.

Bugress, C., & Lund, K. (2000). The dynamics of meaning in memory. In E. Dietrich & A. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in human and machines* (pp. 117-156). Hillsdale, NL: Lawrence Erlbaum Associates.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, *39*, 510-526.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Academic Press.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons*. Journal of Experimental Psychology: Learning, Memory, & Cognition, 23*, 857-871.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. Van Gompel, M. Fisher, W. Murray & R. L. Hill (Eds.), *Eye movement research: A window on mind and brain* (pp. 341-372). Oxford: Elsevier Ltd.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*(3), 249–253.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15*, 245–286.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & performance* (Vol.6, pp. 535-555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Connell, L., & Ramscar, M. (2001). Using distributional measures to model typicality in categorization. In J. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference the Cognitive Science Society* (pp. 226-231). Mahwah, NJ: Erlbaum.

Cook, A. E., Colbert-Getz, J., & Kircher, J. C. (2013). Number-of-features effect during reading: Evidence from eye movements. *Discourse Process, 50*(3), 210-225.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163-201.

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science, 23*, 371-414.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*, 25-62.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic characteristics. *Behavior Research Methods, 37*(1), 65-70.

de Vega, M., Glenberg, A. M., & Graesser, A. C. (Eds.). (2008). *Symbols and embodiment: Debates on meaning and cognition*. Oxford, UK: Oxford University Press.

de Vega, M., Graesser, A. C., & Glenberg, A. M. (2008). Reflecting on the debate. In M. de Vega, A. M. Glenberg & A. C. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 397-440). Oxford, UK: Oxford University Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321.

Deubel, H., O'Regan, K., & Radach, R. (2000). Attention, information processing and eye movement control. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as perceptual process* (pp.355-374). Oxford, UK: Elsevier.

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.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 954-969.

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects during reading. *Quarterly Journal of Experimental Psychology, 61*, 1239-1249.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory & Language*, *27*, 429-446.

Duñabeitia, J. A., Avilés, A., & Carreiras, M. (2008). NoA's ark: Influence of number of associates in visual word recognition. *Psychonomic Bulletin & Review, 15*(6), 1072-1077.

Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods, 40*(3), 705-712.

Durda, K., Buchanan, L., & Caron, R. (2006). WordMine2 [Online]. Available at www.wordmine2.org.

Ehrlich, S., & 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., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777-813.

Evert, S., & Lenci, A. (2009, July 30). Geometric methods in vector space: Distributional semantic models [PDF file]. The *European Summer School in*

Logic, Language and Information (ESSLLI). Lecture conducted from the Association for Logic, Language and Information (FoLLI), Bordeaux. Retrieved from http://wordspace.collocations.de/lib/exe/fetch.php/course:esslli2009:04_geometric_methods.slides.pdf

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

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition, 106*, 1038-1046.

Firth, J. R. (1957). A synopsis of linguistic theory 1930-1955. In *Studies in linguistic analysis* (pp. 1-32). Oxford: Philological Society. Reprinted in F.R. Palmer (Ed.), *Selected Papers of J.R. Firth 1952-1959*, London: Longman (1968).

Fisher, C., & Gleitman, L. R. (2002). Language acquisition. In H. F. Pashler (Series Ed.) & C. R. Gallistel (Volume Ed.), *Stevens' handbook of experimental psychology: Learning and motivation* (Vol.1, pp. 445-496). New York: Wiley.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walk (Eds.), *New approaches to language mechanism* (pp. 257- 287). Amsterdam: Elsevier.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory effects in lexical decision and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 696-713.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178- 210.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*(3), 518-565.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102*, 167-181.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Harris, Z. (1954). Distributional structure. *Word*, *10*(23), 146-162.

Harris, Z. (1970). Distributional structure. In Z. Harris (Ed.), *Papers in structural and transformational linguistics: Formal linguistics* (pp. 775–794). New York: Humanities Press.

Henderson, J. M., Dixon, P., Petersen, 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.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 417-429.

Henderson, J. M., & Ferreira, F. (1993). Eye movement control in reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Psychology: Special Issue on Reading and Language Processing, 47*, 201-221.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54*, 425–443.

Holyoak, K. J., & Spellman, B. A. (1993). Thinking. *Annual Review of Psychology*, *44*, 265-315.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*(3), 427-466.

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of non-fixated words influence the processing of non-fixated words during reading? *European Journal of Cognitive Psychology, 16*, 104-127.

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.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 1612-1627.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Leaning 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., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 20–34.

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, driving and scene perception* (pp. 29-53). New York: Elsevier.

Inhoff, A. W., Radach, R., Eiter, B. M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 56*, 803–827.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431- 439.

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics, 62*, 1474-1484.

Inhoff, A. W., & Topolski, R. (1994). Use of phonological codes during eye fixations in reading and on-line and delyaed naming tasks*. Journal of Memory and Language, 33,* 689-713.

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, *36,* 1-27.

Irwin, D. E., & Carlson-Radvansky, L. A. (1996). Suppression of cognitive activity during saccadic eye movements. *Psychological Science*, 7, 83-88.

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 209–229.

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114*, 1-37.

Juhasz, B. (2008). The processing of compound words in English: Effects of word length on eye movements during reading. *Language and Cognitive Processes, 23,* 1057-1088.

Juhasz, B., Starr, M., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions, and eye fixations. *British Journal of Psychology, 94,* 223-244.

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1560-1579.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675-687.

Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics, 66*, 279-292.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A*, 429-455.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A*(4), 1307–1337.

Kennison, S. M., & Clifton, C., Jr. (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.

Kintsch, W. (2001) Predication. *Cognitive Science, 25*, 173-202.

Kintsch, W. (2007). Meaning in context. In T. K. Landauer, D. S. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 89-105). Mahwah, NJ: Erlbaum.

Kintsch, W. (2008). Symbol systems and perceptual representations. In M. de Vega, A. Glenberg & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 145-163). New York: Oxford University Press.

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,* 13-35.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+2. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1250-1255.

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combing data from eye movement corpora and megastudies. Q*uarterly Journal of Experimental Psychology, 66*(3), 563- 580.

Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child.* Cambridge, MA: Harvard University Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-140.

Lavidor, M., & Ellis, A. W. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language, 80*, 45-62.

Lee, Y. A., Binder, K. S., Kim, J., Pollatsek, A., & Rayner, K. (1999). Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 948-964.

Lee, Y. A., Rayner, K., & Pollatsek, A. (1999). The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast priming technique. *Psychonomic Bulletin & Review, 5,* 624- 634.

Levy, J. P., Bullinaria, J. A., & Patel, M. (1998). Explorations in the derivation of semantic representations from word co-occurrence statistics. *South Pacific Journal of Psychology, 10*, 99-111.

Lima, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language, 26*, 84-99.

Lima, S. D., & Pollatsek, A. (1983). Lexical access via an orthographic code? The Basic Orthographic Syllabic Structure (BOSS) reconsidered. *Journal of Verbal Learning and Verbal Behavior*, *22,* 310-332.

Lin, E. L., & Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1153-1169.

Liversedge S. P., & Findlay J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4,* 6–14.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44*(10), 1013-1024.

Locker, L. Jr., Simpson, G. B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of polysemous words. *Memory & Cognition*, *31*, 505-515.

Louwerse, M., & Zwaan, R., (2009). Language encodes geographical information. *Cognitive Science, 33*, 51-73.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instrumentation, and Computers, 28,* 203-208.

Lund, K., & Burgess, C., & Audet, C. (1996). Dissociating semantic and associative word relationships using high dimensional semantic space. In G. W. Cottrell. (Ed.), *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 603- 608). Mahwah, NJ: Lawrence Erlbaum Associates.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19–40.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*, 102-105.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category-specificity and the brain: The sensory-motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1023–1036). Cambridge, MA: MIT Press.

Masserang, K. M., & Pollatsek, A. (2012). Transposed letter effects in prefixed words: Implications for morphological decomposition. *Journal of Cognitive Psychology, 24*(4), 478- 495.

Masserang, K. M., Pollatsek, A., & Rayner, K. (2009). No morphological decomposition with parafoveal perviews. In S. P. Liversedge (Ed.). Abstracts of the Fifteenth European Conference on Eye Movements (ECEM15), Southampton, August 24 - 26, 2009. *Journal of Eye Movement Research, 3*, special issue, 1-203.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1: An account of basic findings. *Psychological Review*, *88*, 375-407.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye-movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28*(10), 1107-1118.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics*, *46*, 245–253.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*, 578–586.

McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 137–162). Newark, DE: International Reading Association.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is obtained only from the saccade goal. *Vision Research, 46*, 4416-4424.

McDonald, S., & Ramscar, M. (2001). Testing the distributional hypothesis: The influence of context on judgments of semantic similarity. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 611–616). University of Edinburgh.

McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor network. In B. H. Ross (Ed.), *The Psychology of learning and motivation: Advances in research and theory* (Vol. 45, pp. 41-86). San Diego, CA: Academic Press.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*, 547-559.

McRae, K., Cree, G. S., Westmacott, R., & de Sa, V. R. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology, 53*, 360-373.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*, 99-130.

McRae, K., & Jones, M. N. (2013). Semantic memory. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 206-219). Oxford, UK: Oxford University Press.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469-1481.

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(*6), 721-728.

Miller, G., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes, 6*, 1-28.

Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, *34*, 65-79.

Mitchel, J., & Lapata, M. (2008). Vector-based models of semantic composition. *Proceedings of the 46th Annual Meeting of Association for Computational Linguistics: Human Language Technologies* (pp. 236-344). Columbus, Ohio, USA.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 667-682.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review 76*(2), 165-178.

Moss, H. E., & Tyler, L. K. (1995). Investigating semantic memory impairments: The contribution of semantic priming. *Memory, 3*, 359-395.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289-316.

Murray, W. S. (1998). Parafoveal pragmatics. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp.181–199). Oxford: Elsevier.

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging, 7*, 257-270.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Newton P. K., & Barry, C. (1997). Concreteness effects in word production but not word comprehension in deep dyslexia. *Cognitive Neuropsychology, 14*, 481-509.

Niswander-Klement, E., & Pollatsek, A. (2006). The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory & Cognition, 34,* 685-702.

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual arid cognitive processes* (pp. 395-447). New York: Elsevier.

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). New York: Springer.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. *Attention & Performance, 12*, 363-383.

O'Regan, J. K., Levy-Schoen, A., Pynte, J., & Brugaillere, B. (1984). Convenient fixation location within isolated words of different lengths and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 250-257.

O'Regan, J. K., Vitu, F., Radach, R., & Kerr, P. (1994). Effects of local processing and oculomotor factors in eye movement guidance in reading. In J. Ygge & G. Lennerstrand (Eds.), *Eye movements in reading* (pp. 329-348). New York: Pergamon.

Osaka, N. (1992). Size of saccade and fixation duration of eye movements during reading: Psychophysics of Japanese text processing. *Journal of the Optical Society of America A*, *9,* 5–13.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation- verification model for letter and word recognition: The word superiority effect. *Psychological Review*, *89*, 573- 594.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105*(4), 761-781.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford. England: Oxford University Press.

Partee, B., ter Meulen, A. & Wall, R. E. (1990). *Mathematical methods in linguistics*. Kluwer, Dordrecht: Academic Publishers.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8,* 976-987.

Perea, M., & Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 97–120). Hove, UK: Psychology Press.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 767-779.

Perea, M., & Rosa, E. (2000). The effects of orthographic neigborhoods in reading and laboratory word identification tasks: A review. *Psicológica, 21*, 327-340.

Perea, M., & Rosa, E. (2002). Does "whole-word shape" play a role in word recognition? *Perception & Psychophysics, 64*(5), 785-794.

Pexman, P. M., Hargreaves, I. S., Siakaluk, P., Bodner, G., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on word recognition. *Psychonomic Bulletin & Review, 15*, 161–167.

Pickering, M. J. (1999). Sentence comprehension. In S. Garrod & M. J. Pickering (Eds.), *Language processing* (pp. 123-153), Hove, UK: Psychology Press.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology. Learning, Memory, and Cognition 24*(4), 940-961.

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

Plummer, P., Perea, M., & Rayner, K. (2014). The influence of contextual diversity on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 275-283.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*, 174–180.

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language & Cognitive Processes, 20*, 261-290.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 820-833.

Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 726–750.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 148-162.

Pollatsek, A., Perea, M., & Binder, K. (1999). The effects of “neighborhood size” in reading and lexical decision. *Journal of Experimental Psychology: Human Perception & Performance*, *25*, 1142-1158.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*, 1-56.

Pynte, J., New, B., & Kennedy, A. (2007). A multiple regression analysis of syntactic influence in reading normal text. In R. Kliegl & R. Engbert (Eds.), Conference abstracts: 14th European Conference on Eye Movements (ECEM2007). Postdam, Germany, August 19-23, 2007. *Journal of Eye Movement Research,* ISSN1995-8692.

Pynte, J., New, B., & Kennedy, A. (2008). A multiple regression analysis of syntactic and semantic influences in reading normal texts. *Journal of Eye Movement Research, 2*(1), 1-11.

Radach, R., & Kempe, V. (1993). An individual analysis of initial fixation positions in reading. In G. d’Ydewalle & J. Van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 213-225). Amsterdam: North-Holland.

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in words during reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 77-100). Oxford: Elsevier

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65–81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618-660.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception & Psychophysics*, *8,* 21-30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K. (2009). Eye movements in reading: Models and data. *Journal of Eye Movement Research*, *2*(5), 2, 1-10.

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*, 720-732.

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., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*, 468-469.

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41*, 943–954.

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and perceptual span in older and younger readers. *Psychology and Aging, 24*, 755-760.

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 & Psychophysics*, *58*, 734-747.

Rayner, K., Fischer, M. F., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*, 1129–1144.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 779-790.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167-179.

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., & Brown, S. J. (2007). Do readers obtain preview benefit from word n+2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(1), 230-245.

Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46*, 310–323.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14,* 385-388.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829-837.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206–226.

Rayner, K., & Morris, R. K. (1992). Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 163-172.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *Quarterly Journal of Experimental Psychology, 33A*, 351-373.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 327–362). New York: Academic.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Hillsdale: Erlbaum.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. Jr. (2012). *Psychology of Reading* (2nd ed.). New York: Psychology Press.

Rayner, K., Pollatsek, A., Liversedge, S. P., & Reichle, E. D. (2009). Eye movements and non-canonical reading: Comments on Kennedy and Pynte. *Vision Research, 49*(17), 2232- 2236.

Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences, 26*, 507–526.

Rayner, K., Raney, G. E., & Pollatsek, A. (1995). Eye movements and discourse processing. In R. F. Lorch, Jr. & E. J. O'Brien (Eds.), *Sources of coherence in reading* (pp. 9–35). Hillsdale, NJ: Erlbaum.

Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from eye movement priming paradigm. *Psychological Science, 6*, 26-32.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4*, SI 21-49.

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*(5), 1188-1200.

Rayner, K., Slowiaczek, M. L., Clifton, C., & Bertera, J. H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 912-922.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye-movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1290–1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, *27*, 537-544.

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., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science, 17*, 192-193.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Oxford: Elsevier Science.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 274- 280.

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist & S. Everling (Ed.), *Oxford handbook on eye movements,* (pp. 767-786). Oxford, UK: Oxford University Press.

Reichle, E. D., & Laurent, P. (2006). Using reinforcement learning to understand the emergence of "intelligent" eye-movement behavior during reading. *Psychological Review*, *113*, 390-408.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219-237.

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.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*, 4-22.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39*, 4403-4411.

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., Tokowicz, N., Liu, Y., & Perfetti, C. A. (2011). Testing an assumption of the E-Z Reader model of eye-movement control in reading: Using event-related potentials to examine the familiarity check. *Psychophysiology, 48*, 993-1003.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*, 1-21.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z reader model. *Psychological Science, 17*, 742-746.

Riordan, B., & Jones, M. N. (2010). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic representation. *Topics in Cognitive Science*, DOI: 10.1111/j.1756-8765.2010.01111.x

Rohde, D. L. T., Gonnerman, L. M., & Plaut, D. C. (2005). An improved model of semantic similarity based on lexical co-occurrence. *Unpublished Manuscript*.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192-233.

Rosch, E., & Mervis, C. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Roskies, A. L. (1999). The binding problem. *Neuron,* 24, 7-9.

Rubenstein, H., & Goodenough, J. (1965). Contextual correlates of synonymy. *Communications of the ACM, 8*(10), 627–633.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.

Sahlgren, M. (2008). The distributional hypothesis. *Rivista di Linguistica (Italian Journal of Linguistics)*, *20*(1), 33-53.

Schad, D. J., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research, 50*(23), 2600-2616.

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.

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood frequency and neighborhood size effects in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 876-900.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition. *Psychological Review*, *96*, 523- 568.

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., & 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*(2), 402-409.

Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology, 5*, 133-142.

Shallice, T., & Cooper, R. P. (2013). Is there a semantic system for abstract words? *Frontiers in Human Neuroscience, 7*(175), 1-10.

Shaoul, C., & Westbury, C. (2006). Word frequency effects in high-dimensional co-occurrence models: A new approach. *Behavior Research Methods, 38*(2), 190-195.

Shaoul, C., & Westbury, C. (2010a). Exploring lexical co-occurrence space using HiDEx. *Behavior Research Methods, 42*(2), 393- 413.

Shaoul, C., & Westbury, C. (2010b). Neighborhood Density Measures for 57,153 English Words. Edmonton, AB: University of Alberta (downloaded from http://www.psych.ualberta.ca/~westburylab/downloads/westburylab.arcs.ncounts.html).

Shaoul, C., & Westbury, C. (2012). HiDEx: The high dimensional explorer. In P. M. McCarthy & C. Boonthum-Denecke (Eds.), *Applied natural language processing: Identification, investigation, and resolution* (pp. 230-246). Hershey, USA: Information Science Reference.

Sheridan, H., & Reingold, E. M. (2013). A further examination of the lexical-processing stages hypothesized by the EZ Reader model. *Attention, Perception, & Psychophysics, 75*(3), 407-414.

Siakaluk, P., Buchanan, L., & Westbury, C. (2003). The effect of semantic distance in yes/no and go/no-go semantic categorization tasks. *Memory & Cognition. 31*(1), 100-113.

Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*, 451-486.

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35*(6), 1969–1975.

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.

Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science, 12*(4), 485-52.

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

Smith, F., Lott, D., & Cronnell, B. (1969). The effect of type size and case alternation on word identification. *American Journal of Psychology, 82*, 248–253.

Solomon, K. O., & Barsalou, L. W. (2001). Representing properties locally. *Cognitive Psychology*, *43*, 129-169.

Song, D., Bruza, P., & Cole, R. (2004, July 30). *Concept learning and information inferencing on a high-dimensional semantic space*. Paper presented at the ACM SIGIR 2004 Workshop on Mathematical/ Formal Methods in Information Retrieval, Sheffield, UK.

Starr, M. S., & Inhoff, A. W. (2004). Allocation of attention during eye fixation in reading: Use of orthographic information from multiple word locations. *European Journal of Cognitive Psychology, 16*, 203-225.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences, 5*, 156-163.

Staub, A. (2011). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18*, 371– 376.

Staub, A., Rayner, R., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun–noun compounds. *Journal of Experimental Psychology, 33*, 1162–1169.

Stolz, J. A., & Besner, D. (1996). Role of set in visual word recognition: Activation and activation blocking as nonautomatic processes. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1166-1177.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-402). New York: Academic Press.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 247–268). Oxford: Elsevier.

Underwood, G., Clews, S., & Everatt, J. (1990). How do readers know where to look next? Local information distributions influence eye fixations. *Quarterly Journal of Experimental Psychology, 42A*(1), 39-65.

Vainio, S., Hyönä, J., & Pajunen, A. (2009). Lexical predictability exerts robust effects on fixation duration, but not on initial landing position during reading. *Experimental Psychology, 56*, 66–74.

Vanyukov, P. M., Tokowicz, N., Reichle, E. D., & Perfetti, C. A. (2011). *What is the familiarity check? Using event-related potentials to examine how orthographic, phonological, and semantic information affects eye- movement latencies*. Manuscript in preparation.

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301- 326). Oxford: Elsevier.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 47*(25- 26), 3513-3533.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, *47*, 583–600.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron, 24*, 95–104.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14*, 770–775.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A*, 439-456.

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1261–1276.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A re-examination. *Psychonomic Bulletin & Review, 12,* 891-896.

Widdows, D. (2004). *Geometry and meaning*. Stanford, USA: CSLI Publications.

Wilk, M. B., Gnanadesikan, R. (1968), Probability plotting methods for the analysis of data, *Biometrika (Biometrika Trust), 55*(1), 1–17.

Williams, C. C., Perea, M., Pollatsek, A., & Rayner, K. (2006). Previewing the neighborhood: The role of orthographic neighbors as parafoveal previews in reading. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1072-1082.

Yang, S., & McConkie, G. (2004). Saccade integration during reading: Are words necessary? *European Journal of Cognitive Psychology, 76*(1-2), 226-26l.

Yap, M. J., Pexman, P. M., Wellsby, M., Hargreaves, I. S., & Huff, M. J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontier of Human Neuroscience, 6*(72). doi: 10.3389/fnhum.2012.00072

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review, 18,* 742-750.

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*, 971-979.

Yates, M. (2010). Investigating the importance of the least supported phoneme on visual word naming. *Cognition*, 115, 197-201.

Yates, M. (2012). The influence of semantic neighbours on visual word recognition. *Journal of Research in Reading, 35*(2), 215-226.

Yates, M., Friend, J., & Ploetz, D. M. (2008). The effect of phonological neighborhood density on eye movements during reading. *Cognition, 107*, 685- 692.

Yates, M., Locker, L. Jr., & Simpson, G. B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Memory & Cognition*, *31*(6), 856-866.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory & Language, 48*, 779-793.

Zipf, G. (1949). *Human behavior and the principle of least-effort*. Cambridge, MA: Addison-Wesley.