PRONOUNCING PRINTED WORDS: INVESTIGATING A SEMANTIC CONTRIBUTION TO ADULT WORD READING

ELISA COOPER

A thesis submitted in partial fulfilment of the requirements of the School of Psychology, University of East London for the degree of Doctor of Philosophy

April 2013

Abstract

When considering print-to-sound word reading, orthography and phonology are obviously involved. However, another system, that of semantic memory, might also be involved in orthography-to-phonology computation. Whether this occurs is debated in the literature both in the interpretation of behavioural results (e.g., Monaghan & Ellis, 2002; Strain et al., 1995) and in the implementation of semantic memory within computational models of word reading (Coltheart et al., 2001; Plaut et al., 1996). The central aim of this thesis was to investigate whether there is a semantic contribution to orthography-to-phonology computation in healthy adult word reading. Experiments 1-4 used a semantic priming design in which a picture prime was followed either two trials later (Experiments 1, 3, and 4) or one trial later (Experiment 2) by a word target, and this investigated priming of various word types. Regression investigations explored whether semantic features and imageability were unique significant predictors of ELexicon single word reading reaction times while statistically controlling for age-of-acquisition. The two ERP experiments (Experiments 5 and 6) investigated the neurocorrelates of imageability and semantic features and whether there are semantic effects early in the time-course of low frequency word reading. Experiments 1, 2, 4, 5, and 6 and the regression investigations show evidence of a semantic contribution to low frequency regular and low frequency exception word reading. There is also some suggestion of a semantic contribution to high frequency word reading (Experiment 2 and Regression analyses). From the results of the three lines of investigation, it is concluded that semantic information is involved in healthy adult word reading, and these results are best accommodated by the connectionist triangle model of word reading. These investigations also provide information concerning various word types and factors that contribute to “easy” and “difficult” words, semantic memory models and their accounts of priming, and the measures, age-of-acquisition, imageability, and semantic features.

Declaration

This dissertation is the result of my own work, under the supervision of Dr Melanie Vitkovitch, Dr Volker Thoma, and Dr Tom E. Dickins, and includes nothing which is the outcome of work completed in collaboration except where specifically indicated. This dissertation has not been submitted, in part or in whole, for any other degree or qualification. This thesis does not exceed the word limit set by the University of East London in the General Regulations.

Data reported in Chapter 3 were presented as:

Pye, E., & Vitkovitch, M. (2009, September). *Semantic influence on word reading: picture-to-word semantic priming across and intervening item.* Poster session presented at the British Psychological Society Cognitive Psychology Section Annual Conference, Hertfordshire.

Data reported in Chapters 3 and 4 were presented as:

Pye, E., & Vitkovitch, M. (2009, July). *Picture-to-word priming across intervening items provides evidence for a semantic contribution to word naming.* Paper presented at the 24th Annual British Psychological Society Psychology Postgraduate Affairs Group (PsyPAG) Conference, Cardiff.

There are plans for adapted versions of Chapters 3-7 to be submitted for publication: Cooper, E. & Vitkovitch, M. (2013). *Long Lag Semantic Priming of Word Targets and The Importance of What is in the Middle*. Manuscript in preparation.

There are plans for an adapted version of Chapters 8 to be submitted for publication:

Cooper, E. & Vitkovitch, M. (2013). *Investigating whether Semantic Features, Imageability, and Age-of-Acquisition are Significant and Unique Predictors of Single Word Reading Times.* Manuscript in preparation.

Data from four participants reported in Chapter 9 were collected with the help of Dr Melanie Vitkovitch. The data from Experiment 6 were included as part of different analyses using a larger stimulus set by Dr Melanie Vitkovitch, the results of which were presented as a poster at the 15th World Congress of Psychophysiology of the International Organization of Psychophysiology, Budapest. The proceedings of which are published:

Vitkovitch, M. & **Cooper-Pye, E.** (2010). Semantic features and word naming: Do event related potentials vary according to the number of features elicited by an object name? [Abstract]. *Proceedings of the 15th World Congress of Psychophysiology of the International Organization of Psychophysiology, Budapest. International Journal of Psychophysiology, 77,* 309-310.

Table of Contents

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 9

List of Tables

List of Figures

Acknowledgements

As the work of a PhD is the effort of one individual, the phrase “it takes a village” is probably not used too often when referring to this effort. Yet, in reference the amount of support I’ve received throughout this PhD, no phrase seems more fitting. It has taken a village to shore me up. I’m unsure how to express my immense feelings of gratitude to so many people; my words are futile devices, but I can start by sitting down and writing this letter of thanks to all my friends who have helped along the way.

First and foremost, thank you to the 220+ participants who volunteered for my studies, sometimes for nothing in return. Without you, this research quite literally would not have been possible.

With thanks also to UEL and the School of Psychology for taking me on as student and for funding me as I researched. I’m grateful to my supervisors, Melanie, Volker, and Tom. With special thanks to Melanie whose aim was always the same as mine- to see it finished. Melanie, you had an endless capacity for guiding me (and for reading drafts), you always wanted to make me better; thank you for helping me. I’ve learned so much.

Thanks also to other UEL folks. This includes but is certainly not limited to Traci, Peter, Steve, Oliver, and Anita. To Silvia for all your help with Net Station and EEG, Chapter 9 would have been unbearably difficult without your guidance. Thanks to all my many many office mates (both in 141 and the new shiny fish tanks), who kept me and this PhD company, with special honours to Elliott, Haiko, Elley, Emma, Caroline, and Trevor. I’m sometimes a little lonely without you.

I’m also grateful to the MRC Cognition and Brain Sciences Unit, who let me use their library…a lot. Thank you to Dr Peter Watson for his help with the regression investigations of Chapter 8; you always had time, answers, and books for me. To Tristan Bekinschtein and his team, thank you for the use of Net Station and for making it easy. To Rik Henson, for introducing me to Hofstadter’s law, and being as supportive of me finishing this thesis as a line manager could be. Thanks also to Andrea and Tina. Georgie, Clare, and Michael. I’m grateful for the coffee breaks; the knock-on always came just in time. Thanks to Michael also for never tiring of asking if I’d submitted and having a sympathetic ear.

Thank you to my parents (and my big sister) who taught me that being a southern woman means that I’m inherently strong and feisty. Thank you for saying that I can get through anything and can do anything I set my mind to, even though these things might not necessarily be true. Thanks also to my folks for providing me with an infinite pile of starbucks, a bit of sneaky dosh at the end, and with an excellent education in the early years, which included teachers like Mrs Mac, who took a kid who hated science and turned her into scientist (thank you), Miss Cheatham (thank you), and Mr LaGrange (thank you).

To the Johns of John Street and to The Portland Arms folks, thank you for being my adopted families and for keeping me hugged, fed, and sane when there was seemingly no one else. Thanks to James, who reminded me how very important it is to laugh and be happy, and thanks to Stephen for always reminding me of who I am and what friendship is all about.

To the coffee and cake ladies- Lauren, Lauren, and Cara thank you for the chat. Dear bestie Scott, thank you for never caring about this science stuff; it made doing the stuff easier and our time together better.

I will forever be grateful to Tim Pye. I continue to find that much of who I am now is due to the years we spent together. You're the foundation of the adult me. You started this PhD journey with me and by forming who I am today, you have been also been responsible for this PhD's completion. Thank you for lots of things, said and unsaid.

Special thank you prizes for listening go to Dr. Michelle DiMeo and Dr. Samrah Ahmed-Ali. You always heard, you always empathised, and you always knew just what say to have me carry-on just that bit more; you are wise ladies indeed. Sams, dude, thanks also for reading, I've not forgotten all of the reading!

With gratitude to Prof Karalyn Patterson, thank you for reading drafts of my chapters, providing me with comments, and with science discussions. You were my biggest cheerleader during this PhD. You've been a science fairy godmother, paving the way for me, probably, in ways I'll never know. I understand how people name children for you and if they'd let me name this thesis *Karalyn* I would do. Thank you. Thank you. Thank you.

My words fall short when thanking curly-haired Kevin K Wright. Thank you for being the team's coach, domestique, shrink, chauffer, maid, chef, waiter, cheerleader, proof-reader (He read the whole thing! Multiple times!), tear-collector, comforter, and hug-giver, amongst so many other things. You are outstanding. Will either of us ever be able to explain all you've done? You seemed to have an infinite well of patience, tolerance, encouragement, and most of all, love. How you survived me (and the anxiety, stress, and tears) during the last 2 and half years I'll never know, but I'm very glad you did. (Though not the captain), you steered the ship. You're the go-to man and the package. It's just what you do. Endless, infinite gratitude to you for helping me achieve this. You make it always…sometimes…easy.

For Karalyn and For Kev,
whose value of me has never been dependent on whether I finished this thesis, but whose unconditional care, love, and support meant that I did.

Chapter 1

The Thesis Topic and a Review of Computational Models of Word Reading

1.1 Introduction

1.1.1 Introducing word reading

Word reading is performed by healthy literate adults with ease. The systems used to accomplish word reading, however, are ultimately more complex and controversial than the act initially seems, as will be shown. For this thesis, word reading is defined as processing a single word from its printed form to its correct sound form. The terms “word naming”, “single word reading”, and “reading aloud” are also used with this meaning.

When considering print-to-sound word reading, two distinct systems are apparent; they pertain to the initial print form and the final, correct sound form. The print form is composed of the written spelling of a word in letter appearance, known as the orthographic form. The pronunciation form of a word is composed of the correct sounds for the specific combination of letters; this is known as the phonological form. It is evident that both orthographic and phonological systems are involved in word reading, but it is possible that in order to correctly read a written word another system, that of semantic memory, might be involved (Patterson & Hodges, 1992).

1.1.2 Introducing semantic memory

Semantic memory is a form of memory that includes general concept information that constitutes knowledge about the world, including words and objects (Tulving, 1972). For example, semantic knowledge about "a glove" might include encyclopaedic-type information, such as: is non-living, an item of clothing, and worn on the hand. Correct word reading might involve the activation and use of information within semantic memory, but this has yet to be confirmed in the literature as will be presented throughout the introduction chapters. The terms "semantic memory", "semantic information", "semantic knowledge", and "semantics" are also used to indicate this type of knowledge, and as a short-hand to indicate the processing that might be involved in word reading.

There are a number of issues under consideration in the semantic memory literature. One issue is the storage and organisation of semantic information (Anderson, 1993; Becker, Moscovitch, Behrmann, & Joordens, 1997; Collins & Loftus, 1975; Masson, 1991, 1995; Mirmam & Magnuson, 2008). This is reviewed briefly in Chapter 2. Another issue is the activation of semantic memory for the comprehension of printed words and whether or not this is via phonology (Harm & Seidenberg, 2004). This is separate from the focus of this thesis. Principals that guide comprehension research, however, may also apply to investigations of word reading for pronunciation. For example, in a computational model that explains word comprehension, there are connections between semantic memory and orthography, and semantic memory and phonology (Harm & Seidenberg, 2004). The relevant issue of this thesis is whether semantic information might contribute to orthography-to-phonology computation for the purposes of forming the correct phonology of a word, as this issue has not been resolved in the literature (Blazely, Coltheart, & Casey, 2005; Monaghan & Ellis, 2002; Patterson, Lambon-Ralph, Jefferies, Woollams, Jones, Hodges, & Rogers, 2006; Strain, Patterson, & Seidenberg, 1995 to name only a few). It is this controversy that forms the foundation of the work of this thesis.

1.1.3 Controversy in behavioural results

Behavioural research, including work using reaction times and error rates during single word reading, with either factorial designs or regression methods, has failed to reach a conclusion as to whether semantic memory contributes to word reading (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002; Strain et al. 1995; Strain, Patterson, & Seidenberg, 2002). Examples from the human empirical behavioural literature are reviewed in Chapter 2, but to summarise, while some studies implicated a semantic memory contribution when accounting for word reading effects, other investigations have explained these effects as being due to non-semantic characteristics of the stimulus words.

1.1.4 Controversy in computational models of word reading

Computational models of word reading attempt to account for human word reading performance, such as those mentioned above and reviewed in Chapter 2, by simulating them in computer networks. Therefore, as there is a debate in the empirical literature, the computer models also show this controversy. For example the Dual-Route Cascaded model (DRC) (Coltheart, Curtis, Atkins & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the CDP+ and CDP++ models (Perry, Ziegler, & Zorzi, 2007; 2010b; Zorzi, Houghton, & Butterworth, 1998), the triangle model (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg, & Patterson, 1996; Harm & Seidenberg, 1999, 2004), certainly allow, at least theoretically, for semantic information to contribute to phonology through the connections within the models. Models, however, differ in the emphasis and implementation of a semantic contribution for phonology formation, as will be detailed in subsequent sections. Harm and Seidenberg's (2004) connectionist triangle model has implemented semantic memory and there are direct connections between semantic memory and orthography, and semantic memory and phonology. This is in contrast to the DRC and CDP + and ++ models, in which semantic information remains unimplemented (Coltheart et al., 2001; Perry et al., 2007, 2010b). Authors of a more recent computational modelling article highlight the debate surrounding semantic memory and word reading, writing, "The role of semantics in both normal and impaired naming of written words is controversial" (Perry et al., 2007, p. 278) and "One major controversy, however, concerns the role of semantics [in reading aloud a printed word]" (Perry et al., 2007, p. 302). The controversy concerning a semantic contribution to word reading is present throughout the behavioural and computation modelling literature as will be presented in detail in this and the next chapter.

1.1.5 The central aim of the thesis

The programme of research described in this thesis addresses the question of whether semantic information contributes to orthography-to-phonology computation in word reading using quantitative methods. The studies of this thesis use the English language and healthy adult participants. This empirical work includes studies that use factorial and regression designs, and behavioural and neurophysiological approaches.

1.1.6 An overview of Introduction Chapters 1 and 2

The remainder of this chapter and the next further delineate the controversy concerning a semantic contribution to word reading through a review of the literature. The remainder of the current chapter introduces word stimuli and their characteristics that are important both in the literature and in this thesis, as this needs further defining before proceeding with the review. This is followed by a review of the computational models of word reading with an emphasis on the way in which semantic memory is accounted for when orthography-to-phonology is computed. These models provide a theoretical framework for the experiments of this thesis. Additionally, as these models seek to account for human behaviour from experiments in the literature, they also provide an overview of the way in which these behavioural results are currently explained, providing a foundation for understanding the studies reviewed in Chapter 2. The subsequent chapter (Chapter 2) contains a review of the relevant experiment-based literature, mainly behavioural in nature, that has investigated a semantic memory contribution to word reading. These studies have also, in part, informed the computational models of word reading. The research of Chapter 2 provides an empirical foundation for the work of this thesis and concludes with additional aims addressed by the experiments in this thesis.

1.2 Defining word types

Human performance on word reading tasks reveals the ease with which a word is read. Researchers seek to explain why certain words may be read more quickly or more accurately than others by identifying characteristics of words that may be responsible for these effects. Identified characteristics that are central to the ease with which a word is read include (a) word frequency (Balota & Chumley, 1985; Forster & Chambers, 1973; Solomon & Howes, 1951) and (b) the regularity/consistency of spelling-to-sound relationship, as explored in the subsequent sections (Glushko, 1979); these benchmark effects are detailed in the following sections. It should be noted that a third major factor also affects word reading times is word length, as measured either in number of letters or number of syllables (Butler & Hains, 1979; Forster & Chambers, 1973; Frederiksen & Kroll, 1976). This factor is usually eliminated from word difficulty classification by restricting stimuli to single-syllable words that are no more that six to seven letters in length. Below this number of letters, the effects of word length on word reading are minimal (Damian, Bowers, Stadthagen-Gonzalez, Spalek, 2010; Young & Ellis, 1985). These factors and other words measures, such as familiarity, imageability, and age-of-acquisition, are often correlated with one another (Strain et al., 2002; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004), and these factors are detailed in Chapter 2, and discussed in relation to the results of this thesis in Chapter 8, Chapter 9, and Chapter 10. It can be difficult to control all intercorrelated measures when designing experiments that use word stimuli, and, this must be considered when interpreting the results.

Word types are introduced here to ease discussions of the computational models of word reading in this chapter and the studies reviewed in Chapter 2. Throughout the computational modelling review, which follows in Section 1.3, the empirical literature review in Chapter 2, and the investigations of this thesis, the importance of these word types will be discussed further.

1.2.1 Frequency

Word frequency is the number of instances of a particular word (usually per one million words); written frequency is the number of times the word appears in a printed text. This provides a good estimate of how frequently a literate person is likely to encounter that specific word in print form. High frequency words are read aloud more quickly than low frequency words[1], demonstrating that they are processed more efficiently (Balota & Chumley, 1985; Forster & Chambers, 1973; Solomon & Howes, 1951). In the earliest finding, Solomon and Howes, using a tachistocope, presented words that varied in frequency to healthy adult participants and measured the amount of time it took for the words to be correctly identify, gradually increasing the duration of exposure until the participant could report the word after exposure to the word ceased. They found a significant correlation between the two measures, with high frequency words being identified significant faster than low frequency words. Balota and Chumley tested healthy adult participants on a delayed word reading task, with words being either high or low in frequency. A significant frequency effect with high frequency words being read more quickly than low frequency words was found. This effect was found in the three experiments of this series at various delays between word presentation and

[1] The frequency thresholds for high and low frequency words are not identical in all studies of word reading. Examples of high and low frequency thresholds, using Kucera and Francis counts are provided here. Strain et al. (2002) considered high frequency to be above 70 instances per million, and low frequency to be below 30 million. Monaghan and Ellis (2002) considered high frequency to be above 14 instances per million, while low frequency was below 13 instances. Balota and Chumley (1985) considered high frequency to be above 36 instances per million and low frequency to be below 7 instances. Andrews, Woollams, Bond (2005) considers low frequency items to be below 36 per million, whereas Andrews (1982) considered low frequency to be below 20 per million and high frequency to be above 40 per million.

participants being cued to report aloud the word (Experiments 1, 2, and 3) and when the delay interval was filled with a task (Experiment 3). These two noteworthy articles (Balota & Chumley, 1985; Solomon & Howes, 1951) provide a sample of the work that finds the benchmark frequency effect of high frequency words being read faster than low frequency words, when all other factors have been matched.

Because high frequency words are read faster than low frequency words, higher frequency words could be said to be easier to read than lower frequency words. The connectionist triangle model of word reading, which is reviewed in detail in a subsequent section, explains frequency effects in the following way. Higher frequency words have stronger, more efficient connections between orthography and phonology. For a high frequency word, connections are stronger because it is seen more often and consequently that specific phonology is retrieved in reference to that specific orthography more often, reinforcing the connections. Though frequency is accounted for differently in the computational models of word reading, as is noted in Section 1.3, it is accounted for as a factor that partially explains efficiency of orthography-to-phonology computation (Coltheart et al., 2001; Harm & Seidenberg, 2004; Plaut et al., 1996).

1.2.2 Spelling-to-sound correspondence

Orthography-to-phonology efficiency can also be partially defined in terms of spelling-to-sound correspondence. The languages of the world, in written form, can be classified according to how forthright their orthography-to-phonology mappings are, as will be described. That is, whether letters or combinations of letters are pronounced in the same way in all circumstances or whether letters are pronounced differently according to the word in which they appear (Andrews, Woollams, & Bond, 2005; Frost, Katz, & Bentin, 1987; Fushimi, Ijuin, Patterson, & Tatsumi, 1999; Plaut et al. 1996; Shibahara, Zorzi, Hill, Wydell, & Butterworth, 2003; Wydell, Butterworth, & Patterson, 1995; Wydell, Patterson, & Humphreys, 1993). In some languages, such as in Italian and Serbo-Croatian, letter-to-sound mappings are consistent and transparent. Such languages are considered to have "shallow" orthography (Frost et al., 1987; Zorzi, et al., 1998). In other languages, such as English, Hebrew, and Japanese Kanji, sometimes described as "deep" orthographies, orthography-to-phonology mappings are inconsistent and opaque because pronunciation is context dependent, i.e., certain letters can be pronounced in more than one way, and the way in which they are pronounced depends on the whole-word-context in which they appear; an example follows (Fushimi et al., 1999; Plaut et al., 1996; Shibahara et al., 1993; Wydell et al., 1995; Wydell et al., 1993).

Even in the "deep" orthography languages, orthography-to-phonology mappings can also contain *some* consistencies, leading researchers such as Plaut et al. and Fushimi et al. to describe as a "quasi-regular". For example in English, "B" is always pronounced /b/ as in "blouse". No other letter represents this sound, nor can "B" be pronounced any other way. An example of an inconsistently pronounced letter in English is "S" pronounced /z/ as in "hose" but /s/ as in "dose". The /s/ sound could also be represented by the letter "C" as in "cite". The terms regularity and consistency are used to define spelling-to-sound correspondences and are described in the following section. They can be used separately (Sections 1.2.2.1 and 1.2.2.2) or in unison (1.2.2.3), though they account for different aspects of spelling-to-sound typicality.

1.2.2.1 Regularity

Regularity refers to whether a word follows the “rules” of letter-to-sound correspondence within English (see Venezky, 1970). For example, there is a general rule that vowels should be long (or “tense”) preceding a single consonant and short (or “lax”) preceding a doubled consonant. Therefore, “table” (tense) and “babble” (lax) are both regular words. Single syllable words ending in a consonant without a final “e” usually take the lax vowel according to the rules. For example, “hint”, “tint”, “mint”, “print”, “stint”, etc. are all regular words, but “pint”, in which the correct vowel is a tense “I”, is an irregular word. Regularity is treated as a categorical distinction: a word is either regular or irregular. It is easier to make this distinction for single- than multi-syllabic words, which may be one reason why studies of word reading have largely, until recently, restricted their stimuli to single syllable words (Yap & Balota, 2009). However, this dichotomy can be problematic; for example, not all researchers would agree which word out of “lemon” and “demon” is irregular.

1.2.2.2 Consistency

A word is consistent if it has the same pronunciation as other words with the same spelling (Andrews, 1982; Glushko, 1979). Consistency is defined in relation to other similarly spelt words, not in terms of rules. The section of the word most often selected for consistency judgements is the word’s “body”, though smaller units can be used too, e.g., pairs of letters, “kn”, or “ea” (Andrews et al., 2005). In single-syllable words the body is all the letters following the initial consonant or consonant group, for example, the ‘-int’ in “hint”, “tint”, “mint”, “print”, “stint”, etc. If all the words with the same orthographic body rhyme, then they all are consistent. This is true of the body “-and” (“hand”, “land”, “band”, “strand”, etc). However, none of the words ending in “-int”, even regular ones like “hint”, are consistent, because “pint” shares the same body. Glushko (1979) was the first to argue for effects of consistency above and beyond effects of rule-based regularity. In Experiments 1 and 2, he showed a consistency effect in word and non-word reading reaction times of healthy adult participants. In real word reading and in non-word reading, consistent letter strings were read aloud faster that inconsistent letter strings. In Experiment 3, regularity was held constant while words were manipulated on consistency; regular consistent, regular inconsistent and exception words were used. Regular inconsistent words were read more slowly than regular consistent words. These studies by Glushko demonstrate a benchmark effect in word reading: consistent words are read more quickly than inconsistent words when all other factors are matched.

Unlike regular words, consistent words can have different degrees of consistency and this can be quantified, as follows. The consistency of a particular word can be calculated in reference to “friends” and “enemies” within that word’s orthographic body neighbourhood (Jared, McRae, & Seidenberg, 1990; Jared, 1997; Ziegler, Stone, & Jacobs, 1997). To illustrate, words with body rhymes are known as friends, such as “pear” and “bear”. However words, such as “pear” and “dear”, with different body pronunciations are known as enemies. At a simple level, this can be calculated with a count of these friends and enemies. For example, there are 21 words in the “-ear” orthographic neighbourhood. Sixteen are pronounced like “dear” and five like “pear”. If the word itself is counted as a friend, then “fear” has 16 friends and five enemies. “Wear” has five friends and 16 enemies. With additional calculation, consistency can also account for word frequency by summing or averaging the frequency of friends and enemies. For example “sear” is relatively low in frequency and therefore may contribute less to pronunciation that the highly frequent “dear”.

1.2.2.3 Regularity and consistency together

The terms regularity and consistency can be used jointly to define a word, as these terms classify different types of spelling-to-sound correspondence. Since consistency effects occur even when regularity has been controlled, the work of Glusko (1979), reviewed above, shows regularity and consistency are two separate measures of a word’s spelling-to-sound that can be used simultaneously to ultimately capture different aspects of spelling-to-sound correspondence. A word can be classified using terms of regularity *and* consistency, such as the regular consistent word “hand”, the regular inconsistent word “sear”, or the irregular inconsistent word “pear”. Words that are irregular and consistent, though not completely non-existent, are very rare. Therefore in the main, most irregular words also happen to be inconsistent.

1.2.2.4 Conclusions on spelling-to-sound correspondence

The various computational models of word reading favour either regularity or consistency as the more important variable when computing orthography-to-phonology (Sections 1.3.2.1, 1.3.3). One of the routes on the Dual-Route Cascaded model of word reading, reviewed shortly, includes spelling-to-sound rules and therefore favours regularity (Coltheart et al., 1993; Coltheart et al., 2001). By contrast the triangle model lacks rules; connections between the distributed representations of orthography and phonology are weighted toward the pronunciation of friends, especially higher frequency ones, making consistency more influential in this model. There are empirical investigations that test the validity of regularity or consistency, i.e., assess whether orthography-to-phonology computation is better captured by regularity or consistency (for example Andrews; 1982; Andrews et al, 2005; Cortese & Simpson, 2000; Jared et al., 1990; Jared, 1997; Jefferies, Lambon-Ralph, Jones, Bateman, & Patterson, 2004; Perry, Ziegler, Braun, & Zorzi, 2010a; Spencer, 2009; Zeigler, Jacobs, & Stone, 1996). The aim of this thesis, however, is to investigate a possible semantic memory contribution to word reading regardless of whether orthography-to-phonology computation is better simulated by regularity or consistency within word reading models. The results of this thesis' studies, of course, may be better accounted for by a specific model, and this model will use either regularity or consistency to account for orthography-to-phonology computation.

Irregular words and inconsistent words are both read more slowly than regular words and consistent words, respectively (Baron & Strawson, 1976; Glushko, 1979). This finding is important to the aim of this thesis because the literature (reviewed in this chapter and the next) has indicated that a semantic memory contribution to word reading might occur when orthography-to-phonology is less efficient and slow (Plaut et al., 1996; Strain et al., 1995, 2002). If semantic memory contributes to word reading in this circumstance of slow orthography-to-phonology computation, then when investigating the aim of this thesis, difficult words regardless of whether they are irregular and/or inconsistent should be carefully considered.

Following the major trend in the literature the terms regular and exception will be used within this thesis to refer to typical spelling-to-sound and atypical spelling-to-sound, respectively. The precise boundaries of these terms are defined in Chapter 3.

1.2.3 Four word types

The two factors frequency and regularity often interact in human word reading performance (Andrews, 1982; Paap &Noel, 1991; Seidenberg, 1985; Seidenberg & McClelland, 1989; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). They are used to orthogonally divide words into four word "types": high frequency regular, high frequency exception, low frequency regular, and low frequency exception. These four word types are used throughout this thesis, both in the literature reviews and the research itself.

High frequency regular words such as "table" and "hand" are read quickest, and therefore defined as the easiest to read of the four types (Andrews, 1982; Paap &Noel, 1991; Seidenberg, 1985; Seidenberg & McClelland, 1989; Seidenberg et al., 1984). High frequency exception words like "have" and "push" and low frequency regular words like "sear" and "vest" are also read relatively quickly and easily, and are interchangeable when considering word difficulty level (Andrews, 1982; Paap &Noel, 1991; Seidenberg, 1985; Seidenberg & McClelland, 1989; Seidenberg et al., 1984). The most difficult and slowest word type is usually low frequency exception words like "sew" and "caste" (Andrews, 1982; Paap &Noel, 1991; Seidenberg, 1985; Seidenberg & McClelland, 1989; Seidenberg et al., 1984). Furthermore, researchers can choose words to exploit these orthogonal distinctions in order to choose word stimuli that range from "easy" to "difficult" when measuring with these two factors. As a reminder, these factors, however, are also correlated with other measures. It is possible that if not accounted for, other factors can also account for a word being "easy" or "difficult". Specifically, within the context of this thesis, easier and more difficult words (whether because of a strictly orthogonal interaction of frequency and regularity or because of additional correlated measure, as discussed in Chapter 10) could offer an opportunity to test the specific circumstances in which semantic memory might contribute to orthography-to-phonology computation, as will be shown.

1.3 Computational models of word reading

As mentioned previously, computational models of word reading are informed by human performance from a variety of methodological approaches. Computer-based implementations of word reading attempt to accommodate a number of key results from the broader empirical literature, including behavioural investigations and neuroimaging studies with participants who are healthy and participants with reading deficits. The models are not solely designed to accommodate investigations of a semantic contribution to orthography-to-phonology computation. Of note is that, though these models attempt to accommodate human performance, the theories formed from these models also predict human reading performance (Seidenberg & Plaut, 2006).

The majority of the research into word reading, including research that informs the development of the models of word reading, is primarily the result of studies in English with its “quasi-regular” orthography (see Section 1.2.2; Coltheart et al., 2001; Plaut et al., 1996). The triangle model of word reading aims to be a universal model of word reading suitable for all languages (Plaut et al., 1996; Seidenberg & McClelland, 1989) (also see Seidenberg & Plaut, 2006 for discussions of creating a theory of word reading) and at the very least, is suitable for modelling all quasi-regular languages, not only the English language (Plaut et al., 1996, Shibahara et al., 2003). In contrast, the DRC model of word reading is suitable for the interpretation of results from Roman alphabet languages only (Coltheart et al., 2001).

1.3.1 Semantic memory and word reading models

The various computational models of word reading differ in their implementation of the three systems important to word reading- that is orthography, phonology, and semantic memory, as was briefly noted in Section 1.1.4. All models, theoretically, include a semantic system. Models differ in orthography-to-phonology connections and how they work, whether serial cascaded processes or parallel processing, the storage of information, whether local or distributed representations, and the importance of semantic information to phonology computation. The following sections review the computational models of word reading and the implementation of orthography, phonology, and semantics within them, highlighting the role for semantic memory in orthography-to-phonological computation.

1.3.2 Computational models of word reading that have not implemented semantic memory

1.3.2.1 The DRC model of word reading

One computational model of word reading is the Dual-Route Cascaded model (DRC). In the DRC model there are two main routes to word reading; words are read via rules and via stored whole-word representations (lexical entries) (see Figure 1.1) (Coltheart et al., 1993; Coltheart et al., 2001). The non-lexical, grapheme-to-phoneme conversion (GPC) route, correctly computes phonology for only regular words and non-words (letter strings that have legal English spellings, but are not real words and do not have meanings, such as “mand”), as is depicted in the right of Figure 1.1. The other route, the lexical route, produces the correct phonology for irregular words, or any item that cannot be pronounced using rules, and is depicted in the left-side route of Figure 1.1; it is also capable of producing the correct pronunciation for regular words. Words are processed using both routes with elements represented locally in nodes. The two routes interact at the phoneme encoding system.

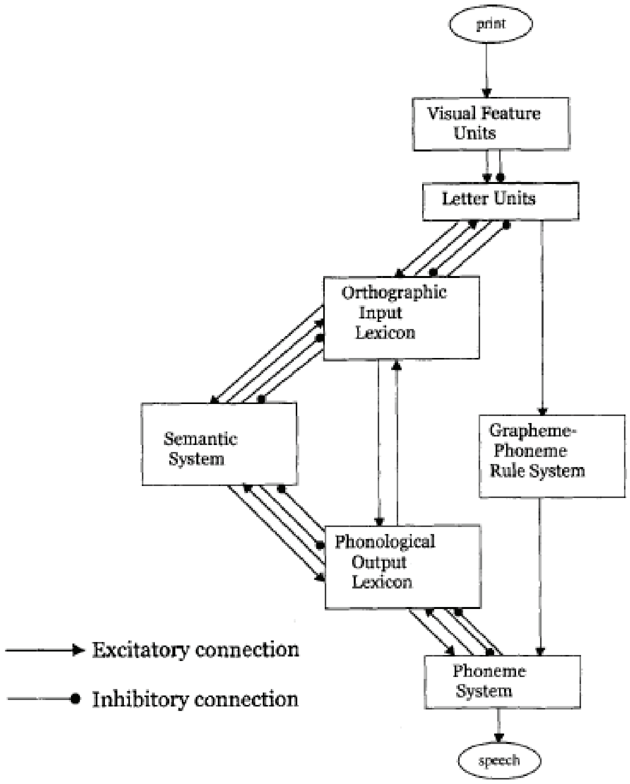

Figure1.1. The DRC Model as depicted in Coltheart et al. (2001), Figure 7, pg 214. The right side of the diagram depicts the grapheme-to-phoneme (rule-based) conversion route (GPC). The left side depicts the lexical route (based on whole word storage), which theoretically can either include or bypass semantics. However, the semantic lexicon is not implemented in this model. Inhibitory (———●) and excitatory (———▸) mechanisms are depicted individually. Copyright © 2001 by the American Psychological Association. Adapted with permission. The official citation that should be used in referencing this material is 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*(1), 204-256. The use of APA information does not imply endorsement by the APA.

The GPC route of the DRC model assembles phonology using English regularity rules as programmed by the modellers (Coltheart et al., 1993; Coltheart et al., 2001). The programmed rules of the English language (Sections 1.2.2.1, 1.2.2.4, 1.2.3) dictate the correct sound correspondence for each letter serially, with each letter's orthographic form being matched to the phonological form. Length effects, i.e., that longer words are read more slowly than shorter words (Butler & Hains, 1979; Forster & Chambers, 1973; Frederiksen & Kroll, 1976), are accounted for on this route as it translates letters to sound from left to right within a word letter-by-letter. The GPC route fails to calculate the correct phonology for an irregular word, even if it has many "friends" (Section 1.2.2.2). For example, "table" is read correctly. The nonword "mand" is read according to the rules, and sounds like "hand". However, "pear" is read to rhyme with "dear".

The lexical (also known as the direct) route of the DRC model uses whole word representations (orthographic lexicon) and their pronunciations (phonological lexicon) (Coltheart et al., 1993; Coltheart et al., 2001; Coltheart, 2006a). It references a programmed whole phonological entry for the whole orthographic entry in the lexicons, somewhat analogous to referencing information in a dictionary. Word frequency is accounted for in the amount of time it takes for the computer model to reference the correct pronunciation in the phonological lexicon (Andrews et al., 2005). The lexical route can produce the correct phonology for all real words, including irregular words.

Low frequency exception words are the slowest processed in the DRC model as the two routes produce conflicting pronunciations. On the rule based GPC route "pear" will be computed to rhyme with "dear". On the lexical whole-word route "pear" is computed correctly, to rhyme with "bear". In resolving this pronunciation conflict, exception words take longer to be read than regular words. Regular words do not create this conflict because both routes arrive at the same suggested pronunciation.

The lexical route also theoretically includes semantic memory (semantic lexicon). Coltheart (2006b) describes the three lexicons (orthography, phonology, and semantics) as separate entities that do not interact for phonology formation, claiming that semantic information would only be accessed for comprehension after phonology has been correctly computed and the phonological encoding system engaged.

Also of note, within this model the semantic system can be bypassed. The DRC model, therefore, has three routes: the non-lexical GPC route (an orthography-to-phonology route), the lexical-bypassing-semantics route (another orthography-to-phonology route), and the lexical-via-semantics route (an orthography-to-semantics-to-phonology route; Coltheart et al. 2001) (see Figure 1.1). DRC modellers argue that words can be read correctly using the GPC route and the lexical-bypassing-semantics route (Coltheart et al.; Coltheart, 2006a, 2006b). Within the DRC model, words are read using only orthography and phonology, but not semantic memory, as it is unimplemented (Coltheart et al., p. 271). The authors state, “We could not consider any effects involving semantics…because the DRC model at present has no semantic system” (Coltheart et al., p. 200). The authors conclude that the lack of an implemented semantic system is a strength of the DRC model when attempting to account for impaired reading (Coltheart et al., p. 245). Similar claims- that semantic memory is not needed to account for healthy or impaired word reading- are present in the literature (Blazely, et al., 2005; Coltheart, 2000; Coltheart, 2004; Coltheart, 2006a, 2006b). Although semantic memory is included in the DRC model in theory, semantic memory does not contribute to orthography-to-phonology computation.

1.3.2.2 CDP+ and CDP ++ models of word reading

There are other models of word reading, the connectionist dual process (CDP+, CDP++) models, that have not implemented semantic memory. The CDP+ and CDP++ models of word reading combine the DRC model's two routes with the connectionist principles used in the triangle model (presented in Section 1.3.3.), claiming to use the successful parts of other models (Perry et al., 2007). The CDP + model builds on the connectionist dual process (CDP) model (Zorzi et al., 1998) and the CDP++ model extends the model to accommodate one and two syllable word reading (Perry et al., 2010b; Yap & Balota, 2009). When modelling semantic memory for phonology computation, these models are the same as the DRC model. Like the DRC model, these models have two main routes to compute orthography-to-phonology; there is a GPC rule-based route, and a lexical whole word storage route.

The CDP+ and CDP++ differ from the DRC model only the implementation of the GPC route. The rules used for assembly of the correct phonology on the GPC route are not programmed explicitly, as they are in the DRC model. Instead the model develops the rules of pronunciation through training a corpus; this is how combinations of letters are pronounced are "learnt", similar to the triangle model (Section 1.3.3), and consistency is important in this model (Zorzi et al. 1998, Perry et al. 2007). In the CDP++ model there has also been a change to the GPC route's grapheme system so that it can now accommodate longer words (Yap & Balota, 2009). These models are described as recreating consistency effects, and regularity and frequency effects well (Perry et al., 2007).

There are no changes to the lexical route in the CDP (+ and ++) models, which is taken directly from the DRC model. Therefore, these models do not include an implemented semantic system and attempt to accommodate human word reading performance without semantic memory (Perry et al., 2007; 2010b). Perry and colleagues state, "We implemented a localist lexical route that is as close as possible to that of the DRC….We discarded the alternative solution of implanting the semantic pathway of the triangle model" (Perry et al., 2007, p. 278). Like the DRC model there are three routes on the CDP+ model: the GPC route (an orthography-to-phonology route), the lexical-bypassing-semantics route (an orthography-to-phonology route), and the lexical-via-semantics route (an orthography-to-semantics-to-phonology route). After training the model on the large CELEX database of over 7,000 words and optimising the model in order to, amongst other factors, balance the input of the two pathways to word reading computation, word reading of a large 3,000 word corpus was simulated by this computer model. The authors provide the results of the computer model's simulation, including the percentage accuracy of how well the model computed the correct phonological output from the orthographic input, including real words and non-words. They evaluate these results and conclude that their model simulates word reading well enough. Since the model does not include semantic memory on the lexical route and word reading is simulated adequately, the CDP+ modelers conclude, "In summary, it appears that word meaning does not have an important contribution in written word naming" (Perry et al., 2007, p. 303). These conclusions and the reasons behind them- that is, word reading being simulated well enough in a computational model without an implemented semantic memory- are identical to those of the DRC modelers.

1.3.3 A computational model that has implemented semantic memory: The connectionist triangle model of word reading

In contrast to the DRC, CDP+, and CDP++ models, there are computational models of word reading that have implemented semantic memory in the process of word reading;

one is the connectionist triangle model of word reading[2]. The question of whether semantic information is involved in the process of computing phonology from orthography is the central aim of this thesis; the triangle model is therefore discussed in finer detail than the previously described models that do not implement semantic memory. In the triangle model, called this because of its visual appearance in diagrams (Figure 1.2), the same computational principles apply to all words, regardless of word type (Harm & Seidenberg, 1999; Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989; Seidenberg & Plaut, 2006).

Orthography, phonology, and semantic memory (represented by the circled name of the system in Figure 1.2) are interacting systems (represented as arrows in Figure 1.2) made up of sets of neuron-like units (see Figure 1.2, in which semantic memory is centrally located), and these three systems operate collectively to produce a word's phonology.

Connections are specified between orthography and phonology, orthography and semantics, and semantics and phonology (Harm & Seidenberg, 2004; Plaut et al., 1996)[3].

[2] The junction computational model of word reading also includes semantic information in the process of phonology computation, and also attempts to model multi-syllable words (Kello, 2006; Kello & Plaut, 2003). This model, only in its pilot stages, is in some ways similar to the triangle model, as it is a single-route model that includes semantic information. Orthography, semantics, and phonology are included, implemented, and joined through a single shared lexicon, and words are read using these three systems. All words are read using the same links, pathways, and information; there are not different routes for different word types (Kello, 2006). However, the use of nodes, lexicons, and local representations within this model, are in contrast to the triangle model. Semantic information is included as text co-occurrence statistics, which is a measure how often a specific word and any other word are observed together in a larger text. High co-occurrence can indicate similarity in meaning. It is claimed that early simulations offered adequate replication of word reading results (Kello, 2006). In further tests of this model, however, it failed to simulate results with semantic variables from a multi-syllable word reading regression analysis (Yap & Balota, 2009). The junction model of word reading includes a contribution from semantic memory to word reading, but it requires further development, and is therefore not discussed further.

[3] Semantic information is modelled using semantic features, and semantic features are described in Chapters 8 and 9.

There is one single semantic system; see Section 2.4.5.2[4]. The three systems are not directly connected to one another (Plaut et al., 1996); instead, the systems are connected through sets of hidden units (represented with empty ovals in Figure 1.2). Hidden units are an additional layer of units that allow for more complex patterns of information to be represented, in a distributed fashion (described below) than would be possible if the systems were directly linked to one another. For example, the hidden units between orthography and semantic memory receive an input from orthographic units, and send an output forward to semantic units and also, if the model has a recurrent rather than a strictly feed forward architecture, backward to orthographic units. When simulating word reading, activation in phonological units gradually builds up from activation inputs from the "intimately related" (Plaut et al., 1996, p.108) phonological (orthography-to-phonology) and semantic (orthography-to-semantics-to-phonology) pathways. Though the triangle model is in fact a two-route model, these pathways are distinctly different from the two routes of the DRC and CDP models. In DRC and CDP simulations, the two routes compute a response by completely different computational principles; certain types of letter strings can only result in a correct response by one route or the other. In the triangle model, there is a single type of procedure for computing phonology: the summation of inputs to phonological units, whether these inputs come straight from orthography and the hidden units connecting orthography to phonology, or via the semantic system.

[4] Harm and Seidenberg (2004) argue that the semantic information one learns as a child is the same semantic information that is linked into orthography and phonology As language is learned in speech form (phonology), as when a child learns to speak, it is linked to the already established semantic information in order to comprehend words. In this way connections between phonology and semantics are formed. As language is learned in print form (orthography), as when a child learns to read and write, this is linked to the semantic information and phonology that have been learned previously.

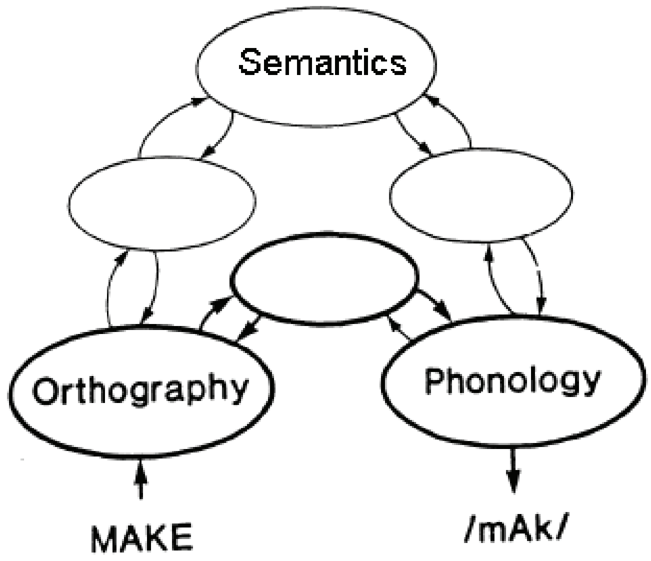

Figure1.2. The Connectionist Triangle Model. In this single route model, orthography, phonology, and semantic memory (meaning) are fully interconnected systems that work together to compute the correct phonology for all words. Semantic memory is implemented in this model. Copyright © 1989 by the American Psychological Association. Adapted with permission. The official citation that should be used in referencing this material is Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568. The use of APA information does not imply endorsement by the APA.

1.3.3.1 Distributed representations

In the triangle model, orthographic, phonological, and semantic forms are represented in a distributed fashion. Distributed representations are used to represent information through a complex pattern of activity over the sets of units, including the hidden units. With this type of representation, a small number of units can create a large number of patterns. A metaphor that can be used to understand distributed representations is to liken it to the illuminated time on a digital clock (Le Voi, 2005). Numbers on a digital clock are not represented by one large whole number. Instead different segments are illuminated or not illuminated in unison to create a pattern, and this pattern forms a number that constitutes the correct time. So, for example, in distributed orthographic representations, there will be considerable overlap between the set of orthographic units activated for the word "pint" and the units activated for words like "pin", "pine", “hint”, “print” etc. This contrasts with orthographic representations in the DRC model, where there is a single node in the orthographic input lexicon for the whole word “pint”, with nothing to indicate its similarity of spelling to “pin”, “pine”, “hint”, and “print”.

1.3.3.2 Word reading in the triangle model

In the triangle model, word reading is simulated in the following way. A word’s orthographic form (initial input pattern) produces a pattern of activity in the orthographic units, which represents the word in letter form. Then, via weighted connections and hidden units, activation spreads to the associated semantic pattern (orthography-to-semantics) and to the associated phonology (final output pattern) (orthography-to-phonology and orthography-to-semantics-to-phonology). The output pattern for a word is computed over time as the model settles into a stable pattern of phonological activation that builds from both the phonological pathway and the semantic pathway.

1.3.3.3. Weighted connections

As mentioned previously, connections within the model are weighted. These weights determine the strength of the activation that flows along the connections and, during training of the model, the weights are gradually adjusted using a learning algorithm and backpropagation with error correction. Through training on a large corpus, the weights on connections between the specific input and correct output patterns are gradually strengthened and connections to incorrect patterns of output are gradually weakened. For example, with the word “pint”, the model is concurrently learning to pronounce the 12 other English words with the orthographic pattern “_int” for which the correct phonology includes a short vowel “I” (“hint”, “lint”, “mint”, “tint”, “dint”, “print”, “stint”, “sprint” etc.). Therefore many training trials on “pint” with back-propagated error will be required to shift the pattern of phonological activation from the short “I” of “mint” etc. to the long “I” of “pint”.

An example is now described using the orthography-to-phonology pathway, but these principles can also be applied to the semantic pathway within the triangle model. To train the model, it is presented with letter strings which are locked onto the appropriate grapheme units. These grapheme units feed activation forward to the hidden units, and these hidden units then feed forward activation to the phonological units. In this way the model computes an initial phonological code. This phonological pattern of activation is then compared to the correct target phonological pattern (provided by the modellers), and the amount of difference between the two phonological patterns, called error, is computed by the model. The model uses the error to adjust the weights backwards through the connections (backpropagation); for example, some connection weights between the phonological units and the hidden units will be strengthened, while others will be weakened. Through many iterations of this feed forward and feed backward process, the model gradually adjusts its weights across the units until the connections produce optimal performance, i.e. the phonological output of the model matches the intended target phonology better than it matches any other possible phonological output. After training, the model uses this weighted set of connections to simulate reading aloud when it is presented with any letter string, whether or not the orthographic form was present in the original training corpus. These subsequent presentations can also alter the weights of the connections within and across the sets of units, using the same backward and forward connections as were used in training, though these would likely not affect the weights as much as the original training.

When a stable and settled pattern of phonological activation is reached repeatedly, as occurs over training, an attractor basin in the phonological system is formed by the increased weights amongst the phonological units of the correct response. In other words, the long "I" of "pint" gradually becomes the favoured pronunciation, not only because of the activation arriving from orthography via the hidden units but also because of the activation from the phonemes corresponding to the other elements of "pint". The attractor basin for "pint", however, will be shallow, because "pint" is both a lower frequency word (not many training exposures) and an exception word. Words like "hint", "mint", and "print" will have deeper attractor basins because they all reinforce one another. Before the completion of training, or even after this if there is noise in the system, "pint" will always be in danger of capture by the incorrect, but much stronger, attractor basins for the 'enemy neighbourhood' of words with a short "I".

1.3.3.4. The triangle model and semantic information

Semantic information contributes to the computation of phonology with activation flowing via weighted connections from orthography-to-semantics-to-phonology (see Figure 1.2). For example, despite the large enemy neighbourhood of "save", "gave", "pave", "brave "etc., a high-frequency exception word like "have" will develop its own deep attractor basin simply because it occurs so frequently; but a low-frequency exception word like "pint" will always be at risk of influence from the phonology of its many conflicting neighbours. The activation of the meaning of "pint", however, in the semantic component of the model, will send activation to the correct long "I" of its pronunciation. In this way, semantic knowledge can contribute to the process of settling on the correct pronunciation when the orthography-to-phonology computation is less efficient (Seidenberg & McClelland, 1989).

Likewise the triangle model also uses clean-up units to settle noisy or partial patterns of activation. These clean-up units enhance the attractor basins within the latest instantiation of the triangle model (Harm & Seidenberg, 2004). For example, all semantic units are connected to clean-up units and the clean-up units are connected back to the semantic units. Therefore any partial activation of semantic units is passed to the clean-up units and the clean-up units then alter this partial or noisy activation by feeding activation back to the semantic units, pushing the activation into a stable, settled pattern in the nearest attractor basin.

1.3.3.5. Division of labor

Both pathways can contribute to the computations of phonology; but since connections are weighted, one pathway might be more efficient than the other under certain circumstances or for certain word types, resulting in a division of labor [sic] (Plaut et al., 1996; Strain & Herdman, 1999). For example, when direct orthography-to-phonology computation is less efficient, as with low-frequency exception words like “pint”, a significant contribution from the semantic representations will be more likely because the settling process on the direct pathway occurs more slowly. For other word types, e.g. the high frequency regular and consistent word “hand”, orthography-to-phonology will be particularly efficient, giving little time and opportunity for a semantic information contribution. Any activation flowing between semantic memory and phonology would likely occur after the phonological units had reached a settled level of activation; therefore semantic information would be activated but would not contribute to the response.

Plaut and colleagues emphasise the importance of all three systems in normal reading. They write, "This [model's simulation] supports a view of the normal word reading system in which there is a division of labor[sic] between the phonological and semantic pathways such that neither pathway alone is completely competent and the two must work together to support skilled word and nonword reading" (Plaut et al., 1996, p.99).

1.3.3.6. Frequency and consistency in the triangle model

In the triangle model, the effects of frequency and consistency are simulated by the strength and efficiency of the connections (Seidenberg & McClelland, 1989; Plaut et al., 1996). If a word is highly frequent, then because of a higher number of exposures, the appropriate connections gradually become stronger. Frequency affects not only the connections between orthography and phonology, but also orthography and semantics, and semantics and phonology (Harm & Seidenberg, 2004). Consistency of spelling-to-sound correspondence also affects the strength of the connections between orthography and phonology. If a word has many "friends", and the "friends" are highly frequent as well, connections work in favour of the word's correct pronunciation. If a word has many "enemies", then translation from print-to-sound is less efficient because connections are weighted for the alternative pronunciation of the letters.

1.3.3.7 Conclusions on the triangle model

Triangle modellers claim that an account of reading that includes a semantic component accounts for human performance more successfully than models that do not allow for it, such as the DRC model (Plaut et al., 1996). They argue that the triangle model, despite its simplicity of processing all words via the same computational principles, better simulates the reading performance observed in healthy and impaired word reading than the DRC model. In particular, the triangle model can account for the dramatic frequency-by-consistency interaction observed in the reading aloud of people with semantic dementia simply with reference to the patients' central deficit to semantic knowledge, for which there is overwhelming evidence; the DRC model, by contrast, assumes additional deficits in the orthographic input lexicon or the phonological output lexicon or both (Section 2.2.1), such as the frequency by consistency interaction, i.e. low frequency inconsistent words are responded to more slowly than high frequency consistent words, and word reading performance of patients with a semantic impairment, respectively. Furthermore, an account of the graded consistency effects, as opposed to the dichotomy of regular and exception words, is a logical effect in the triangle model, but requires a much more complex explanation in the DRC model.

Within the triangle model, it is largely a matter of a word's orthography-to-phonology efficiency as to whether semantic information has a chance to contribute to the settling of phonology. In the investigations of this thesis, special attention will be paid to the most difficult word types because, if semantic memory is involved in healthy word reading, then, according to the triangle model, less efficient words are most likely to provide time enough for this influence to occur. For the purposes of the current research, the studies presented in subsequent chapters are concerned with investigating a contribution of semantic information before final correct output phonology is achieved, whether this is as a result of orthography-to-semantics-to-phonology, or orthography-to-phonology-to-semantics-to-phonology interaction. In the experiments of Chapters 3-6 and 8, the marker of stable phonological activation is the correct reading aloud of the word stimuli. Evidence of semantic information influencing word reading times will constitute evidence of a semantic contribution to word reading.

1.3.4 Conclusions on models of word reading

The main computational models of word reading differ in their inclusion of a semantic contribution to orthography-to-phonology computation, reflecting the continuing debate within the human empirical literature. The triangle model of word reading includes an implemented semantic system and allow for a semantic contribution to word reading, while DRC, and the CDP+ and CDP++ models, do not have an implemented semantic memory system and claim it is not involved in orthography-to-phonology computation.

The results from the research of this thesis will be discussed in terms of the current computational models (see Chapter 10). The next chapter (Chapter 2) reviews the empirical literature from the word reading domain that has helped to shape the controversy involving orthography-to-phonology computation and has informed the computation models reviewed in the current chapter. The next chapter's review also helps to identify converging methods that can be used to investigate the central aim of this thesis, and aids in establishing specific additional research aims.

Chapter 2

A Literature Review of Behavioural Empirical Research

2.1 Introduction

This chapter reviews behavioural empirical research that has used factorial designs with orthogonal manipulations, smaller scale regression studies[5], and semantic priming designs to investigate a semantic contribution to word reading in healthy readers, as it is pertinent to the central research aim. The review of the semantic priming literature directly applies to the first set of behavioural experimental studies (Chapters 3 to 7). In Chapter 8 research directly relevant to a larger scale regression study will be reviewed in detail prior to the report of regression investigations, although a brief summary of the regression literature will be included here. Similarly, in Chapter 9, research relevant to the neurophysiological experiments of this thesis will be reviewed, but very briefly anticipated here. As will be shown, this research, however, fails to provide a definitive conclusion as to whether there is a semantic contribution to word reading, and it highlights the need for further research.

[5] The term "smaller scale" here is used to describe investigations that have performed regression analyses with a small number of items; these are often performed in conjunction with investigations using factorial manipulations.

2.2 A brief review of three research areas

Before reviewing in detail directly relevant behavioural literature, other areas of research that are pertinent to the topic of this thesis will be briefly summarized for completeness. Three areas of note are: (a) word reading in adults with semantic dementia, (b) research using lexical decision task in a healthy population, and (c) research using ERP methodology to investigate word reading in a healthy population.

2.2.1 Word reading in adults with semantic dementia

An area of research related to the aim of this thesis is word reading in people with semantic dementia - a form of dementia that results in progressive damage to the anterior temporal lobes of the brain and loss of semantic knowledge. This patient group often demonstrate impaired word reading (Hodges, Patterson, Oxbury & Funnell, 1992; Jefferies, et al. 2004; Patterson & Hodges, 1992; Patterson et al. 2006; Warrington, 1975; Woollams, Lambon-Ralph, Plaut & Patterson, 2007), struggling to read difficult words correctly, especially low frequency exception words (Patterson & Hodges, 1992). For example, exception words are regularised, e.g., “brooch” is read to rhyme with “pooch”. Whether this reading deficit is attributed to a semantic impairment or non-semantic impairments fosters the debate concerning a semantic contribution to orthography-to-phonology formation and whether a computation model adequately simulate their impaired word reading is often a performance benchmark for computational models of word reading.

Since semantic memory is impaired in this population and word reading performance is affected, some researchers maintain that semantic memory is therefore implicated as contributing to correct word reading, especially when orthography-to-phonology is not efficient, as with low frequency exception words (Graham, Hodges, & Patterson, 1994; Patterson & Hodges, 1992; Patterson et al., 2006; Woollams, Cooper-Pye, Hodges, & Patterson, 2008; Woollams et al., 2007). Moreover, evidence suggests that the greater the impairment to semantic memory, the more impairment there is to word reading performance, with the largest impairment for low frequency exception words (Patterson & Hodges, 1992; Woollams et al., 2007).

A recent investigation of reading performance in people with semantic dementia used observations from a large number of people over the progression of the disease (Woollams et al., 2007). Through the analysis of word reading errors and computational model simulations, low frequency words and exception words were more likely to be read incorrectly as semantic knowledge was increasingly impaired. It was concluded that low frequency exception word reading is reliant on a semantic contribution. Moreover when reading is less efficient low frequency regular words and high frequency exception words may also receive some contribution from semantic memory, since specific errors increased with loss of semantic knowledge.

Not all researchers, however, agree that semantic memory deficits are responsible for word reading deficits in the semantic dementia population (Coltheart, 2004; Coltheart, Tree, Saunders, 2010). Intact word reading in single case studies of individuals with semantic dementia has been found (Blazely et al., 2005). Because these cases show a dissociation between impaired semantic memory (as measured on standardized tests) and word reading, it is argued that though there is damage to the semantic system, this is not responsible for the poor word reading performance reported in this patient group by other researchers, and consequently semantic memory is not needed to correctly read a word (Blazely et al., 2005; Coltheart 2006a; Coltheart, 2006b Coltheart et al. 1993; Coltheart et al., 2001; Coltheart, Saunders, & Tree, 2010; Coltheart, Tree, et al., 2010). Instead researchers claim that individuals with semantic dementia who show impaired word reading must also have damage to other brain regions near to the anterior temporal lobes, which support reading (Coltheart, 2004; Coltheart, Saunders, et al., 2010; Coltheart, Tree, et al., 2010).

Impaired word reading performance of people with semantic dementia is one domain that influences the development of word reading computational models, and whether a semantic memory contribution to orthography-to-phonology transformation is emphasised or eliminated depends on the interpretation of this data (Coltheart et al., 2001; Coltheart 2006a; Coltheart, Tree, et al. 2010; Patterson et al. 2006; Plaut et al. 1996; Woollams et al., 2007). Even though healthy individuals are the focus of this thesis, the data from other populations highlight the controversy as to whether semantic information contributes to word reading.

2.2.2 Research using lexical decision

Results from research using lexical decision tasks have also informed the understanding of orthography-to-phonology computation. Lexical decision is a paradigm in which participants decide, by responding with a button press, whether or not a stimulus item is real. With word stimuli, participants decide whether a letter string forms a real word or a nonword. Used in conjunction with reading time data, lexical decision data can be a strong benchmark for computational model simulations (Balota et al., 2004). However, lexical decision may reflect decision processes (Carreiras, Michelli, Estevez, & Price, 2007; Tyler, Voice, & Moss, 2000). Performance in lexical decision has also been accounted for through the matching of a stimulus to a stored lexical representation; this could mean the use exclusively of stored orthographic and/or phonological representations (Andrews, 1982; Carreiras et al., 2007; Damian & Als, 2005; Joordens & Becker, 1997), and/or the use of semantic representations (Balota, et al., 2004 Carreiras et al., 2007; Hauk, Patterson, Woollams, Watling, Pulvermuller & Rogers, 2006) which may bias the participant towards the use of semantic memory for matching purposes (i.e. if there is a semantic representation for this item, then it must be a real word). Therefore, while this task has clearly contributed to understanding some aspects of word recognition, lexical decision may not prove particularly useful in understanding the role of a semantic memory contribution to orthography-to-phonology computation. This large literature will not be reviewed, though where relevant points can be made (such as in the semantic priming or ERP reviews) this literature will be referred to.

2.2.3 Research using ERP methodology

Electroencephalography (EEG) has been used to investigate the time course of word reading, including a semantic memory contribution to the process. Electrical potentials from neurons in the brain (EEG) are measured and time-locked to stimulus presentation to create event-related potentials (ERPs) (Coles & Rugg, 1996; Handy, 2005; Luck, 2005a). ERPs reveal neural activity that occurs in response to stimuli, including neural processes that are involved in word reading (Coles & Rugg, 1996; Handy, 2005; Luck, 2005a; Rugg & Coles, 1996). ERP methodology is of particular interest to the current thesis.

Effects that may reflect phonological processing have been found in the time course, at around 200ms after stimulus presentation (Abdullaev & Posner, 1998; Bentin, Mouchetant-Rostaing, Giard, Echallier & Pernier, 1999; Dien, 2009; Grainger, Kiyonaga, & Holcomb, 2006; Hauk & Pulvermuller, 2004; Hauk, Patterson, et al. 2006; Proverbio, Vecchi & Zani, 2004; Sereno, Rayner, & Posner, 1998). To show that semantic information has the potential to contribute to word reading, evidence of a semantic variable's influence on ERP data would need to be shown prior to this time point[6]. ERP, as it is sensitive to time course, will be particularly useful in this aim (Handy, 2005; Luck, 2005a; Otten & Rugg, 2005).

The ERP literature has shown semantic effects early in the time course of word reading, as early as 150ms (Abdullaev & Posner, 1998; Hauk, Davis, Ford, Pulvermuller, & Marlsen-Wilson, 2006; Landi & Perfetti, 2007). Further effects of semantic variables, specifically semantic features, have been found around 200ms post-stimulus (Kounios et al., 2009). However, the studies noted here used lexical decision, delayed naming, or tasks that emphasise the use of semantic memory. Of interest is whether similar early semantic effects can be found using the semantic measures imageability and semantic features in two experiments, respectively, using a task that is less likely to involve complex decision processes and may not explicitly bias word reading via semantic memory, such as silent word reading. Chapter 9 presents a full review of the relevant ERP literature and presents the final studies of this thesis, which use ERP methodology to investigate potential early semantic activation effects.

[6] Semantic processing using ERPs is often studied by focusing on data from around 400ms post-stimulus (N400). This data may be linked to context, comprehension, and integration of total semantic information (for a review see: Hinojosa, Martin-Loeches, and Rubia, 2001).

2.3 A literature review of behavioural empirical research that has used factorial manipulations and regression methods

The remainder of the literature review focuses on studies concerning a semantic contribution to orthography-to-phonology within a healthy population using word reading. The review is divided into three sections (a) factorial manipulations and regression methods, (b) larger scale regression studies, and (c) semantic priming paradigms.

2.3.1 Semantic variables

There are several ways in which semantic memory representation of a word can be measured and quantified. For example, it can be measured using imageability (the strength of the mental image of the item), concreteness (whether the word refers to a definite object), ambiguity (the number of divergent meanings of a word), number of semantic neighbours (number of items that are similar in meaning), or number of semantic features (the number of descriptive features that are listed when defining an item). Though some of these measures are subjective, such as imageability, others could be considered more objective, such as number of semantic features. These measures however are not infallible. Using subjective measures, it is not always clear as to what participants are rating. More objective measures give reassurance that one aspect only is being quantified, but some are still listed by participants, e.g., semantic features. A rich semantic representation may be best accounted for by more than one semantic measure (Strain et al. 1995). As will be shown, within the investigations of this thesis more than one measure of semantics is used.

Semantic measures are tools researchers use to study semantic information, including studying a semantic contribution to orthography-to-phonology computation. Items that have richer representations in semantic memory may measure as high in a semantic measure, such as imageability, concreteness, semantic features, or semantic neighbours. To explain, the following paragraph describes an example using the semantic measure of imageability.

A word such as "vest" is high in imageability as participants usually rate it as easy to conjure a mental image of this item. It is more difficult to call to mind a mental image for "vain", which is judged by participants as low in imageability. Strain et al. (1995) give three explanations as to why a highly imageable word might have a rich semantic representation, and these are described here. Firstly, a highly imageable word might contain its own inherent meaning that is represented in semantic memory and therefore may not be reliant on sentence context for meaning. In contrast a low imageable word would be reliant on its sentence context for a rich meaning. Secondly, a highly imageable word may have a greater number of semantic features than a low imageability word. An example can be found in describing a "vest", which is a non-living, item of clothing, worn on the upper half of the body. To describe a low imageable item like "vain" there is no definite list of features. In this way the semantic measure semantic features may make up part of what is being measured by imageability; this is discussed further in Chapter 8. Thirdly, a highly imageable item may also have more connections to other parts of the brain. For example, "vest" may connect beyond the visual image, such as to regions of the brain that store sensory images. "Vain" may not be connected in such a way. The way in which imageable items are semantically rich is transferable to other semantic measures as well.

Using semantic measures in conjunction with word types may offer the opportunity to research a semantic contribution to word reading. As described in Section 1.3.3, in the triangle model a semantic contribution may only have time to contribute when orthography-to-phonology is less efficient, e.g., low frequency exception words. For example, consider a low frequency exception word that is semantically rich, such as “pear”. Within the triangle model of word reading, orthography is activated and processing proceeds to phonology; semantic memory may be involved either through orthography-to-semantics-to-phonology and/or through orthography-to-phonology-to-semantics-to-phonology (Plaut et al., 1996; Harm & Seidenberg, 2004). As orthography-to-phonology for “pear” is less efficient and slow - this is known from human performance data and word reading models (Andrews, 1982; Paap & Noel, 1991, Seidenberg, 1985; Seidenberg & McClelland, 1989; Seidenberg et al., 1984) - its rich semantic representation could provide information. The word “pear” might therefore be read faster than a low frequency exception word that is semantically poor, such as “dose”, because of this rich semantic contribution. In this way, word types and a semantic measure can be used to study a semantic contribution to orthography-to-phonology computation.

2.3.2 A review of factorial studies and small scale regression analyses

In the literature, there is research using factorial manipulation and smaller scale regression analyses that employ semantic variables and word types (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002; Shibahara et al., 2003; Strain et al., 1995, 2002; Woollams, 2005). Additionally, there are research projects that use both factorial and small scale regression designs, as complementary methods (Strain et al. 2002; Monaghan & Ellis, 2002; Strain & Herdman, 1999). Factorial methods attempt to control stimuli on a number of factors while carefully manipulating them on a limited number of factors of interest. This may lead to having a restricted number of stimuli. Additionally, controlling stimuli on all factors can be difficult. A regression study uses a large number of stimuli that vary along a continuum on an assortment of measurements. The impact of each independent variable on the dependent variable, usually reading or lexical decision reaction times, is accounted for in the regression analysis while statistically controlling for the other factors. A recent article, which uses regression alone, champions the strength of using the two methods together, “It is likely that the two approaches [factorial and regression] will provide complementary constraints on [word reading] theory development” (Balota et al., 2004, p. 313). Both approaches are used within the investigations of this thesis as complementary methods, and these experiments are described and discussed in subsequent empirical chapters. Research using factorial methods and/or smaller scale regression studies that are relevant to this thesis are now presented.

Investigations by Strain et al. (1995) form a cornerstone for research into a possible contribution of semantic memory in healthy word reading, and have been built on by others (e.g., Balota et al., 2004; Coltheart et al. 2001; Harm & Seidenberg, 2004; Seidenberg & Herdman, 1999; Shibahara et al., 2003; Woollams, 2005). Strain et al. investigated whether semantic information is involved in the computation of phonology. They factorially manipulated word stimuli by imageability, called imagery by Paivio (1971), and word type (Section 1.2.3), and measured word reading times and word reading error rates of healthy adults.

The first experiment of this series used an orthogonal manipulation of frequency, regularity, and imageability, creating eight word difficulty types and these are listed from easiest to most difficult: (1)high frequency regular high imageability, (2)high frequency regular low imageability, (3)high frequency exception high imageability, (4)high frequency exception low imageability, (5)low frequency regular high imageability, (6)low frequency regular low imageability, (7)low frequency exception high imageability, and (8)low frequency exception low imageability. Results of Strain et al.'s (1995) first experiment revealed significant main effects of regularity, frequency, and imageability in reaction time analyses. High imageability words were read faster than low imageability words. Low frequency exception words showed the largest imageability effect, but the predicted three-way interaction of frequency, regularity, and imageability only revealed a trend towards significance. In Strain et al.'s second experiment, which only used the low frequency types listed above, there was a significant interaction between regularity and imageability; that is, there was a significant impact of regularity on reaction times for low but not for high imageability words. In keeping with this, high imageability low frequency exception words were read faster than low imageability words of the same type. These, in both of the first two experiments, 'difficult' words (low frequency exception words) that were high in imageability were read more quickly than their low imageability counterparts. This likely occurred because (1) low frequency exception words have less efficient orthography-to-phonology computation and therefore provided enough time for additional information to contribute to the phonological computation and (2) the additional semantic information inherent in high imageability words contributed to the computation of a phonological response, giving the items an advantage, as compared to low imageability exception words. From these results, it was concluded that imageability makes difficult words easier to read. Therefore the speed of low frequency exception word reading was 'boosted' when this word type was high in imageability as opposed to low in imageability.

In the third and final experiment of this series Strain et al. turned their attention to word reading errors, notably regularization errors (e.g., reading “sew” to rhyme with “new”); the goal was to seek converging evidence for a semantic contribution, from the factor of imageability in the reading of difficult word types. If previous observed beneficial semantic imageability effects are reliant on having enough time during orthography-to-phonology computation, then ‘forcing’(actually encouraging via a deadline procedure) the participants to make their reading responses quickly should reduce the semantic contribution and therefore increase the number of errors, as compared to ordinary (non-speeded) reading. Words with efficient orthography-to-phonology links, such as low frequency regular words, should not be affected by this speeding, as under normal reading circumstances the orthography-to-phonology computation for this word type is sufficiently efficient that any semantic information does not have time to contribution to, and does not need to, contribute to the computation of a response. Indeed, in Experiment 3, low frequency regular words were not affected by the response deadline: that is they showed no increase in error rate as compared to the previous experiment where reading was not speeded. However, words with less efficient orthography-to-phonology characteristics also now had to be read quickly. Words that were previously ‘boosted’ by semantic information, i.e., high imageability low frequency exception words, should now show an increase in errors, as compared to Experiment 2, because semantic information no longer has time to contribute to activating the correct response. This prediction was also confirmed: high imageability low frequency exception words showed an increase in reading errors, relative to non-speeded reading in Experiment 2.

Based on these results, Strain et al. (1995) concluded that imageability (a semantic measure) interacts with word difficulty, providing evidence that semantics knowledge can contribute to word reading in healthy adults. Moreover, since phonology is computed in the same way for all word types within the triangle connectionist mode, but only the most difficult words show the semantic imageability effect, Strain et al. argued that although semantic information is probably accessed for all word types automatically, it is only with more slowly processed words that semantic information has the opportunity to make a real contribution to computing a phonological response. This conclusion has not gone unchallenged;, criticisms of it will be reviewed in Section 2.3.2.1.

Strain and Herdman (1999) extended the behavioural work of Strain et al. (1995) and the computational modelling word of Plaut et al. (1996) using factorial designs and regression methods to investigate a semantic memory involvement when orthography-to-phonology is less efficient. In the factorial designs of these investigations, reading performance, as measured with the reaction times of healthy adult readers, were assessed and used to measure individual differences in efficiency of orthography-to-phonology computation. In good readers orthography-to-phonology might be efficient; therefore semantic information would not have time to contribute to phonology formation. In contrast, poorer readers might have less efficient orthography-to-phonology computation skills, making them slower readers, allowing time for a semantic memory contribution to phonology computations.

In Strain and Herdman's (1999) investigations, participants were divided into reading skill groups (high, medium, low) and they read low frequency words manipulated on regularity and imageability, from Experiments 2 and 3 of Strain et al. (1995). Results from all participants and from high skilled readers alone matched those of Strain et al., i.e., an imageability effect for low frequency exception words only. However, for low ability readers, with supposed inefficient orthography-to-phonology mappings, there was instead an imageability effect with low frequency regular and exception words. In the second experiment, Strain and Herdman used regression methods to study the possible evidence for a semantic contribution to word reading as may have been evidenced in Experiment 1. Lexical predictors[7] and imageability, a semantic predictor, were entered as predictors of word reading times. In addition to other significant measures, imageability was also a significant predictor of the variance in word reading reaction times. A significant interaction of imageability and regularity revealed that imageability was a significant predictor for exception word reading times, but not for regular words. Moreover, reading times differed for low imageability exception and regular words, but there was no significant difference for high imageable words.

Because imageability effects are found when word types are more difficult (all readers in experiment 1 and prediction of word reading times experiment 2) or when reading ability is poor (low skill readers in experiment 1), Strain and Herdman concluded that a semantic contribution to word reading occurs when orthography-to-phonology computation is less efficient whether due to word factors or reader skill. The conclusions are in keeping with those of Strain et al, yet, like Strain et al. there were critiques to this study, which are discussed in a subsequent section.

There is other research in the literature that has also investigated imageability effects. DeGroot (1989) investigated the storage and organisation of semantic memory. Experiment 5 might provide evidence of a semantic contribution to word reading. She measured word reading reaction times of healthy adults in a using factorial manipulation of frequency (high and low) and imageability (high and low). Main effects

[7] Lexical predictors are measures that might capture orthography and phonology properties of a word. In this study they were initial phoneme, frequency, length, bigram frequency, and regularity.

of frequency and imageability were found, and there was no significant interaction of the two. Highly imageable words were read faster than low imageable words. Since imageability is thought to be a semantic variable, this study may indicate a semantic contribution to word reading. However, unexpectedly, there appears to be an imageability effect for both word types, including the more efficient high frequency words. However, this study does not fully investigate the word types as it does not control or manipulate regularity type, and with this additional manipulation the results may differ. Therefore though these results hint of a semantic contribution to all word frequency types, the results may be limited to the sample of words used in this experiment.

More recently, whether the factor of imageability is truly a measure of semantic richness has been investigated, and these studies also provide evidence of imageability effects during word reading (Woollams, 2005). Whether the measure ambiguity is a semantic measure was also investigated in this series, though the focus of this brief review will be imageability. In Experiment 1, the reaction times of healthy adults were measured as they read well matched single word stimuli that were low in frequency and manipulated on regularity (regular or exception) and imageability (high or low) or ambiguity (ambiguous, with many divergent meanings, or unambiguous with one consistent meaning). Significant effects of imageability and ambiguity in low frequency exception word reading were found, replicating benchmark effects of previous studies, including Strain et al. (1995; 2002). High imageability words and ambiguous words were read faster than their low imageability or unambiguous counterparts. In Experiment 2, similar to the final experiment of Strain et al. (1995), Woollams sought to eliminate semantic effects of Experiment 1. She used this as a way of identifying effects as being truly semantic in nature, as opposed to capturing orthographic or phonological properties of a word. Matched nonwords, which are likely read using orthography-to-phonology connections, as they do not have semantic representations, were included with the list of original stimuli and healthy adult participant reaction times were measured during this task. By mixing this non-word stimuli with the real word stimuli all reading, even real word reading, will be biased to primarily use orthography-to-phonology connections. If the imageability and ambiguity effects of Experiment 1 are semantic in nature, and all stimuli in Experiment 2 are read primarily via the orthography-to-phonology pathway, then these semantic effects should be eliminated. Word reading reaction times revealed only a significant ambiguity effect with low frequency exception words, but no imageability effect with any word type. Since the imageability effect from Experiment 1was eliminated in Experiment 2 in these circumstances when the stimuli were read using primarily orthography-to-phonology connections, the author argues that imageability is a factor that measures semantic information. Whereas, ambiguity, as this effect remains, is not capturing semantic information, but more likely captures orthographic and phonological information. Because imageability affects low frequency exception word reading in Experiment 1, and from Experiment 2 results imageability is said to measure semantic information, this research also suggests a semantic contribution to word reading.

2.3.2.1 Criticisms of imageability effects

Though the previous review section has suggested a semantic contribution to word reading through significant imageability effects, there have been criticisms as to whether these effects implicate this. The primary criticisms have been levied by Monaghan and Ellis (2002) and Ellis and Monaghan (2002) on the work of Strain et al. (1995, 2002) and Strain and Herdman (1999). Monaghan and Ellis and Ellis and

Monaghan presented data that implicated an age-of-acquisition interaction with the orthography-to-phonology computation instead of imageability.

Age-of acquisition refers to the age at which a word is learned. This can be measured by observing children and noting when words are added to their vocabulary (Gilhooly & Logie, 1980; Morrison, Chappell & Ellis, 1997), or measured by asking adults to rate when a word was learned (Bird, Franklin, & Howard, 2001; Cortese, & Khanna, 2008; Gilhooly & Logie, 1980; Morrison et al., 1997). High age-of-acquisition refers an older acquisition age, and low age-of-acquisition refers a younger acquisition age. Age-of-acquisition and imageability are negatively correlated; high imageability words are learned at a younger age than low imageability words. Moreover, whether or not age-of-acquisition is a semantic measure is now unclear[8]. However, in this account, age-of acquisition was regarded as affecting mappings between orthography-to-phonology connections[9]. If age-of-acquisition is a measure of orthography-to-phonology connections and previously published imageability effects are instead due to uncontrolled age-of-acquisition differences between the conditions, then previous claims of a semantic contribution to word reading could be nullified.

Monaghan and Ellis (2002) measured word reading times of healthy adult readers across a series of studies. The stimuli were single words that had been factorially manipulated on word type (frequency (high or low) and regularity (regular or exception)), while controlling or manipulating age-of-acquisition (high or low) and

[8] See Chapter 8 for a detailed account of age-of-acquisition. Other interpretations of this variable do not eliminate the possibility that it is, in part, a semantic variable. However, for the purposes of the debate, Ellis and Monaghan (2002) and Monaghan and Ellis (2002) considered age-of-acquisition to be capturing orthographic and phonological properties of words.

[9] Chapter 8 discusses whether age-of-acquisition affects arbitrary connections most (Lambon-Ralph & Ehsan, 2006). Connections from orthography-to-semantics and phonology-to-semantics are more arbitrary than orthography-to-phonological connections. Therefore if age-of-acquisition affects arbitrary connections most, then it should also affect semantic connections.

imageability (high or low). Experiment 1 investigated whether the benchmark frequency by regularity interaction occurs even when stimuli are controlled on age-of-acquisition. Word stimuli were manipulated on frequency and regularity while matching on all other factors, including age-of-acquisition. The benchmark frequency by regularity interaction was found even when age-of-acquisition was controlled. This shows that frequency and regularity capture factors unique from age-of-acquisition and that low frequency exception words are, indeed, the most difficult of the four word types (Section 1.2.3), even when matched on age-of-acquisition. Experiment 2 investigated whether age-of-acquisition interacts with word spelling-to-sound. Words were manipulated on regularity and age-of-acquisition while matching on many other factors, including frequency and imageability. When regularity and age-of-acquisition are manipulated, there is a significant interaction of age-of-acquisition and regularity. Early (low) age-of-acquisition exception words were read more quickly than frequency matched late (high) age-of-acquisition exception words, but there is no significant difference between the regular words that differed in age-of acquisition. The authors conclude, therefore, that age-of-acquisition interacts with word spelling-to-sound correspondence, just as Strain et al. (1995) argue that imageability does. However, Monaghan and Ellis control for imageability while Strain et al did not control age-of-acquisition.

Of particular interest in this series are Experiments 3 and 4 as Monaghan and Ellis directly attempt to replicate the imageability effects of the Strain et al. (1995) series while controlling for age-of-acquisition. Experiment 3 used low frequency words manipulated on regularity and imageability (similar to Strain et al.'s (1995) Experiment 2) while controlling for age-of-acquisition during stimuli selection. This experiment failed to replicate the imageability by regularity interaction of Strain et al, i.e., high imageability low frequency exception words being read more quickly than low frequency exception words, but no such effect in low frequency regular word reading. In Strain et al. this interaction was central to the claim that semantic information, as captured by imageability, contributes to word reading especially with the most difficult word type. Experiment 4 replicated Strain et al.'s Experiment 2, directly, by using identical methods, design, and stimuli, so age-of-acquisition was not controlled for in the stimuli. Instead, effects of age-of-acquisition were statistically accounted for by co-varying out age-of-acquisition effects in the analyses. The results without controlling for age-of-acquisition were initially the same as Strain et al: there was an interaction of regularity and imageability in low frequency word reading with high imageability low frequency exception words being read more quickly than low imageability low frequency exception words, but this was not found for low frequency regular word reading. However, the imageability effect failed to remain significant once age-of-acquisition was co-varied out in the statistical analysis. When Monaghan and Ellis reanalyzed the original data from Strain et al. (1995) using the same parameters this was also true; no imageability effect was found in low frequency exception word reading.

These studies by Monaghan and Ellis (2002) appear to undermine results of Strain and colleagues (1995), which were previously thought to indicate a semantic contribution to word reading, as captured by imageability. Monaghan and Ellis concluded that there was not convincing support for a semantic contribution to word reading because age-of-acquisition effects, which are different from imageability and measure orthographic and phonological (non-semantic) properties, are found in low frequency exception word reading when imageability is controlled for, but the opposite is not true: imageability effects are not found in word reading, including low frequency exception word reading, when age-of-acquisition is controlled for. They write, “In sum we would argue that evidence for any semantic contribution to the rapid naming of any class of familiar written words is currently very weak….But when skilled readers name written words as quickly as possible, they appear to do so using direct mappings between orthography and phonology, even when the word in question is of low frequency and imageability" (Monaghan & Ellis, p. 192). Other published research that replicate and extend Strain et al.'s research, without controlling for age-of-acquisition, such as the work of Strain and Herdman (1999), were also subject to the same critiques of Monaghan and Ellis and Ellis and Monaghan (2002).

The debate continued in Strain et al. (2002) and Ellis and Monaghan (2002) with both using regression analyses of word reading reaction times to support their arguments. Strain et al. used regression analyses to further investigate an impact of imageability on word reading including age-of-acquisition using these and other factors as predictors on healthy adult word reading times. Results revealed that, among other factors imageability was a significant predictor of word reading times, whereas, age-of-acquisition was not. Ellis and Monaghan responded with their own similar regression analyses, claiming the previous regression results of Strain et al. were due to one outlier reaction time; when this was removed, imageability was no longer a predictor of reading times, but age-of-acquisition was. Both also commented on the interpretation of an interaction. This is discussed in the next section (2.3.2.2). This debate concluded with each side defending its original position. Whether the data demonstrate an imageability effect on spelling-to-sound typicality, and ultimately a semantic impact on word reading is still unclear (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002; Strain et al., 1995, 2002; Woollams, 2005).

Worth noting when attempting to draw conclusions about this research, is that in experiments where age-of-acquisition was not controlled, the research also has experiments that identified semantic measures. In Experiment 3 of Strain et al. (1995), speeded naming was used to reduce semantic effects in word reading. Subsequent to this manipulation, low frequency exception words that were labeled ‘high imageability’ no longer showed an advantage. However, Monaghan and Ellis have provided information that these ‘high imageability’ words were low (earlier, younger) in age-of-acquisition. However, if Strain et al.’s (1995) Experiment 3 indeed eliminated semantic effects, then even if the ‘imageability’ effect is due to age-of-acquisition, then the effects elimination indicates that age-of-acquisition is semantic in nature. Likewise Woollams et al. makes a similar argument. Her low frequency stimuli words were manipulated on imageability (and regularity), but not controlled on age-of-acquisition, similar to Strain et al. (1995). In Experiment 2 of her series, real word reading is mixed with non-word reading to eliminate semantic effects that may have been present in a previous experiment, and no imageability effects are found. Yet, if the imageability effects are actually due to uncontrolled age-of-acquisition and the effect is eliminated, then this indicates that age-of-acquisition is semantic in nature. Using this logic, Monaghan and Ellis (2002) and Ellis and Monaghan (2002) may be correct that age-of-acquisition affects word reading and interacts with word type, but they may be incorrect when claiming that age-of-acquisition effects do not indicate a semantic contribution to word reading. It is possible that age-of-acquisition is semantic in nature, and effects of age-of-acquisition on word type actually indicate, at least partially, a semantic contribution to word reading; this is discussed further in Chapter 8.

Important lessons can be learned from this literature and applied to the empirical research of this thesis. This literature demonstrates that the topic of a semantic contribution to orthography-to-phonology is relevant and unresolved, and it reinforces the need for further investigation. It also demonstrates that correlated measures are difficult to control in factorial manipulations, even for the most experienced researchers, but though it is difficult, it is important to control as many factors as is possible in order to make interpretation of significant effects simple. Likewise, when it is not possible to control all possible factors, there are complementary methods that can also be used, such as regression analyses that statistically control for confounded variables. Additionally this series shows how important it is to use a large number of items and predictors in a regression analysis. These lessons are applied in the empirical chapters of this thesis. For example, in Chapter 8 of this thesis the complementary method of regression is used to statistically control for possible confounded variables, and a large number of items (nearly double of that used in Strain et al., 2002) and predictors (both imageability and age-of-acquisition are included amongst others) are used.

2.3.2.2 Imageability effects and controlling age-of-acquisition

There have also been investigations into a semantic contribution to word reading that have controlled age-of-acquisition other than those discussed above. This series of experiments was performed in English and a language other than English with manipulation of regularity and imageability (Shibahara et al., 2003). Of note is that, in contrast to Strain et al. (1995), age-of-acquisition was controlled.

Experiments 1 and 2 of this series used the English language and replicated Experiments 1 and 2 of Strain et al. (1995) (Experiment 1 of this series) and Experiment 4 of Monaghan and Ellis (2002) (Experiment 2 of this series). Reaction times of healthy adult readers were measured when reading single word stimuli manipulated on frequency, regularity and imageability in Experiment 1 and regularity and imageability (low frequency words only) in Experiment 2. Results similar to Strain et al. were found;

an imageability effect was evident in low frequency exception word reading (when age-of-acquisition was not controlled), suggesting a semantic contribution to word reading. Moreover in response to the critiques of Monaghan and Ellis and Ellis and Monaghan (2002) age-of-acquisition was entered in the analysis as a covariant, similar to Experiment 4 of Monaghan and Ellis. A main effect of imageability was no longer significant, however, the significant interaction of imageability and regularity in low frequency words remained significant, just as Monaghan and Ellis found. The interpretation of an interaction without a main effect is discussed in the debate between Strain et al. (2002) and Monaghan and Ellis. Shibahara et al. and Strain et al. (2002) argue an interaction and no main effect is actually predicted by the triangle model in healthy word reading and should be interpreted. Shibahara et al. (2003) asserted, similar to Strain et al. (1995, 2002), that since imageability affects low frequency exception word reading (even when age-of-acquisition is statistically controlled for) semantic information contributes to orthography-to-phonology computation in the English language, especially when computations were slow and inefficient.

In Experiment 3, to investigate a semantic contribution to difficult words further, Shibahara et al. (2003) used the language Japanese Kanji, which is logographic or morphographic in nature. It has a more opaque orthography than English, meaning words are more difficult to read. If, according to the triangle model of word reading, difficult words use semantic information to compute phonology and words in this language are more difficult than quasi-regular English words, then the most difficult words in this language must rely on a semantic route to word reading. Japanese Kanji words were divided into three levels of word difficulty using regularity. The most difficult word type, low frequency exception Japanese-origin pronunciation words, showed an imageability effect, with high imageability words being read more quickly than low imageability words. This effect remained even when age-of-acquisition effects were removed statistically. Since an imageability effects was present in two languages, the authors conclude that semantic effects are present during word reading. Since these semantic imageability effects are present with difficult word types and are more reliable in the language with a more opaque orthography, the authors conclude that this semantic contribution to word reading is more likely with difficult words. Since the triangle model simulates a semantic contribution to orthography-to-phonology computation and this contribution is more likely when the orthography-to-phonology pathway is less efficient, the authors interpret their results as favouring the triangle model of word reading, though they also indicate that theoretically, at least, the DRC and CDP+ models could also accommodate these results.

2.3.2.3 Conclusions on factorial and small scale regression studies

Research that has used factorial designs with orthogonal manipulations and small scale regression methods has provided information concerning the topic of this thesis. The preceding review, however, demonstrated how this research can foster debate more than resolution. In this review some researchers concluded that there is evidence in favour of a semantic contribution to orthography-to-phonology computation (see Strain & Herdman, 1999; Shibahara et al. 2003, Strain et al., 1995, 2002; Woollams, 2005). Others argued the same evidence did not show a semantic contribution to word reading (see Ellis & Monaghan, 2002; Monaghan & Ellis, 2002). These studies have left the question of a semantic contribution to word reading unanswered.

2.3.3 A brief literature review of relevant larger scale regression investigations

To this point, the review has focused on research that used factorial designs and small scale regression analyses to investigate a semantic knowledge contribution to orthography-to-phonology computation. There are also studies that have used large scale regression studies alone to investigate word reading (Balota et al., 2004; Brown & Watson, 1987; Cortese & Khanna, 2007; Pexman, Lupker, & Hino, 2002). Large scale [10] regression studies offer a complementary design with which to study word reading. This type of study can statistically control correlated variables and can include semantic variables[11]. To pre-empt, the investigations reviewed in brief here and more thoroughly in Chapter 8 suggest that there is semantic memory contribution to word reading, but the conclusion is less than definitive. These studies ultimately do not resolve the issue. There remain valid investigations that could be performed using regression designs to investigate whether semantic information contributes to word reading. Therefore a larger scale regression study is conducted as part of the investigations of the current thesis to address relevant issues. In the following section, highlights of the research using regression methods are presented. An expanded review of the articles mentioned below and reviews of other relevant regression literature are available in Chapter 8.

Of note in the literature are the large scale regression analyses of Balota and colleagues (2004). They investigated the factors that influence lexical decision and word reading

[10] The term “large scale” is used to describe investigations that have primarily used regression methods alone with many predictor variables. This is a different set of studies to the smaller scale studies that have been used in conjunction with factorial designs though the methods are the same.

[11] As with studies using factorial methods, some investigations that use regression methods have been used to study word reading without including a semantic component. In these studies, only lexical components like orthography or phonology have been examined. For example Weekes (1997) only includes lexical variables (such as frequency and length) in a regression analysis. Work such is this, which do not include semantic variables, are not included in this review.

reaction times young and old healthy adults, using word reading reaction times that have recently been made available for research purposes (Balota et al., 2007). The English Lexicon Project (ELexicon) is large database of word reaction times, both reading and lexical decision, and lexical variables available online, and it is these that are used in the regression analyses of Balota et al. (2004). To briefly summarise, in addition to other variables, semantic measures were significant predictors of word reading reaction times, with Cortese and Fugett's imageability ratings being the most robust semantic measure of the six semantic measures entered into the analysis. Since measures that are thought of as semantic, i.e., imageability, were significant predictors of word reading reaction times Balota and colleagues conclude, "We believe these results are most consistent with a view in which meaning…contributes to the processes involved in reaching a sufficient level of information to drive a lexical decision or a naming response" (2004, p. 312). Results of this large regression support a semantic contribution to word reading, yet criticisms formed from results of other regression analyses have emerged (briefly summarized below).

Similar to the criticisms of Monaghan and Ellis (2002) and Ellis and Monaghan (2002) on the conclusions of Strain et al., (1995, 2002), Cortese and Khanna (2007) offered an alternative explanation of the significant imageability effects found in the regression analyses of Balota et al.'s (2004). Cortese and Khanna performed their own large scale regression analyses using the same word reading reaction times as Balota et al., which were from the ELexicon project, and the same predictors, including imageability. In addition, age-of-acquisition was also included. In summary, focusing on the results with imageability and age-of-acquisition, when age-of acquisition was not included in the regression analysis; so, the analysis was identical to Balota et al., the results were the same to Balota et al. as well, i.e., imageability was a significant predictor of word reading times. In contrast, when age-of-acquisition was included in the regression analysis, it, not imageability, was a significant predictor of naming reaction times. Since, in their analyses age-of-acquisition is a significant and unique predictor of word reading times, but imageability is not, the authors conclude that age-of-acquisition affects word reading, but imageability does not. Though age-of-acquisition was a significant predictor of reading times and imageability was not, unlike previous interpretations of age-of-acquisition (for example, Monaghan & Ellis, 2002), age-of-acquisition was described as being possibly semantic in nature. Since the significant predictor of the analyses is described as being semantic in nature the authors conclude that their results do not eliminate the possibility that word reading may receive a contribution from semantic memory. These effects and the interpretation of age-of-acquisition are discussed further in Chapter 8.

2.4 Semantic priming paradigms

Thus far, relevant investigations using factorial manipulations of a semantic variable, and regression designs with semantic variables as predictors of word reading times have been presented. Additionally, semantic priming designs, which incorporate factorial methods, provide another way in which a semantic contribution to word reading can be investigated. The following section, the final review section of this chapter, highlights semantic priming studies relevant to the aim of this thesis, including recent investigations that may be indicative of a semantic contribution to orthography-to-phonology translation, and which provides the starting point for the experiments of the current thesis.

Priming is the improvement of response performance to a target item. In consecutive or item-separated prime and target designs, the first item is the “prime”, and the item-of-interest that follows is the “target” or “probe”. The dependent variable is performance on the target item, such as reading, naming, or button press reaction times, with primes and targets being either related or unrelated. In priming paradigms a related prime item can ‘boost’ performance on the target item as compared to a target in the unrelated prime condition.

Relationships between primes and targets can be manipulated in a number of ways. The focus here will be on research that has used semantic priming in which primes and targets are related by meaning. In a semantic priming experiment, target items can be responded to more quickly in the semantically related prime condition because they have benefited from shared meaning (Damian & Als, 2005). For example, using the word “kettle” as a prime followed by the word “toaster” as a target, “toaster” may be read faster after reading the word “kettle” than when it is preceded by an unrelated prime such as “shoe”, because the “kettle” and “toaster” share semantic information and category membership (small electrical kitchen appliances). Therefore, the resonse “toaster” may be boosted by the related, activated semantic information of “kettle”, but more is said on this in Section 2.4.1.

Semantic priming is an extensive topic (for reviews see McNamara, 2005; Neely 1991; see Meyer & Schvaneveldt, 1971 for semantic priming’s inception). These types of studies can use a variety of contrasting designs including differences in intervals between primes and targets, responses to target items, modalities of primes and targets, and relationships between prime-target pairs. Additionally semantic priming effects can be interpreted in other ways - that is not simply the result of shared meaning, but the result of non-semantic effects. One issue in the literature is whether significant semantic priming might be due to automatic or strategic processes (Neely, 1991; Tse & Neely, 2007). These issues and theoretical accounts of the mechanisms of semantic priming and semantic memory models are reviewed in the following sections, respectively. Over the course of this section, the topics mentioned above will be briefly explored with a focus on semantic priming within the context of target word investigations because target word reading is relevant to the central aim of this thesis, which involves orthography-to-phonology computation.

2.4.1 Possible mechanisms of semantic priming and models of semantic memory

There are two explanations for significant priming effects; semantic and non-semantic. That is, the actual representation thought to provide the priming advantage can be semantic or non-semantic; in another sense, all priming effects of the type considered here, where the prime and target are category coordinates and not associates (e.g., "kettle" and "tea", or "mouse" and "cheese"), must originate initially from the semantic relationship between prime and target. Models of semantic memory attempt to account for significant priming effects. Here explanations of priming are reviewed and semantic memory models are presented.

Specifically within semantic sources of priming, there are three possible mechanisms. First, significant priming could result from a direct semantic effect on the processing of the target, specifically shared semantic information between the semantically related prime and target word (Damian & Als, 2005). A related prime is named and semantic information about this item is activated. This semantic representation is still activated when the target is presented, and since the target shares meaning with the prime, processing of the target will benefit from the prime's remaining semantic activation. This explanation may be best accounted for by distributed representation models of semantic memory, which are presented next, though spreading activation models could also account for priming in this way, if activation survives over the length of time between prime and target.

One group of semantic memory models can be classed as distributed representation models (Becker et al., 1997; Hinton & Shallice, 1991; Joordens & Becker, 1997; Masson, 1991; 1995; Plaut, 1995). Distributed representations can be used to model many types of information (see Section 1.3.3.1), including semantic knowledge. Within models that use distributed representations, knowledge is not represented as a whole node, but as a pattern of activation across a network. Priming occurs because the pattern of activation is similar between the related prime and target items. Therefore, few changes in the pattern of activation are needed to activate a representation of the target from the activated pattern of prime. This results in a faster response to the target. However, details of the various distributed models differ, and this is reviewed in Section 2.4.4.

The second way that a target word could benefit from the semantically related prime is through automatic spreading activation, with activation disseminating automatically to related items without the participants' influence. This is best accounted for by spreading activation models (Anderson 1993; Collins & Loftus, 1975), in which concepts are represented in local nodes (Anderson 1993; Collins & Loftus, 1975; Neely, 1977). These nodes are linked together in a network based on related-ness. For example the concept of "kettle" may be linked to other small kitchen appliances, such as "toaster", and would also be linked with other associate-related nodes, "tea". When the prime is processed, activation from its semantic node will spread automatically activating information in related items' nodes, probably including the target. Then, when a target from the same category as the prime is presented, it will be named faster than an unrelated target, because the semantic representation of the related target will be pre-activated from the spreading activation.

The third directly semantic account of priming is through controlled strategic expectancy of possible target items by participants, as they use the prime to predict and self-generate possible targets. For example, if participants receive the prime "kettle", then they may specifically guess potential target responses, including thinking about the target word "toaster"[12]. This explanation relies on participants knowing the purpose of the experiment (including the relationship between prime and target), the pattern of stimuli, and maintaining this throughout an experiment (Neely, 1991; Vitkovitch, Cooper-Pye, & Leadbetter, 2006). Strategically guessing "toaster" ultimately originates from the semantic category relationship between the prime and target. It is, surely, the only way to guess "toaster" from "kettle".

Significant priming effects could also arise in the following manner: although the link between prime and target can only be semantic, the direct advantage to the target could derive from a non-semantic representation, namely that of either orthography or phonology. In other words, if the semantic representation of the prime "kettle" activates the semantic representation of "toaster", whether automatically by shared features in distributed representations or by spreading activation of separate nodes, or strategically

[12] Incorrect guesses of the target item, which occurs when targets are in the unrelated prime condition, or when the target in the related prime condition is simply incorrectly guessed and the whole category is not also activated, could result in the slowing of target response time. Therefore studies into strategic effects often look for effect of facilitatory priming (quickening of response) or inhibition (slowing of response).

by participants guessing, then the semantic representation of "toaster" may also send activation to its orthography and/or phonology. Within this thesis "non-semantic" is used to discuss word target priming that could be due to non-semantic representations, whether automatic or strategic.

In the empirical priming experiments of this thesis, it will be helpful to be able to argue against non-semantic accounts of significant priming. Ways which attempt to rule out non-semantic accounts include certain design aspects of semantic priming studies and specific analyses, which will be described in this chapter and in Chapters 4 and 7. Furthermore, there are additional reasons as to why it is unlikely that orthography and/or phonology representations are responsible for priming effects. For example, Plaut (1991) argued that the phonology of an intended response (such as naming a prime picture) should be activated very late in the gradual activation of a semantic representation, because small late changes in semantic memory could make an early (but incorrect) phonological response candidate unhelpful or even harmful to producing the correct response. Plaut demonstrated in a connectionist model that when the system settles almost completely on a specific semantic representation, the corresponding phonology is still largely undetermined. If this is true for the phonology of the prime, then imagine how much less true early-pre-activation of phonology must be for a semantically related (and not yet seen) target word. Priming effects that are arguably due to a direct semantic mechanism, especially the mechanism of residual activation of the prime item itself, may be more readily interpreted as reflecting a semantic contribution to word reading, and within in this thesis, priming due to a "semantic" explanation is used with this specific meaning.

2.4.2 Semantic priming research of note

Recently Tse and Neely (2007) investigate whether the semantic activation in priming occur because of strategic or automatic mechanisms. They measured target word lexical decision times reaction times of healthy adult participants (i.e., how quickly participants decided if the letter string on the screen was a real word or a fake non-word) in a semantic priming paradigms with associated prime-target words (e.g., “mouse” prime “cheese” target). Since lexical decision is used, results may not necessarily add to issue of whether semantics contributes to orthography-to-phonology computation, but methodological points and word types are of interest to this thesis. Target words were manipulated on lexicality (real word or non-word) frequency (high or low) but not regularity. One condition included a letter search task, in which participants decided whether a given letter was present in the prime word, while in the other condition prime words were read silently. Letter search, it is claimed by the authors, will disrupt controlled strategic processes, as participants would not be able to devote resources in order to strategically predict a target. However, automatic processes should still occur as they will not be interrupted by the letter search task. If semantic priming of target words is due to automatically shared semantic information then the results of the letter search condition and the silent reading condition should be the same, both should show a priming effect. However, if semantic priming is due to strategic pre-activation of target information then the silent reading condition should show priming, whereas the letter search task should not. Tse and Neely also argue that high frequency target words that are in the associated prime condition could be more easily guessed by participants than low frequency targets; therefore priming of high frequency word targets, but not low frequency word targets, could indicate the use of strategic expectation. No priming effects in the high frequency target word condition during the prime letter search task might especially indicate priming due to strategic expectation, as the letter search task interrupts the participants' ability to guess the high frequency items associated with the prime. Low frequency target priming in the prime letter search condition may indicate automatic processes because low frequency targets are difficult to guess, therefore difficult for participants to strategically predict; moreover automatic process can proceed even when participants are occupied with searching for a letter.

Tse and Neely (2007) found priming of both high frequency and low frequency word targets in a silent reading condition. However, in the letter search condition high frequency target word priming was not found, but low frequency word target priming was found. The authors therefore argue that the priming found in the silent reading condition may be due to different sources for each frequency. Since high frequency priming is interrupted by the letter search task and high frequency words are easier to guess, this priming in the silent reading condition is likely due to strategic expectancy. However, since low frequency priming is found in the letter search task, and low frequency targets are difficult to guess, it is likely due to an automatic mechanism. They write, "Expectancy contributes to silent reading priming for high frequency but not low frequency because it is difficult to use the prime to generate relatively inaccessible low frequency words as potential targets…" (p. 1158).

Also worth noting here is a series of two studies that use semantic priming paradigms to investigate the same question as this thesis (Cortese, Simpson, & Woosley, 1997), i.e., whether semantic information contributes to orthography-to-phonology computation. If reading words does indeed involve semantic information, as suggested by the triangle model of word reading, then reading a prime word will activate semantic information, then when reading the related target word, orthography-to-phonology will also be computed via semantic memory, and since the prime and target in the related condition have similar semantic information the target can benefit from remaining activation in of the prime. Therefore significant semantic priming in this paradigm would be indicative of a semantic contribution to word reading. In Experiment 1, healthy adult participants read target words (no task was performed to prime word) in a priming paradigm where associated prime-word-target-word pairs (e.g., "mouse"-"cheese") were presented sequentially. Target words were low frequency words manipulated regularity (regular or exception). A significant main effect of priming was found with both low frequency regular and low frequency exception words being primed. In Experiment 2 low frequency target words were factorially manipulated on regularity (regular or exception) and also imageability (high or low). Significant main effects of regularity, imageability, and priming were found. There were also significant interactions: There was a regularity by priming interaction, i.e., exception target words were primed, whereas regular target words were not; there was also a regularity by imageability interaction, i.e., exception target words showed an imageability effect, consistent with the results of Strain et al. (1995) (Section 2.3.2).

Since related targets were primed, indicating that words are read via semantic memory and also because semantic information in the form of imageability affects the reading of difficult words, Cortese et al. (1997) conclude that the results from these semantic priming designs can be interpreted as offering support for a semantic contribution to word reading. However, some design aspects might mean that the priming results could be due to non-semantic causes. It is possible that with this design, e.g., associated and consecutive primes and targets, the significant priming effects could be explained by non-semantic mechanism. Specifically strategic expectancy could be involved, which means that orthography and phonology of the target word could have been strategically pre-activated. With these paradigms, it is difficult to eliminate this as the possible source of a priming effect.

This literature (Cortese et al., 1997; Tse & Neely, 2007) demonstrate (a) the usefulness of semantic priming paradigms to explore a semantic contribution to orthography-to-phonology translation (b) the interest in word types in these priming paradigms (c) the use of design elements to investigate explanations of significant priming effects. Some evidence of semantic activation in low frequency target word tasks is provided, but cannot provide a definitive conclusion about a semantic contribution to orthography-to-phonology computation, especially using semantic priming designs. There are design aspects that could be improved to maximise the likelihood that priming effects are due to residual semantic activation of the prime (Section 2.4.1). Design aspects include improvements on the relationship between prime and target, non-sequential presentation, and naming of every item to disguise the paradigm. These are discussed in more detail in the remainder of this chapter and the next.

2.4.3 Associative versus conceptual prime-target pairs

Within the two reviewed experiments (Cortese et al., 1997; Tse & Neely, 2007), the prime and target were associatively related, for example, “mouse” and “cheese”. Associative relationships may not necessarily measure a semantic similarity as the items may not share similar semantic information or semantic features (Becker et al., 1997). The relationship of the associated prime and target may not be categorical or conceptual in nature, such as with “kettle” and “toaster”[13]. The relationship that is measured by associate pairs is not clear. Priming of associated targets could occur because of non-

[13] Categorically related prime-target pairs may also be associates, but primarily they share a relationship of meaning, i.e. a semantic relationship.

semantic links, such as co-occurrence (McKoon & Ratcliff, 1992) or those measured by free association production norms (McNamara, 2005). Another reason to avoid using associated prime-target pairs is because this type of relationship may increase the possibility of prediction of a specific target (Tse & Neely, 2007). Therefore, priming could be due to pre-activation of the non-semantic information of the target item with associate pairs. These are reasons not to use associate pairs. With categorically related prime-target pairs, priming is more likely due to the shared semantic relationship, and these types of pairs are used within the priming experiments of this thesis (Experiments 1-4).

2.4.4 Intervening items (lag) and semantic memory models

Though prime-target pairs can be presented consecutively, as in Tse and Neely (2007) and Cortese et al. (1997), prime and target can be separated by an intervening item (Becker et al., 1997; Joorden & Becker, 1997 (which are reviewed shortly); Wheeldon & Monsell, 1994, with repetition priming). Intervening items, or filler items, between prime and target presentation can be referred to in terms of “lag”. Because the presentation of intervening item(s) between prime and target also takes time, lag is also confounded with time. The time between the onset of stimulus items can be called stimulus onset asynchrony (SOA).

Effects of lag, in both time and items, are studied in their own right (McNamara, 1992; McKone, 1998), but this is not the focus here. Important to the empirical priming paradigms of this thesis in Chapters 3-7, using an intervening item lag within a semantic priming design is a useful methodological choice as evidence from priming experiments with a lag can also contribute to explanations of priming (i.e., whether significant priming might be due semantic or non-semantic explanations) . Intervening item(s) between prime and target might interrupt controlled strategic non-semantic processes such as episodic recollection (Wheeldon & Monsell, 1992), implicit memory (McKone,1998), rehearsal of the prime item (Joordens & Besner, 1992; Masson, 1995; Neely, 1991), which might result in the pre-activation of orthography and phonology. Intervening items may also interrupt priming due to automatic pre-activation of orthography and phonology, as described below. Hence priming effects in designs with intervening items between prime and target can be more readily interpreted as being due to shared semantic information, specifically shared residual activation of the prime, which is of key importance to this thesis (Section 2.4.1).

Models of semantic memory attempt to account for semantic priming effects, including effects in priming experiments with lag. In spreading activation model, activation is short lived and decays over time and over distance between items within the network, but is not effected by the number of intervening items (Anderson, 1993; Collins & Loftus, 1975; Masson, 1995; McNamara, 2005; Ratcliff & McKoon, 1988). Within Anderson's models it is calculated that activation decays in under one second (Masson, 1995; McNamara, 2005; Ratcliff & McKoon, 1988). Priming over a lag is possible, but unlikely in spreading activation models (Anderson, 1993; Collins & Loftus, 1975; Ratcliff & McKoon, 1988), as intervening item(s) also bring additional time in which activation will decay. Therefore, the semantic activation of the prime, and the activation that spread from the prime to related items' semantic, orthographic and phonological forms will have decayed over the time created by the lag, leaving no activation for the target's benefit. A design with a lag between prime and target minimises possibility of priming due to automatic pre-activation of the target's orthography and phonology.

Distributed models of semantic memory can account for priming over a lag (Becker et al., 1997; Joordens & Becker, 1997; Masson,1995; Plaut, 1995). The activation of a concept is not all or nothing, and activation occurs with a pattern of activity across links. For priming to occur over a lag, the pattern of the prime's activation must not be completely replaced by the pattern of activation of the intervening item. Therefore intervening items between prime and targets are important as they have the potential to disrupt residual activation of the prime. A greater number of intervening items may more likely disrupt all of the prime's activation. To pre-empt, filler qualities affecting priming are explored in Experiments 3 and 4 of this thesis and discussed in Chapter 7.

One type of distributed model predicts priming across many unrelated filler items (Becker et al., 1997; Joordens & Becker, 1997). Within these models, priming is driven by connectionist learning. When a prime is named, the network's weights change, and learning occurs, resulting in a biased pathway for that item's information. A target, when it is related to the prime, can benefit from these long term changes, as information between the prime and target is similar. This model asserts that the learned information from the prime results in long-term changes and will remain for the target even over an especially long lag (Joordens & Becker, 1997); therefore priming should be found in the presence of many intervening items and a long time. However filler items cause learning too and have the potential to affect the weight changes from prime learning.

There are other types of distributed models (Masson, 1991; 1995; Plaut, 1995). A pattern of activation across many units within the network causes the activation of the one concept, and the target word, as it would have a similar pattern of activation, would benefit from this. The pattern of activation for an unrelated filler item between prime and target has the potential to "turn off" the pattern of activation of the prime. The nature of the intervening items may be important as a certain type of item may be more likely to disrupt remaining activation (Masson, 1995). In later version of these models activation of more than one item's pattern at a time is allowed, therefore both prime and filler could be activated simultaneously allowing for priming of a target (Masson, 1995; Plaut, 1995). Priming over more than one item has not been modelled (Deacon, Hewitt, & Tamny, 1998; Deacon et al., 2004; Masson, 1991, 1995; McNamara, 2005; Plaut, 1995).

Semantic priming found over a long lag may arguably provide evidence that the priming of the target is due semantic explanations, and provide a strong argument that this shared activation is due to remaining semantic activation of the priming item (Section 2.4.1). Priming found over a long lag could not be due to automatic pre-activation of the target information, including automatic pre-activation of target's orthography and phonology, as spreading activation is short-lived, according to spreading activation accounts (Anderson, 1993; McNamara, 2005; McKoon & Ratcliff, 1992). Long lag paradigms may also minimise controlled strategic pre-activation of the target word, as participants would, at the very least, have to hold all target possibilities in mind over the lag (other ways to reduce strategic pre-activation are discussed in subsequent sections). Therefore, evidence of word target priming in long lag paradigms may provide evidence of a semantic contribution to word reading. Relevant recent literature in this domain will now be reviewed and this will lead to the first empirical study of this thesis.

2.4.5 Priming research with word targets

2.4.5.1 Word-prime-to-word-target priming over a lag

Semantic priming of a word target from a word prime over a lag has been investigated, (Becker et al., 1997; Joordens & Becker, 1997; Joordens & Besner, 1992; Schvaneveldt, Meyer & Becker, 1976) and these semantic priming studies are briefly summarized below as they show priming of target words over a lag. Some studies reviewed here use word reading, however, the majority of these investigations use button press tasks. Semantic priming effects with a word target (from a word prime) over a long lag in tasks other than word reading have been found, but may be dependent on the type of task used or the design details of the experiment. Non-reading tasks include lexical decisions and other forms of button press. As discussed in Section 2.2.2, lexical decision tasks may not necessarily contribute to the aim of this thesis concerning orthography-to-phonology computation. This type of task has, however, contributed to distributed representation models of semantic memory and has also commented on a semantic contribution when using target word stimuli.

The studies reviewed here measured reaction times in a lexical decision tasks of healthy adult participants in priming paradigms with word primes and word targets with at least one intervening item. Schvaneveldt et al. (1976) used word-prime-to-word-target compound cue triplet paradigm which included a prime-filler-target condition, which is of interest here. In Experiments 1 and 2 of this series a significant priming over one intervening item was found. Joordens & Becker (1997) found significant priming in three experiments over a long lag with up to eight intervening items. However, though Becker et al. (1997) failed to find significant priming using lexical decision task at various lengths of lag, they did find priming over zero, four or eight lags with a semantic decision button press task. Computer simulations using their distributed representation model (Sections 2.4.1, 2.4.4) also replicated their behavioural findings. Moreover Deacon et al. (1998) and Deacon et al. (2004) failed to find priming over one intervening item (whether visible or masked) using ERP measures. Some of these results suggest that with tasks other than word reading priming of a target word (from a word prime) is possible over (at least) one intervening item (Becker et al. 1997; Joordens & Becker, 1997; Schvaneveldt et al, 1976); however word-prime-to-word-target priming over one item is not consistently found, as Deacon et al. (1998, 2004) show with ERP measures, and it is, possibly, most likely found with an explicitly semantic task (Becker et al, 1997).

With healthy adult participants and measuring word target *reading* reaction times, semantic priming of word-prime-to-word-targets with the task of word reading has been investigated, but significant priming effects are not reliably found, i.e., target words are not reliably read faster in the related prime condition as compared to the unrelated prime condition; results are inconsistent, especially in paradigms with an intervening item(s) (For reviews see Neely, 1991; McNamara, 2005). McNamara (1992) found significant priming of word targets (from word primes) over one intervening item, using word target reading, but not over two. When considering his review, McNamara (2005) argues, that on average, word target priming (by a word prime) with word reading over one intervening item is half of what it is with no intervening item. Through a series of experiments and computer simulations, Masson (1991) failed to find significant priming over an intervening item, later arguing (Masson, 1995) that it can be found, but the qualities of intervening items are important as they may interrupt activation of the prime. Joordens and Besner (1992) found reliable significant priming effects over one intervening item in their series of experiments. They claim that these effects are not due to strategic prediction or checking techniques, but to shared semantic information because the intervening item would interrupt these types of processes. Neely (1991) claims that significant priming is more easily captured when prime-target pairs are associates (e.g., “mouse”-“cheese”), as compared to category coordinates (e.g., “kettle”-“toaster”. However, priming in this circumstance could offer strong evidence for a semantic contribution to word reading (Section 2.4.3) as prediction is more difficult, and the relationship is purely in meaning.

Whether semantic priming occurs in word-prime-to-word-target paradigms with word reading over a lag is not clear. There may be evidence of priming over a lag with studies that do not use word target reading tasks, but these do not ultimately answer whether priming can be found with word target reading, or whether semantic information contributed to word reading because the task is something other than orthography-to-phonology computation. There is some suggestion that semantic priming might occur with a task of word reading, but effects are inconsistent, and improvements on the design, e.g., the relationship of the prime and target, could be made to firmly establish that effects are due to shared semantics as opposed to non-semantic mechanisms. Additionally, inconsistent results in the target word reading studies could be due to the prime’s modality (word modality in these studies).

Within the triangle model of word reading, partial activation of semantic representation pattern is possible when reading a word (Harm & Seidenberg, 1999, 2004) (Section 1.3.3). If the semantic pattern of a word prime is only partially activated, then there may be less semantic information available for a related target word to benefit from, especially over a lag (as the filler item may disrupt what activation is available for sharing), than, for example, if a whole semantic pattern is activated when reading a word prime. This may make priming effects more difficult to find. Paradigms using primes of a different modality, such as a picture prime, may offer more definitive evidence of a semantic information contribution to orthography-to-phonology computation. To name a picture one needs a detailed and full activation of a semantic pattern. Indeed, there may be evidence for semantic information contribution to orthography-to-phonology computation from cross-modal priming designs, specifically paradigms that use picture primes and word targets.

2.4.5.2 Picture-prime-to-word-target priming

In cross-modal semantic priming designs, primes and targets are not in the same modality, for example, primes may be pictures and targets may be words, or visa-versa. The aim of the majority of cross-modal studies has been to investigate whether pictures and words access the same semantic information in a similar way, with past work generally favouring one shared semantic system (Bajo, 1988; Carr, McCauley, Sperber, & Parmelee, 1982; Glaser & Glaser, 1989; Seifert, 1997; Smith & Magee, 1980). Recent work also favours a shared semantic system therefore both words and pictures would access the same semantic representation. (Barber &Kutas, 2007; Patterson, Nestor, &Rogers, 2007; Price, Moore, Humphreys, & Wise, 1997; Pulvermuller et al., 2009; Riddoch, Humphreys, Coltheart, & Funnell, 1988; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; Vitkovitch & Cooper, 2012). This is important because if pictures and words did not access the same set of semantic representations, then even if naming a prime picture activated the whole semantic pattern, when the related target word was read, it could not benefit from any remaining semantic activation because it would be accessing a set of semantic representations different from the prime picture.

Some of these studies also have the potential to produce results that may implicate a semantic contribution to word reading, even though these investigations do not necessarily aim to investigate this primarily. Specifically the picture prime-to-word target conditions of cross modal studies may have the potential to implicate a semantic route to target word reading and these studies are reviewed in this section. Additionally given that the long lag paradigm may provide data that is most relevant to the aim of this thesis, it is also relevant to review in particular long lag cross-modal investigations.

In order to name an object, semantic information in semantic memory must be accessed (Humphreys, Riddoch, & Quinlan, 1988; Lambon-Ralph, McClelland, Patterson, Galton, & Hodges, 2001; Wheeldon & Monsell, 1992, 1994). Long lasting semantic effects from a picture prime are well established in the literature (Damian & Als, 2005; Lee &Williams, 2001; Tree & Hirsh, 2003; Vitkovitch & Humphreys, 1991; Vitkovitch & Rutter, 2000; Wheeldon & Monsell, 1992). Naming a prime object picture activates semantic information, and then connections from semantic memory to phonology will be used for the picture to be named, using connections between semantic memory and phonology, and also possibly affecting the weights of connections in semantic memory because of this exposure (Damian & Als, 2005). Usually effects with a *picture target* (as opposed to the word target focus of this thesis) result in inhibition (also called inhibitory priming)-that is the slowing of a response to a related target item (as opposed to the traditional speeding of facilitatory priming)-, and there is disagreement about the specific mechanisms involved in inhibition, and this is briefly presented in the subsequent paragraph (Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Oppenheim, Dell, & Schwartz, 2010). Understanding whether or not inhibition and facilitation priming effects occur in semantic memory is important to the central aim of this thesis because effects that occur outside of semantic memory cannot provide information about a semantic contribution to word reading. Therefore where (i.e., in semantic memory or not) inhibition and facilitation effects occur is discussed in the subsequent paragraphs.

With picture-prime-to-picture-target naming (i.e., not cross-modal, but within-modal prime-target pairs) inhibition is thought to occur because when the first picture (the prime) is named, the links between semantic memory and phonology are strengthened for that item's name (Vitkovitch & Humphreys, 1991; Vitkovitch, Humphreys, & Lloyd-Jones, 1993). The links between semantics and phonology for competitor items that are similar are either weakened (Oppenheim et al., 2010) or simply not strengthened (Howard et al., 2006). When the related target picture is named, it is named more slowly than when in an unrelated prime condition because its links between semantics and phonology have been previously suppressed or not strengthened. Using these principles, inhibition with a *word* target is possible by the same mechanisms. This would indicate that the links between semantics and phonology are strengthened for one specific item, and not broad range of semantically similar items, otherwise priming would always be expected. Additionally, inhibition may not be expected with word targets as an orthography-to-phonology contribution would influence the computation of phonology as well as a semantic pathway. Therefore inhibition effects are described as occurring outside of semantic memory.

As described by Damian and Als (2007) facilitatory priming effects, if not strategic in nature or due to automatic pre-activation of orthography and phonology, are due to shared semantic activation, whereas inhibition (with picture targets) is likely due shared processing outside of semantic memory, such as the links between semantics and phonology[14] (Howard et al., 2006; Oppenheim et al., 2010). Therefore the focus of this thesis is facilitatory priming effects as these likely indicate effects within semantic memory. It may possible for a target word to benefit from these facilitatory effects if semantic information is used in orthography-to-phonology computation.

When a word target is read correctly, orthography-to-phonology must be computed. The question of this thesis is whether this computation involves semantic information. In these priming paradigms, if word target reading is quickened when preceded some trials earlier by a related picture prime, relative to unrelated prime-target trials, then the word target may be drawing on remaining semantic activation of the prime picture (especially if the paradigm makes non-semantic explanations not very likely). Therefore facilitatory priming effects, especially with word target reading, are of interest to this thesis.

2.4.5.2.1 Picture-prime-to-word-target priming with no lag

Facilitatory priming of word targets from related picture primes has been found when there are no intervening items (Bajo, 1988; Bajo & Canas, 1989; Hines, Czerwinski, Sawyer, Dwyer, 1986), and these are reviewed below. Whether priming occurs when masking is used is unclear (Carr et al. 1982; Hines et al., 1986), and these are reviewed

[14] See Vitkovitch and Cooper (2012) for evidence that inhibition is not due to strategic processing. Vitkovitch and Cooper used a long lag prime(word)-filler-filler-target(picture) semantic priming paradigm similar to that used in Vitkovitch et al. (2006) and the priming paradigms presented in Chapters 3-7 of this thesis. On half of the prime(word)-filler-filler-target(picture) trials, participants completed a counting backwards task directly following prime word naming. This counting task, with its difficult computational aspect and aloud vocal component would prevent participants from holding the prime in mind and thinking actively about possible related items, but would not prevent any automatic processes that might be responsible for inhibitory effects. Therefore if inhibitory effects on the count backwards task trials are similar to the inhibitory effects found on the non-count-backwards-trials, i.e., no significant interaction of task, then the inhibitory effects are likely due to automatic processes, not strategic conscious anticipation. Vitkovitch and Cooper found a significant main effect of relatedness, revealing inhibition of related target trials, no main effect of task condition, and no interaction of the inhibition effect with task condition. The authors, therefore, concluded that inhibitory processes are not likely due strategic conscious anticipation by participants.

below as well. These studies, however, do not necessarily implicate a semantic contribution to word reading, as will be shown in this review.

In Experiment 1 of Bajo (1988), which used a cross-modal (picture-prime-to-word-targets) consecutive semantic priming paradigm with no intervening item, participants' button press reaction times were measured as they either verified the target word name via button press, with this task perhaps focusing on processes similar to word reading, or verified the target word category via a button press, with this task emphasising semantic meaning information about the target word. Significant priming was not found with a name verification task, but was found with the category verification task. In Experiment 2 of Bajo, participants read word targets, in the same paradigm as Experiment 1. Significant facilitatory priming was found when instructions for participant emphasised the semantic relationship between prime and target and mixed the picture and word trials in a block. Moreover, Bajo and Canas (1989) found significant priming of word targets with word target reading in a similar cross-modal paradigm to Bajo when the semantic relationship was revealed to participants. Using healthy adult participant in a cross-modal (picture-prime-to-word-targets) consecutive (no intervening item lag) semantic priming paradigms, with tasks of button press responses and word reading, Bajo (1988) and Bajo and Canas (1989) found significant priming of word targets in the related prime picture condition, as compared to the unrelated prime condition. It is possible that these effects indicate a semantic contribution to target word reading because a speeding of word target reading was found in the related condition (Bajo, 1988; Bajo & Canas, 1989); some methodological points, however, must be highlighted, as these significant effects could be interpreted as being non-semantic or as not being indicative of ordinary word reading, as they may have been found in circumstances where semantic processing was emphasised.

If the results of Bajo (1988) and Bajo and Canas (1989) are semantic in nature, then they may have only occurred because word reading was biased to use semantic memory and non-semantic explanations of the priming effects cannot be eliminated. Priming was not found in the name verification task, but was only found when the category or semantic relationship was emphasized. There is not an intervening item and participants were informed of the prime-target relationship, both of these factors make non-semantic explanations more likely than if the paradigm used an intervening item and was covert in nature. Therefore the priming effects obtained in these studies could be due to pre-activation of non-semantic information.

Priming of word targets has been obtained in a similar experimental paradigm, but this one minimised non-semantic strategic explanations of priming. Carr et al. (1982) used a cross-modal paradigm (picture-prime-word target) with no intervening item and picture primes presented below perceivable conscious thresholds. Participants' target word reading reaction times were measured. Significant word target priming was found (Carr et al., 1982). This type of priming would be less readily explained using strategic expectation because the prime cannot be detected easily. Therefore, it cannot be used to predict the target item. Carr et al. found that word priming might be greater when prime pictures were high category exemplars. Priming was also greater for participants with a longer response time (Carr et al., 1982). Priming of a word target with a masked picture prime might suggest that priming is not due to strategic conscious pre-activation of target information, but it does not eliminate the possibility of significant automatic non-conscious priming.

Hines et al. (1986) used a consecutive cross-modal semantic priming paradigm and measured healthy adult participant target response times; of particular interest is the picture-prime-to-word-target condition. Like Carr et al. (1982) primes were below identification threshold. In contrast, to the results of Carr et al., Hines et al. failed to find significant priming of word targets. However, in a different condition, where prime pictures were at a perceivable threshold, priming of word targets was found in the related prime picture condition. Because priming was found when the prime picture could be identified, but was not found when the prime picture was not perceivable, the authors conclude that significant priming effects are likely due to strategic processes, including possibly conscious controlled pre-activation of the target's non-semantic information (orthography and phonology). It does not eliminate, however, the possibility that significant priming is still due to shared semantic information between the prime and target. However, as no intervening item is used in this paradigm, non-conscious automatic pre-activation of non-semantic information can not be eliminated as an explanation either.

Though significant semantic priming of word targets has been found using cross-modal paradigms, those just reviewed do not definitively implicate a semantic contribution to word reading. Some studies found priming when the task emphasised semantic processing (Bajo, 1988; Bajo & Canas, 1989), and none used an intervening item (Bajo, 1988; Bajo & Canas, 1989; Carr et al., 1982; Hines et al., 1986); so, there is no way to eliminate non-semantic explanations of priming, even with studies using masked-priming. There are further investigations that could be performed while improving some design aspects; for instance, designs using intervening items to reduce the likelihood of non-semantic explinations. Cross-modal priming studies that have used intervening items will now be reviewed, as they may provide stronger evidence of a potential for a semantic contribution to word target reading (Section 2.4.4).

2.4.5.2.2 Picture-prime-to-word-target priming with a lag

Recent research by Vitkovitch et al. (2006) may have found support for a semantic contribution to word reading. In a long lag cross-modal conceptual semantic priming paradigm in which the priming design was not revealed to participants, the authors investigated processes that are shared between pictures and words when they are named. Vitkovitch et al.'s (2006) two experiments focused on cross-modal target picture naming and whether word primes could create semantic competition (inhibition) for picture targets. However, a word-prime-to-picture-target design is rarely used in isolation. Both experiments had a number of conditions that mixed pictures and words, using them as both targets and primes. A picture-prime-to-word-target condition produced the results of interest to the central aim of this thesis. Using healthy adults, and measuring word target reading reaction times, Vitkovitch et al. found priming of word targets from a picture prime with two intervening filler items, using a task of stimulus naming (including target word reading). This facilitatory priming effect, however, was not pursued further as interference effects were the main focus. As the facilitatory priming effects are promising, this investigation is detailed here and in the start of the next chapter.

The experiments used line-drawings of objects as primes and object names as word targets in one condition of the experiments. In the primed condition primes and targets were semantically related category coordinates. For example, the prime "kettle" was in picture modality and the target "toaster" in word modality. In both experiments, when a target word was read aloud after a category related prime picture and two intervening items, facilitatory priming was found in planned comparisons. Experiment 1 included one condition with two intervening filler items separating picture primes and word targets, which was intermixed with other non-lag conditions. Significant semantic priming was found in this condition. Experiment 2 sought to replicate the two intervening filler item condition of Experiment 1 and is the experiment of interest.

In Experiment 2, stimulus quadruplets were presented in the order of prime, filler, filler, target. An example of a related quadruplet is, “kettle” (prime) is a picture, “wheel” (filler 1), “fairy” (filler 2), and “toaster” (target) is a word (PFFW). Prime-target pairs were designed to appear in each modality and modality combination. Therefore prime and target pairs were either cross-modal, that is picture-prime-to-word-target (PFFW) or word-prime-to-picture-target (WFFP), or within-modal, that is picture-prime-to-picture-target (PFFP) or word-prime-to-word-target (WFFW). This was a between subjects manipulation. The modality of the two filler items was also manipulated. Each stimuli list contained filler pairs in modalities of word-then-word, word-then-picture, picture-then-picture, and picture-then-word. The cross-modal condition of picture-prime-to-word-target (PFFW) produced results that are of interest to this thesis.

Target words, e.g., the word “toaster”, were read aloud more quickly after naming two unrelated intervening filler items and a preceding semantically related prime picture, e.g., the prime picture “kettle” not the unrelated picture of a “boot”. This effect was found with two intervening filler items and 16 seconds between prime and target items. The authors were not able to fully explain facilitation effect, nor was the effect the focus of their investigations. It was speculated that this effect might be due to a strategic pre-activation of speculated targets. However, given the design of this priming study, including the two intervening items, and the discussion above of the advantages of long lag experiments, potential priming due to non-semantic explanations, including pre-activation of the target’s orthography and phonology (whether strategically or automatically), is not very likely. Since the possibility non-semantic explanations are reduced due to design aspects of this paradigm and because significant target word facilitatory priming was found in the related prime picture conditions, this priming result could instead reflect genuine shared semantic activation, with target word reading benefiting from residual activation of the prime picture, thus indicating that target words are read via semantic memory. Therefore the priming results of Vitkovitch et al. might demonstrate a semantic contribution to orthography-to-phonology computation. It is the exploration of this priming result that begins the empirical work of this thesis in the following chapter, as it pertains directly to the central aim of this thesis.

2.5 Additional aims addressed by this thesis

The research of this thesis aims to investigate whether there is a contribution of semantic information to orthography-to-phonology computation. Over the last two chapters the controversy concerning a semantic contribution to word reading has been reviewed. The review highlighted word types and computational models of word reading and their differences in semantic memory implementation. Word reading research that used factorial and regression methods and their failure to produce clear results was presented, and the potential for neurophysiological approaches was identified. Semantic priming literature, including how effects can be explained, and models of semantic memory have been reviewed. A recent cross-modal semantic priming design may have uncovered evidence of a semantic contribution to word reading, and this was reviewed.

Original work to the area of word reading and investigations of a semantic contribution to orthography-to-phonology computation will be performed as part of this thesis. It will use experimental designs and quantitative methods across the programme of research.

Within the review the relevance of using complementary methods was also highlighted, and this thesis will use factorial, regression and neurophsyiological approaches to capitalise on this. ANOVA and t-test analyses are used with word reading reaction time data collected in a series of semantic priming experiments that include a factorial manipulation of target word types, regression analyses are used with data from single-word reaction times in a large previously published dataset (ELexicon), and ERP measurements collected during two single word reading experiments with factorial manipulations are analysed using ANOVA and t-test analyses. Additional aims of this thesis are to:

(1) Investigate whether a semantic contribution is evident in word reading using long lag cross modal semantic priming paradigms and various word types as targets.

(2) Investigate whether the semantic variables of semantic features and imageability are significant predictors in regression analyses while accounting for other known variables, including age-of-acquisition and using ELexicon single word reading reaction times.

(3) Use ERP measures to investigate the neuralcorrelates of the semantic measures of imageability and semantic features in a silent word reading paradigm, with a specific interest in whether semantic effects are present early in the time course of word reading prior to phonological processing (as defined by the literature), thus showing the potential for a semantic contribution to word reading.

It is expected that the research of this thesis of will add to the literature by investigating the controversy of whether a semantic contribution occurs in healthy adult orthography-to-phonology computation. Using three complementary lines of investigation (semantic priming experiments, regression analyses, and ERP experiments) within the same program of research has the potential to provide definitive converging evidence as to whether this phenomenon occurs. This research will provide information on the reviewed semantic memory models (Section 2.4.1, see Chapter 7), on word reading in healthy adults, including whether semantic information contributes to various word types (Section 1.2.3) or whether a semantic contribution is only seen with the most "difficult" word types (low frequency exception words), as previously found in the literature. Information will also be provided about the semantic measures of imageability and semantic features, and age-of-acquisition. The results of this programme of research will be considered in relation to the reviewed computational models of word reading (Section 1.3), and comment will be provided as to which model best accounts for the results of this thesis' investigations.

Chapter 3

Semantic Priming Experiment 1

3.1 Introduction

In the previous chapter, literature utilising various behavioural designs to investigate whether semantic information contributes to orthography-to-phonology computation was reviewed. This review highlighted a cross-modal long lag semantic priming design that may have implicated a semantic contribution to word reading (Vitkovitch et al., 2006).

3.1.1 Priming Effects

Experiment 2 of Vitkovitch et al. (2006) found significant priming from a related prime picture over two intervening filler items to a target word, replicating an effect found in a condition of Experiment 1. The focus of the research of Vitkovitch et al., however, was not facilitatory priming, and the authors did not focus on the interpretation of the priming from a related picture prime to a word target. The authors admitted that they were unsure how to interpret this "unexpected" (p. 723) and "unanticipated" (p. 719) semantic priming effect on word targets. They write, "We are unable in the present experiments to provide a conclusive answer as to why there is a picture-to-word facilitation effect over two unrelated trials" (Vitkovitch et al. p. 723). They considered a possible non-semantic explanation of the significant priming effects and a pre-activation of the target's phonology or orthography through strategic prediction. However, in view of the preceding reviews, this does not now seem the most likely explanation.

There is, however, an alternative interpretation of these results, which Vitkovitch et al. (2006) do not categorically eliminate. These priming effects could instead indicate a semantic involvement in word reading. As discussed in Chapter 2, when a related picture prime is named, the semantic information for that concept will be activated. When a target word is then named two intervening items later, if it is processed using semantic information, e.g., from orthography-to-semantics-to-phonology, then a target word could benefit from semantic activation remaining from the processing of the prime, and the response to that target word will be faster compared to a word in the unrelated prime condition[15]. Semantic priming effects, a quickening of response, may be indicative of shared semantic activation between the prime and the target (Damian & Als, 2005). Also, see Section 2.4.5.2. By focusing on facilitatory priming, if a successful argument is made against pre-activation of orthography and phonology (whether strategic or automatic), then significant priming effects are likely due to shared semantic activation. Inhibitory effects may not provide the same assurances as it is possible that inhibitory effects occur outside semantic memory (see Section 2.4.5.2). Additionally, whether or not a semantic contribution occurs with some or all word types in healthy adult word reading could also be revealed with a semantic priming design and significant facilitatory priming effects.

[15] Thanks to an anonymous reviewer for this alternative interpretation.

It is the priming of target words by a picture prime in Vitkovitch et al. (2006) that is of interest here, and it requires further investigation. In addition to revealing a possible semantic contribution to word reading that is worth investigating, there may be advantages to replicating the design used in this experiment, which will be briefly summarised.

3.1.2 Advantages of long lag semantic priming designs

In order for priming to reveal a semantic contribution to word reading, aspects of the priming design must offer reasonable arguments against non-semantic explanations of priming (whether automatic or strategic), giving favour to arguments of priming due shared semantic activation between the prime and target (Section 2.4.1). There are, therefore, several advantages to using a long lag semantic priming design, as was discussed in Section 2.4.1, such as that of Experiment 2 of Vitkovitch et al. (2006), to study target word reading.

One advantage of using a semantic priming paradigm is the constraint it can offer over confounding variables. Each item within the paradigm is presented only once to each participant. Within a full counterbalancing of materials, target words are compared to themselves in primed (related prime-target pairs) and unprimed (unrelated prime-target pairs) conditions. Therefore when comparing a specific related and unrelated condition, any priming effects cannot be due to uncontrolled factors that may differ between target words, as could occur in other paradigms (e.g., Ellis & Monaghan, 2002; Monaghan & Ellis, 2002; Strain et al., 1995, 2002).

Another advantage of the priming design used by Vitkovitch et al. (2006) is the "covert" priming design and intervening items between category-coordinate prime and target. The participants were not informed of the relationship between stimulus items or the priming design. They were told that it was a naming experiment, investigating the naming of individual items. Moreover, unlike other priming designs, every item in the quadruplet was named, so the target should not have seemed any different from the other stimuli. Participants, therefore, were not overtly made aware of the priming design.

Additionally, prime-target pairs were separated by two intervening filler items and related and unrelated quadruplets were randomly mixed. The time delay between prime and target may limit automatic processes, such as spreading activation from the prime's orthography to the target's orthography, as such activation is considered short-lived (Sections 2.4.1. and 2.4.4.). As the order is unpredictable, due to mixed quadruplets of related and unrelated prime-target pairs, prediction from a prime to a target may be difficult as participants may not know which items are important. Expectation would require guessing and holding items in mind for all stimuli (Neely, 1991; Vitkovitch et al., 2006). Therefore the design of Experiment 2 of Vitkovitch et al. (2006) offers several advantages that promote the likelihood that remaining activation of prime's semantic representation is the leading explanation of any facilitatory priming effects.

3.1.3 Manipulation of target word type

As presented in detail in Chapter 1, words can be categorised into types using frequency and regularity, and are associated with benchmark effects that computational models attempt to explain. Word types are a valuable tool when investigating a semantic contribution to word reading with healthy adult readers, and investigating whether the target words within the Vitkovitch et al. study are of various types, e.g., low frequency exception, could provide important information concerning the aim of this thesis. Different word types may explain the circumstances in which a semantic contribution, indexed by priming, will be observed in healthy adult reading. Such effects may also be accommodated by one computational model of word reading more than another. It is also possible that a specific pattern of priming effects across the various word types might be indicative of strategic effects as explained by Tse and Neely (2007) when accounting for high frequency word effects (Section 2.4.2) However, as noted, the present design makes this less likely.

Not all of the computational models of word reading can accommodate semantic priming of the four various word types. The DRC and CDP+ models emphasise the efficiency of orthography-to-phonology links. These models would likely attribute priming effects to a strategic process than to semantic involvement. However, since theoretically, at least, semantic memory is still included in these models, if it was implemented then it is possible for these models to account for priming with a semantic account. Within the triangle model, semantic memory is implemented and has time to contribute to word reading with difficult word types, such as low frequency exception words. Moreover it does not eliminate the possibility of semantic contribution to all word types. In the current studies, this would include semantic priming effects.

3.1.4 Reanalysis Vitkovitch et al.'s (2006) data

As a reminder, the four commonly studied word types are: high frequency regular, high frequency exception, low frequency regular, low frequency exception. Vitkovitch et al.

(2006) did not manipulate target word type when designing the prime-target pairs, but their target words may nonetheless be comprised of the four types. A reanalysis of Experiment 2 from Vitkovitch et al. was performed taking this into consideration. Reaction times from the cross-modal picture prime-to-word target between related and unrelated conditions (prime picture, filler item, filler item, word target sequences) were used in this reanalysis of participant reaction times, as this is (a) the condition that revealed priming for word targets in the original experiment, and (b) pertains directly to the thesis aim.

There were 44 eligible[16] target word stimuli out of a possible 48. Thirty-five of these were considered low frequency words (Sections 1.2 and 1.2.1) with frequencies below 14 per million in Kucera and Francis (1967) (provided by provided by Balota et al., 2007) and below 17 per million in Celex written frequencies (Baayen, Piepenbrock, & Gulikers, 1995, provided by Medler & Binder, 2005); only nine words were high frequency. As noted in Chapter 1, spelling-to-sound regularity is easiest to classify in mono-syllable words. When using strict rule based regularity for one syllable words (Section 1.2.2.1.), as calculated by the Neighbourhood-Watch program (Davis, 2005), only half (22) of the eligible words, can be classified according to their spelling-to-sound regularity as the other half of the stimuli words are two or more syllables in length. Of these 22 words, nine were exception words and 13 were regular words. Using these measures therefore, there are five low frequency exception words, eight low frequency regular words, four high frequency exception words, and five high frequency regular words across the stimulus lists of Vitkovitch et al. However, these words were not equally divided across the four stimulus lists.

[16] Words omitted were: "cotton reel" and "lorry" as they did not appear in either frequency database, and "sofa" and "screw" as they did not fall into the same frequency bin (i.e. the words were high in frequency in one database and low in frequency in another database).

Because of the design, however, only two of the lists for any one participant were in the picture-prime-to-word-target condition of interest. Each participant had a maximum of 12 reaction times in the related condition and 12 reaction times in the unrelated condition. Therefore, a reanalysis considering the four word types would be based on a very low number of stimuli or sometimes none at all for a particular list. In other words, a factorial reanalysis of semantic priming with frequency and regularity in an orthogonal manipulation was not possible due to nature of the original design and the low number of maximum data.

A reanalysis considering low frequency words only is, however, possible. Eighty percent of the original eligible word targets used by Vitkovitch et al. (2006) were low frequency words, as defined above. If the priming effect of Experiment 2 in Vitkovitch et al. is present when these low frequency word targets are considered in this reanalysis, then this may indicate two things. First, it may indicate that the priming results of Vitkovitch et al. are unlikely to be due to non-semantic strategic expectancy, because as argued Tse and Neely (2007) low frequency words are less predictable than high frequency words. Secondly, it could also indicate that semantic information contributes to the reading of a written word. The triangle model of word reading predicts that in the processing of low frequency words with less efficient orthography-to-phonology, there is time in the computation process for semantic information to contribute.

Comparison analyses were performed using participant word reading reaction time means and medians for low frequency target words in the related and unrelated conditions; (Vitkovitch et al., 2006, used medians). The priming effect approached significance with participant means, $t_p(39) = 1.72$, SEM = 14.35, p =.09, two-tailed, but was not significant with medians, $t_p(39) = 1.10$, SEM = 17.40, p = .28, two-tailed. As these were the first analyses performed as part of this thesis' investigations, there were no specific predictions as to whether the effects would be facilitatory or inhibitory; two-tailed tests were, therefore, employed (Field, 2000). Though the priming effect failed to reach significance, the means and medians revealed a strong trend for priming. Related target word means were 25ms faster than the unrelated target word means, and related medians had a 19ms advantage. Items analyses may provide more information.

Item analyses of all eligible target word reaction times for stimuli in the picture prime-to-word target condition were performed using the original Vitkovitch et al. (2006) data. This analysis was not reported for the original data in the article. The priming effect failed to reach significance either by means, $t_i(43) = 0.34$, SEM = 20.70, $p = .74$, two-tailed, or by medians, $t_i(43) = 0.52$, SEM = 20.85, $p = .60$, two-tailed. Items analyses restricted to low frequency target word reaction times also failed to reach significance by means, $t_i(34) = 1.64$, SEM = 19.59, $p = .11$, two-tailed, though median analysis approached significance, $t_i(34) = 1.87$, SEM = 19.72, $p = .07$, two-tailed. Low frequency related item means were 32ms faster on average than targets in the unrelated condition. Medians demonstrate that related targets were read 37 ms faster than word targets in the unrelated condition.

The marginally significant primed difference in the reanalysis of participant reaction times for low frequency word targets (25ms) was similar to that of the original analysis (21ms). The primed difference in the items analysis (32ms) was larger than the original participant analysis (21ms). Yet, both participant and item analyses failed to reach significance because the analysis was based on a subset of the original data. However, even with a loss of power there remained a trend towards priming with low frequency target words. Therefore further investigation of semantic priming with low frequency target words is merited.

3.1.5 Aims

The semantic priming experiment presented in this chapter aims to replicate and extend the work of Vitkovitch et al. (2006) to explore further the possible semantic priming effect with word targets, and to extend the design by including manipulations of target word type. The case has already been made for investigations to begin with low frequency word targets, and target items of this first experiment will be low frequency regular words and low frequency exception words. Carrying out this work may reveal evidence of a semantic contribution to word reading in the form of semantic priming. Of particular interest, in light of the aim of this thesis, are significant facilitatory priming effects, i.e. the quickening of response to a related item, and not inhibition (Section 2.4.5.2), i.e. the slowing of a response to a related item. Priming effects, if not strategic in nature, are arguably due to shared semantic activation. Tools used to reduce the likelihood of an explanation of non-semantic strategic effects, if significant priming is found, are as follows: the experiment's design, i.e., long-lag and covert priming (Section 3.1.2), and post-experiment questionnaire responses.

3.1.6 Predictions

On the basis that Vitkovitch et al. (2006) find picture-prime-to-word-target priming using 80% low frequency target words, that there is a trend in the reanalysis of the data from Experiment 2 of Vitkovitch et al. for low frequency word target priming, and that other empirical word reading work suggests a semantic contribution might occur with low frequency words (e.g., Strain et al., 1995, 2002), priming might be expected in regular and exception target word conditions of this study. Therefore, in all priming experiments of this thesis (Chapters 3-6), initially ANOVA analyses will be used to explore the possibility of priming in all included target word types. Special attention in the literature has been paid to low frequency exception words, as detailed in Chapters 1 and 2, and the triangle model of word reading predicts that semantic information has time to contribute to word reading when the connections from orthography-to-phonology are less efficient, as with low frequency exception words; therefore, in the event of non-significant priming results in the initial ANOVA, in all priming experiments of this thesis, specific comparisons (planned t-tests) will explore low frequency exception target word reading, and power was calculated with a view to the comparison of these two conditions, that is, related low frequency exception target words versus unrelated low frequency exception target words.

3.1.7 Power to detect effects

Using the G*power program (Faul, Erdfelder, Lang, & Buchner, 2007), power calculations based on a replication of the picture-to-word condition of Vitkovitch et al.'s (2006) Experiment 2 were undertaken. The power to detect a small effect (effect size = .32) using a sample size of 48 was .72. The current study aimed to extend the original study and increase the power to detect an effect through increasing the number of stimuli (from 24 to 56 picture-prime-target-word pairs), and also matching on a number of factors. Furthermore, on the basis of the earlier work by Vitkovitch et al., the broader literature demonstrating facilitatory priming effects, and the word reading literature, one-tailed t-tests were used to investigate the planned comparison.

3.2 Methods

3.2.1 Participants

Forty-eight volunteers (12 male) with an average age of 32 years from the University of East London participated in this experiment. Participants were entered into a prize draw for book tokens in exchange for their participation. All reported English as their first language; nine reported fluency in a second language. All reported normal or corrected-to-normal vision.

3.2.2 Ethical details

Ethical approval was obtained for all experiments of this thesis from the University of East London, University Research Ethics Committee. The research for all studies of this thesis was performed in accordance with this approval and with the BPS Code of Ethics and Conduct. This included, but was not limited to, providing information about the specific study, the rights of participants to the volunteer, and the anonymity of their data. Also signed informed consent was obtained prior to the participation, and participants were debriefed after participation.

3.2.3 Stimuli and design

The picture prime-to-word target conditions of Vitkovitch et al.'s (2006) long lag, cross modal, semantic priming paradigm was used in this experiment as a within participants design. To create the target word manipulation of the current study, new stimuli were selected. Each participant received stimulus quadruplets in the sequence of prime picture, filler item, filler item, target word (PFFW). Low frequency regular and low frequency exception words were used as targets. Prime pictures were line drawings of objects. Primes were systematically rotated to create related and unrelated conditions with the target words. The experiment's design including the selection and rotation of stimuli is described in the following sections. See Appendix A for complete stimulus lists.

3.2.3.1 Target words

Targets were 56 low frequency words. All values in section 3.2.3.1 and subsections are also presented in Table 3.1.The majority of words (40 out of 56) were one syllable and none were more than two syllables in length. Target words had a Kucera and Francis (1967) written frequency count of less than 16 per million, average 7.18, (provided by, Balota et al., 2007) and a Celex lexical database count of less than 14 per million, average 6.34, with the exception of "wool", which has a Celex count of 22 per million (Baayen et al.,1995, provided by Medler & Binder, 2005). Words were also high in imageability and "picture-able" as to replicate Vitkovitch et al. (2006), in which all target words also appeared in picture form. For the purposes of the current experiment, having words that represent items from a particular category was important as prime-target pairs were designed to be category coordinates. Target words were names of concrete items from a variety of semantic categories.

Using the specific terminology for this thesis (Section 1.2.2.4), half (28) of the target words were regular in their vowel spelling-to-sound correspondences and the other half (28) were exception in their vowel spelling-to-sound correspondences. The following sections details the creation of these regularity lists

3.2.3.1.1 Spelling-to-sound correspondences

Words with difficult spelling-to-sound correspondences, as discussed in Chapter 1, were chosen as exception target words. Exception words violated ‘major correspondences’ as defined by Venesky (1970). For example, the major pronunciation correspondence of the vowel pair “ea” is as it is in “beak”. Therefore the word “steak” is in violation of this and for the purposes of this experiment would be classed as exception.[17] Words were considered regular if they followed major correspondences in Venesky. Words were also considered according to consistency with “friends” and “enemies” measures, as discussed in Section 1.2.2.2.

3.2.3.1.2 Friends and enemies analysis

A post-hoc examination of spelling-to-sound consistency of the regular and exception target words lists was undertaken to identify “friends” and “enemies” of target stimuli, as described by Jared et al., (1990) and Jared (1997), and as used by Woollams (2005). One-syllable target words that were included in the database of spelling-to-sound consistency published by Ziegler et al. (1997) were collated. Low frequency regular target words on average had a greater number of friends than enemies and a greater summed frequency of friends than enemies. Eighteen of the eligible 23 regular target words had no enemies. The low frequency exception target word list on average had more enemies than friends and the sum frequency of enemies was greater than the sum frequency of friends. Only three of the eligible “rule-breaking” exception words had no enemies. Therefore, on the whole, exception words were also more inconsistent and had more enemies than regular words.

[17] Words with “ea” as the vowel are an inconsistent “neighbourhood” (See chapter 1), but this is not true of all exception target items in the current experiment, nor are all regular words in the experiment from consistent neighbourhoods.

3.2.3.1.3 Target word lists

Each regularity list (regular and exception) was divided in half to create four lists, two lists of each type of regularity, with 14 target words per list. Counterbalancing of materials across participant made it possible for participants to receive each regularity type in the related and unrelated conditions while only seeing each target word once. The two lists of each regularity type were also matched as far as was possible.

Initial phonemes were matched across regular and exception lists before being divided. Upon being divided, initial phonemes were distributed as evenly as possible across both lists within a regularity type. Multiple words from one category within a regularity type were distributed between the lists, and miscellaneous or unique category items were also divided between lists. Therefore two lists within a regularity type did not have examples from one category only, or several examples that could not be easily categorised. Effort was made to evenly distribute two syllable words between lists within regularity types.

The target words were matched on Kucera and Francis (1967) frequency and Celex frequency (Baayen et al., 1993) across and within regularity type. The two target word types (regular, exception) were also matched on concreteness, imageability, age-of-acquisition, and familiarity (provided by the MRC Psycholinguistic database, Wilson, 1988). The target word types were matched as closely as possible given the constraints on stimuli selection, e.g., the semantic relationship of prime and target. In addition the measure of priming is within word type, in that words are compared against themselves in related and unrelated conditions; therefore, any small differences in the stimulus measures should not have an impact on priming within word type.

The target words were assessed for their category dominance; that is the percentage that the target word was produced in response to the general category, and the percentage the target word was produced first in response to the category. The category norms of Van Overschelde, Rawson, and Dunlosky (2004) were used. Regular target words and exception target words did not differ on either measure.

	Target Word Type		
Measure	Low Frequency Regular	Low Frequency Exception	
Number of Friends per list	342	51	
Number of Enemies per list	10	77	
Average number of Friends per item in list	19	2.83	
Average number of Enemies per item in list	0.56	4.28	
Summed frequency of Friends per list	12,862	5,293	
Summed frequency of Enemies per list	2,214	9,637	*p-values*
KF Frequency	6.41	7.95	*0.24*
Celex Frequency	6.17	6.51	*0.78*
p-values for frequency matching	*0.32*	*0.67*	*n/a*
Concreteness	598.00	602.11	*0.58*
Imageability	576.48	591.00	*0.08*
Age-of-Acquisition	302.42	321.08	*0.47*
Familiarity	494.54	478.00	*0.39*
Number of letters	4.57	5.14	*0.04*
Number of syllables	1.15	1.39	*0.05*
Category dominance: Total produced	32%	31%	*0.94*
Category dominance: First produced	10%	8%	*0.21*

Table 3.1. Mean target stimulus values for Experiment 1 on a variety of measures as a function of target word type with p-values from condition-matching analyses. See text for details.

3.2.3.2 Prime pictures

Prime pictures were black line drawings on a white background from the International Picture Naming Project (IPNP) (Szekely et al., 2004). Related primes were created by pairing each target word with a picture from the same semantic category (e.g., jug-prime and vase-target, not bread-prime and butter-target). Also, as far as was possible items were chosen so as not to be strongly associated. The word association norms of Nelson, McEvoy, and Schreiber (2004) were used to explore whether targets were produced as associates when the prime items were given as a cue. Of the 47 prime items, 36 did not have their targets produced as associates. Of the 11 primes for which their target was produced as an associate, only two were the dominant response. Unrelated primes were created by pseudo-randomising the prime pictures within each list and assigning them to specific target words. Prime pictures from within a specific list now appeared before a target word from an unrelated semantic category.

3.2.3.3 Semantic rating of prime-target pairs

Semantic ratings were collected for related and unrelated prime-target pairs. Ten volunteers, separate to those that participated in the priming study, rated each prime and target pair individually on a scale of 1 to 7. Participants were asked to rate each pair in terms of how similar the two words were in meaning, with 1 being extremely dissimilar in meaning and 7 being extremely similar in meaning.

Average rating for unrelated prime-exception word target pairs was 1.5 on the 7 point scale, and the average rating for unrelated prime- regular word target pairs was 1.3, and there was no significant difference between these two conditions $t(25)= 1.5$, SEM = 0.15, $p=.14$[18]. Average rating for related prime-exception word target pairs was 5.7 on the 7 point scale, and the average rating for related prime-regular word target pairs was

[18] An omission in the rating study meant that ratings for two of the unrelated pairs were unavailable for analysis.

5.8, $t(27) = .32$, SEM = 0.23, $p = .75$. The prime-target pairs were well matched with respect to semantic similarity across regularity conditions.

3.2.3.4 Intervening filler items

Filler items were selected in keeping with Experiment 2 of Vitkovitch et al. (2006). Two filler items were randomly assigned to each target word and stayed with their target word whether in the related or unrelated prime picture condition. In total, there were 112 filler items, half of which were highly picture-able item names as words and half of which were pictures from the IPNP (Szekely et al., 2004). Information about the filler items was provided using N-watch program (Davis, 2005). Fillers appeared in one of the following patterns before a target word: fillers were either both words (three per list), both pictures (three per list), a word followed by a picture (four per list), or a picture and followed by a word (four per list). In this way, half of the target words per list received a picture filler beforehand and the other half received a word filler. The majority of filler items were from a miscellaneous category, e.g., coin, though some were from a specific category that was not represented in the target or primes of the experiment, e.g., ball. With the exception of six items, filler items were one or two syllables in length. The frequency range of filler words using CELEX written frequency (Baayen et al., 1993) was 0 per million minimum to 1068.25 per million with an average frequency of 37.85 per million. Using Kucera and Francis (1967) frequency measures, the frequency range for filler words was zero per million to 1004 per million, with an average of 42.11 per million. Using a strict rule-based judgement of spelling-to-sound regularity (Section 1.2.2.1) 47 of the filler words were regular in their spelling-to-sound correspondence, and 65 of the filler words were exception in their spelling-to-sound correspondence.

Care was taken to ensure that initial phonemes did not overlap between prime, fillers, and targets, so that priming could not be due to phonological overlap, and that none of the items in the set of four could be associate primes. Examples of related and unrelated sequences are presented here and can also be used, generally, for the remainder of the priming experiments. A related sequence with a regular words target is “trousers” (picture), “handcuffs” (filler picture), “cage” (filler word),“vest” (target word). An example of a related sequence with an exception words target is “ring” (prime picture), “camera” (filler word), “dice” (filler word),“brooch” (target word). An example of an unrelated sequence with an exception word target is “tie” (prime picture), “candle” (filler picture), “pram” (filler picture), “sieve” (target word).

3.2.3.5 Counterbalancing of stimuli

The four lists, two for each regularity type, contained 14 prime-filler-filler-target quadruplets. Each participant received all four lists, with two target regularity types in both a related-prime-picture and unrelated-prime-picture condition. For example, one regular word target list of quadruplets would appear in the related prime picture condition and the other regular word target list of quadruplets appeared in the unrelated prime picture condition, and the same for exception word targets. Within regularity-type, which list appeared in the related or unrelated prime condition was counterbalanced across participants; when data collection was completed, the four lists had been presented in the primed and unprimed conditions an equal number of times. Participants were randomly assigned to one of four counterbalances. Prime-filler-filler-target quadruplets were presented together as a set of four, but beyond this, presentation was randomised, so the two regularity conditions and the relatedness of the prime-target pairs were shuffled. All items were viewed only once by each participant.

3.2.4 Procedure

As this study is in part a replication, the procedures from Experiment 2 of Vitkovitch et al. (2006) were followed. Each participant entered the testing room and sat one metre away from the computer screen. Participants were given a microphone and held this close to their mouths. They were instructed to name aloud each item as it came on the screen as quickly and as accurately as they could. They were further instructed that the item could either be a word or a picture. If it was a word then they were to read it aloud, if it was a picture they were to name it aloud. Participants were also informed of the fixation crosses between trials and that they were not to name this. Participants were not informed of the priming paradigm or the relation between items.

E-prime 2.0 (Psychological Software Tools, Pittsburgh, PA) was used to present the stimuli, collect the response time of the participants, and the button press coding of the experimenter. Each stimulus was presented one at a time in the centre of a computer screen and remained on the screen until it was named. Words were presented in a clear black font (century gothic 24 point) on a white screen. The black line drawings also appeared on the same colour background. The software recorded the reaction time of the participant from the onset of the stimulus until the initial sound of the response. The trial sequence was as follows: fixation cross (two second), stimulus (until response), blank screen (three seconds), and an illustration is provided in Figure 3.1. During the blank screen, the experimenter pressed a button to code whether the response was correct, incorrect, or voice key error (from the microphone not registering the response to stutters), and noted down the specific error. Participants named a set of practice trials before beginning the test items. Practice trials consisted of four sets of four items always in a picture prime, two fillers of varying modality, and target word order, and these items were unique to the practice. After naming half of the test items, i.e., 28 sequences of randomly presented prime-target quadruplet sets, participants received a break of five minutes before completing the second half the experiment. After completing the experiment, participants verbally completed a post-experiment questionnaire about the stimuli (see Appendix B). This short series of questions investigated whether participants noticed the priming paradigm, the relationship between primes and targets, or were attempting to predict items. Their responses were noted down by the experimenter. Each testing session lasted no more than 40 minutes.

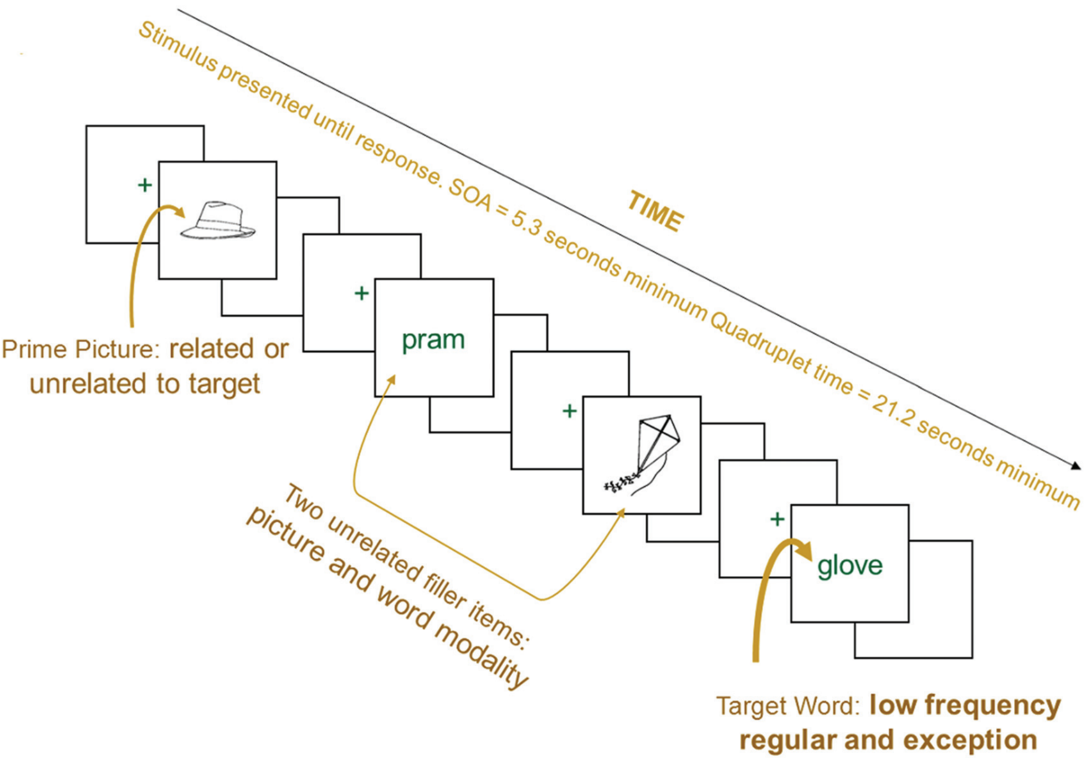

Figure 3.1. Illustrated example of the procedure for one quadruplet of stimuli from Experiment 3 (see text for details)[19].

[19] The images of the hat and the kite are from "A Standardised Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity," by J. G. Snodgrass and M. Vanderwart, 1980, Journal *of Experimental Psychology: Human, Learning, and Memory, 6* (2), p174-215. Copyright by Life Science Associates (LSA). Adapted with permission of the copyright holder.

3.3 Results

Target reading times were removed from the dataset when the prime was incorrectly named (9.2% of target responses), there was a prime voice key error, such as a stutter or microphone malfunction (2.9% of target responses), the target was incorrectly named (7.7% of target responses), or there was a target voice key error (1.7% of target responses). In total 21.5% of target responses were removed.

The percentage of target naming errors was calculated per condition. As mentioned above, the misreading of a target word composed over 7% of total responses. These proportions were used in the errors analysis.

Box plots of mean participant naming times per condition were examined. Based on these, one participant was eliminated from any further analysis as their naming time mean was found to be an outlier in all four experimental conditions. Results are therefore reported for 47 participants[20].

All priming data in this thesis, for this and subsequent experiments, was trimmed using reading reaction time thresholds. Naming times that were longer or shorter than these thresholds were removed. The lower bound threshold was 300ms and the upper bound threshold was 1500ms (Howell, 1992).[21] Reaction times above or below thresholds may not reflect a true reaction time. This was less than 0.7% of possible target responses. These thresholds were not implemented for median reaction times.

[20] Analyses were also performed with 48 participants. They did not differ from the reported results.

[21] As this is a cross modal priming experiment in which pictures were included, a reaction time threshold criteria was used, as this has been employed by other studies in the cross modal priming literature (Damian & Als , 2005; Howard, et al., 2006; Lupker; 1988).

After the removal of the participant outlier and the more extreme reaction times, the reaction time for the exception word data were not significantly skewed as measured in the Kolmogorov-Smirnov and Shapiro-Wilk analyses. Reaction times of the regular target words were significantly positively skewed according to the Shaprio-Wilk analysis. ANOVA and t-test analyses are robust against this type of violation (Field, 2000; Howell, 1992; Kinnear & Gray, 2000; Tabachnick &Fidell, 2007). Data used in the analyses reported here have not been transformed (Howell, 1992). However the same analyses, both ANOVA and t-test, were performed with log linear transformed data and results are the same as those reported below.

To check whether the counterbalancing of materials had an effect, an ANOVA was carried out on reaction time means with counterbalance, of which there were four, entered as a between subjects factor. There were no significant main effects, nor were there any significant interactions with counterbalance order.

The analyses described below were performed on this and the subsequent priming experiments of this thesis. As they are common across experiments, they are only detailed once, here. Analyses of variances (ANOVAs) were carried out on participant (F_p) and item (F_i) naming time means, as is prevalent in the literature. To investigate whether there was a difference between primed and unprimed low frequency exception target word conditions, as there were specific predictions of facilitatory priming in this condition, planned comparisons (t-tests) were also performed on participant mean naming times in the low frequency exception word condition when priming effects were non-significant in the ANOVA (Howell, 1997, p351). Given that Vitkovitch et al. (2006) reported medians, where mean analyses returned a near significant result, analyses were re-calculated using medians, as reaction time data can be (non-significantly) positively skewed. Where medians were analysed, the alpha level was adjusted. Analyses of participant mean error rate were also conducted. When considering the error rate analyses presented Chapters 3-6 of this thesis, it is worth noting that though it is standard to add a constant of 0.001 to the data and then transform it before analysis, the reported error analyses were performed on the original non-transformed percentages.

	Target Word Condition					
	Regular Target Words			Exception Target Words		
	Related	Unrelated	Primed difference	Related	Unrelated	Primed difference
Mean RT	642.88	640.36	+2.52	651.00	654.00	-3.0
SD	140.34	135.07		129.41	138.07	
Median RT	629.27	626.14	+3.13	625.49	636.73	-11.24
SD	146.28	125.64		116.74	132.81	
Error %	3.3%	3.0%	+0.3%	11.9%	12.5%	-0.6%
SD	6.1%	6.6%		11.0%	12.3%	

Table 3.2. Experiment 1 results. Reaction times (RT) reported in ms, error rate (Error %) reported in percentages, and standard deviations (SD) for target words (correct-only) as a function of target word type and priming (related/unrelated) conditions for Experiment 1.

3.3.1 Word reading time analyses

Reaction time means and medians are presented in the table above. Note the primed difference column. There is almost no primed difference for regular target words in both means and medians, and exception word means show a similar pattern. However, the primed difference for exception words in medians reaction times is much larger.

A within-subject ANOVA was initially performed on target naming times with regularity and priming as the 2 (exception, regular) by 2 (related, unrelated) factors. There were no significant main effects of either regularity or priming, regularity $F_p(1,46) = 2.167$, $MS_e = 2568.03$, $p = .15$[22]; priming $F_p(1,46) = .002$, $MS_e = 1744.24$, $p = .97$; $F_I(1,54) < 0.01$, $MSe = 1695.44$, $p = .99$. The interactions of the two factors also failed to reach significance, $F_p(1,46) = .185$, $MS_e = 1942.68$, $p = .67$, $F_I(1,54) = 0.17$ $MS_e = 1695.44$, $p = .68$.

Planned comparisons examined priming within the exception word conditions. Mean reading times showed a very small priming trend in the exception word condition, this trend failed to reach significance, $t_P(46) = 0.393$, $SEM = 7.65$, $p = .35$. However, median reading times that showed a similar pattern for priming approached significance, $t_P(46) = 1.40$, $SEM = 8.03$, $p = .08$, as displayed in Figure 3.2. Exception target words were named, on average, 11ms faster in the related condition than in the unrelated condition.

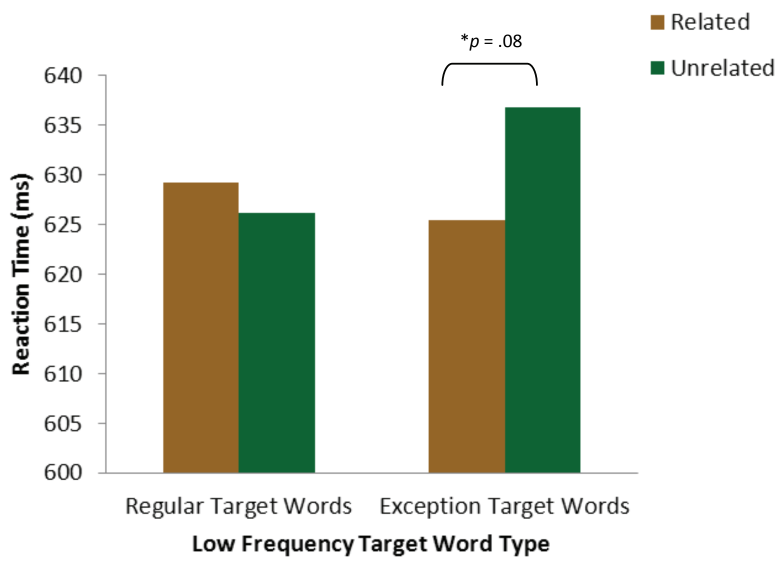

Figure 3.2. Median reaction times of Experiment 1 as a function of target word type and priming (related/unrelated) conditions. The priming effect in the exception word condition approached significance, as is specified in the figure (*).

[22] RT analyses were also performed with log transformed data. The only difference from the reported results was a the regularity main effect that approached significance, $F_P(1,46) = 3.28$, $MS_e = .001$, $p = .08$.

3.3.2 Error rate analyses

Mean error rates were calculated as percentages (see Table 3.1.). The number of misread or mispronounced target words for each participant per condition was divided by the total number of target words in the condition, i.e. 14. Error rates were analysed in a two factor ANOVA.

The main effect of regularity was significant, $F(1,46) = 60.37$, $MS_e=0.01$, $p <.001$, with more errors being made to exception target words. There was, however, no main effect of priming, $F(1,46) =0.015$, $MS_e = 0.01$, $p = .90$. The interaction was also not significant, $F(1,46) = 0.16$, $MS_e = 0.01$, $p = .69$.

3.4 Discussion

Experiment 1 aimed to replicate and extend the semantic priming investigations of Vitkovitch et al. (2006), by including a manipulation of regularity for low frequency target words. Extending the design to include a manipulation of target word types provided the opportunity to investigate whether semantic activation might be observed with only certain word types. Of particular interest were low frequency exception words which with less efficient orthography-to-phonology with which there may be more time for a semantic contribution, as suggested by the triangle model of word reading described in the Chapter 1.

Clearly significant semantic priming effects were not found in Experiment 1, which used only low frequency target words with a manipulation of spelling-to-sound regularity. The main effect of priming and the interaction of priming and word type both failed to reach significance. Vitkovitch et al. (2006), in which the majority of the targets were low frequency regular words, found priming. Additionally a re-analysis of their data showed indications (non-significant trends, only) of priming with the low frequency portion of their target reaction times. Therefore, priming in a low frequency experiment of the same design might be expected.

Though no priming results reached significance at the critical .05 level in the current experiment, low frequency exception word means showed a very small non-significant priming trend, and a comparison of median participant reaction times for low frequency exception target words approached significance. In line with the prediction, when a related prime picture was presented two unrelated filler items before a low frequency exception target word, the target word was read, on average, 11ms faster than in the unrelated condition.

There is, however, an indication of a difference between mean and median analyses as medians showed a larger primed difference that neared significance that was not present in the means. Though mean reaction times were trimmed, it is still possible that there is more variability in these reaction times, than in the medians, e.g. the smallest standard deviation is in medians for the primed low frequency exception condition. Upon examining the distributions of the two measures in the two conditions, medians had a steeper distribution than the means; the means' distribution was wider and flatter. The kurtosis was not significant as the z-score fell within $\pm$ 1.96 (Tabachnik & Fidell, 2007). Using Shapiro-Wilk analysis of normality, mean and median reaction times to low frequency exception words were not significantly skewed (Field, 2000). The means' distribution may stem from (a) the way in which means are calculated, i.e., means were affected by extreme scores (Howell, 1992) and (b) the variance and spread of word reading reaction times, i.e., response times varied individually across participants, but each participant's response times contained wide distribution of scores. This perhaps indicates that there is high variability in the reaction times. As medians are not affected by extreme scores, they likely reflect a genuine priming effect, i.e., possible low frequency exception word priming by a related prime picture.

Post-experiments questionnaire data was examined to determine whether participants were using a strategy. When questioned about related items in the experiment, 10 participants noticed categories of items, (15 participants reported incorrect relationships, e.g. associates or identical pictures and words). Three participants reported noticing a pattern, but this pattern was of modality. Five reported that they attempted to predict items, but they were not guessing prime-targets, but likely modality or associates. Therefore any indication of priming is not likely due to strategic prediction by participants as the majority of participants (42) did not report predicting, none reported correctly predicting any items (including targets), and none of the participants identified the design of the experiment, the quadruplets, or primes and targets.

There is a trend in the means for a regularity effect with exception words being read more slowly than regular word, but this was not significant in mean reaction time analysis. There was, however, a highly significant difference in error rate. Regularity effects, which usually reveal themselves in reaction times, may have instead emerged in accuracy rates. There is a trend for a significant priming effect in the exception word condition, without this priming advantage a main effect of regularity may have been obtained, as reaction times in the unrelated exception word condition are greater than those of the regular word conditions (see Table 3.2.). Also as noted previously, there is variability in the data, and this may have also effected whether regularity effects were found.

The priming results from the current experiment, however, are less than definitive and there is a failure to clearly replicate the findings of Vitkovitch et al. (2006). Methodological differences between the current experiment and Experiment 2 of Vitkovitch et al. could account for the failure to replicate to original results. Differences of note might possibly be a difference in strength of semantic relationship for prime-target pairs and a lack of high frequency target words as stimuli in the current experiment. These and other design elements are discussed in detail in the following chapter, which presents a follow-up semantic priming experiment.

Chapter 4

Semantic Priming Experiment 2

4.1 Introduction

The next experiment sought to capture a significant semantic priming effect in a further attempt to replicate the results of Vitkovitch et al. (2006), as this may provide evidence of a semantic contribution to word reading. In Experiment 1, a significant priming effect was expected, but not clearly found. In creating Experiment 2, changes to the design were made to maximise the potential to clearly find a significant semantic priming effect.

4.1.1 Differences between Experiment 1 and Vitkovitch et al. (2006), and the design of Experiment 2

4.1.1.1 High frequency target words

One difference is the type of target word used, and it is possible that this may have contributed to the difference in effects between Experiment 1 of this thesis and Experiment 2 of Vitkovitch et al. (2006). The Experiment 1 contained only low frequency words as targets, for reasons described previously, whereas the target words of Vitkovitch et al. were ultimately a mixture of word types that included high frequency words, though the majority of target words were low frequency. Despite being few in number, the high frequency targets may have contributed to the significant effect of Vitkovitch et al. (2006).

There are also additional reasons for including high frequency target words. Triangle modellers claim that high frequency words have efficient links. There are especially quick links between orthography and phonology, and also quick connections to and from semantic memory, including efficient links between orthography and semantics, and semantics and phonology (Harm & Seidenberg, 2004). Therefore, it is possible, though unlikely, to find semantic effects, in this experiment marked by semantic priming, with high frequency target words as orthography-to-semantics-to-phonology could be very quick. Including regular and exception high frequency words in an experiment with a low frequency word type manipulation as a within subjects design provides the opportunity to explore a semantic contribution to word reading using all four word types (high frequency regular, high frequency exception, low frequency regular, and low frequency exception), as described in Sections 1.2.3 and 1.2.2.4, and also provide the opportunity to investigate any possible differences between word types in receiving this contribution.

Also, Tse and Neely (2007) claim that high frequency word targets are more predictable than low frequency word targets (Section 2.4.2). They argue that priming of high frequency target words, but not low frequency target words, in their associate pair experiment may implicate conscious expectancy of targets by the participants in a paradigm with associated prime-target pairs. The same could also be true in the current design. Were this pattern of results found in the current investigation, despite the design, it could suggest that priming might be due to non-semantic strategic effects (Section 2.4.1; see Section 4.1.2).

Therefore, in Experiment 2, high frequency words are included as targets in addition to low frequency. There also may be another explanation for the differing results between Experiment 1 and Vitkovitch et al.'s Experiment 2.

4.1.1.2 Semantic relationship strength and lag

Another difference may be that semantic activation from the prime picture to the target word in Experiment 1 with its stimulus pairs did not last over the two intervening items. As original, i.e., chosen for Experiment 1 specifically, stimulus items were used in Experiment 1, it is possible that the relationship between primes and targets was weaker than the pairs of Vitkovitch et al. (2006), even though semantic similarity ratings for related pairs of Experiment 1 were high (Sections 3.2 and 3.3). Finding semantic priming across a long lag may be dependent on the strength of semantic relationship between prime and target, e.g., the overlap in semantic features (Becker et al., 1997; Joordens & Becker, 1997; McRae & Boisvert, 1998). In distributed representation models of semantic memory, semantic priming over a lag is possible, but priming is dependent the prime's representational pattern remaining activated in spite of the fillers' activation and priming may be affected by the number of items (Becker et al., 1997; Hinton & Shallice, 1991; Joordens & Becker, 1997; Masson, 1991; 1995; Plaut, 1995). Therefore, the filler items of Experiment 1, which were original to that experiment, may have replaced all of the semantic activation of the related prime, leaving no remaining activation from which the target word could benefit. This may be more likely if there is a weaker semantic relationship between prime and target, i.e., fewer semantic features in common. This may not have occurred with prime-target pairs and filler items of Vitkovitch et al.'s Experiment 2. Initial conclusions therefore indicate that these reasons, i.e., a weaker semantic relationship and/or the differences in specific filler items used, might account for why clear priming was not found in Experiment 1 of this thesis, when it was found in Experiment 2 of Vitkovitch et al. With fewer filler items, it is less likely that a filler item's representation might replace the semantic activation left by the prime. With the stimuli of Experiment 1, e.g., prime-target pairs, semantic activation for the prime may remain over one item. By decreasing the number of filler items between prime and target (lag) to one intervening filler item in Experiment 2, semantic priming might be found. It also has the additional benefit of reducing the number of items needed for the design so there are enough stimulus items, e.g. filler items, for high frequency targets to be included.

Behavioural literature supports this notion that semantic priming of a word target over one intervening item might be more likely than with two intervening items. Priming, with word primes, has been found with word targets over one intervening item, though with the task of naming this effect is not reliable (Joordens & Besner, 1992; Masson, 1991) (Section 2.4.5.1). In word prime to word target designs, priming with two filler items is often half of that found with one intervening filler item or not found at al (McNamara, 1992, 2005). There could be a chance that priming is detected in Experiment 2 with one filler item when it was not found with a two filler items in Experiment 1.

By removing one filler item, the time between prime and target items will be reduced, and this needs consideration. In Experiment 2, removing an intervening item reduces the time from prime onset to target onset to 10 seconds[23]. In spreading activation theories of semantic memory (Section 2.4.1) semantic activation is short-lived (Anderson, 1993,; Collins & Loftus, 1975). However within these theories, semantic activation is so short-lived that it would not be able to last over 10 seconds (Damian & Als, 2005; Masson, 1995; McNamara, 2005). Therefore, the chance of finding priming due to spreading activation, including the pre-activation of a target word's orthography (a non-semantic locus of priming), should not be increased by reducing the number of intervening filler items from Experiment 1 to Experiment 2. Therefore, keeping at least one intervening item within this priming paradigm is important. It reduces the chances that significant priming might be due to automatic spreading activation and pre-activation of non-semantic systems, i.e., orthography and phonology. However, the design changes, including the reduction of the number of filler items, and filler items in the same modality as the target, may increase the chance of priming due to non-semantic strategic effects. In the event of significant priming effects, design aspects and special analyses are used as tools to argue against explanations of strategic pre-activation (see Section 4.1.2).

4.1.1.3 Word modality filler item

In addition to reducing the number of intervening filler items from two to one, the modality of the filler item must be considered, as there may be an opportunity to tighten the design for Experiment 2. The fillers of Experiment 1 replicated the design of Vitkovitch et al. (2006); half of the target words were preceded by a word filler and the other half were preceded by a picture filler. It is possible that naming a filler picture before a target word, as occurs in Experiment 1 resulted in a switch cost (Monsell,

[23] As items are displayed until response, the precise time is dependent on the participants' response.

2003), meaning target words preceded by a picture filler would be read more slowly than target words preceded by a word filler as it might take time to switch between the processes of picture naming and word reading. This may have created variability in the data of Experiment 1.

An analysis was undertaken to investigate whether switching costs created variability in the data of Experiment 1. Word target naming times were analysed using preceding filler modality as a factor. A within subjects two (related/unrelated) by two (regular/exception) by two (preceding word filler/ preceding picture filler) ANOVA was performed. There was a significant main effect of regularity, $F(1,45) = 8.20$, $MS_e = 3074.91$, $p < .01$, with exception target words being named more slowly than regular target words (exception 650.5 ms, regular 633.9 ms). There was also a trend towards a significant main effect of preceding filler type, $F(1,45) = 3.35$, $MS_e = 3601.18$, $p = 0.08$. Target words preceded by a word filler were (636.5ms) read more quickly on average than when preceded by a picture filler (647.9ms). No other main effects or interactions were significant, including effects with priming.

The post-hoc analysis revealed that switching costs might be present in Experiment 1, and this may have created variability in the data. Though the filler results only approached significance, it is worth noting that the averages per condition were calculated from a maximum of seven items, possibly resulting in a lack of power. Regularity effects, which were originally only present in error rate, were now revealed in reaction times, indicating one reason for a lack of regularity effects in the original analysis may have been variability created by switching costs. However, of importance is that no interactions with priming were revealed. Nevertheless, by taking these results into consideration when designing Experiment 2, variance in the reaction times due to switching costs could be reduced. Therefore, only fillers in word modality will be used in Experiment 2.

4.1.1.4 Blocking of word target difficulty type

As a reminder, Experiment 2 will have both high and low frequency target words, unlike Experiment 1 that only had low frequency target words. In Experiment 2, stimuli triplets (prime, filler, target) will be grouped according to target word type. Experiment 2 will be blocked by frequency, as in Experiment 1, but will also extends the blocking to include regularity type as well, e.g. all low frequency exception target word triplets will appear together. Blocking stimuli by target word type segregates low frequency and high frequency words into separate halves of the experiment (See 4.2.3). Therefore, for half of the participants, low frequency items will appear first, rendering the first half of Experiment 2 like the whole of Experiment 1, i.e., all low frequency word targets appear together and no other word type comes before them. This also allows for analyses that can compare Experiment 1 data with this low frequency first block data (see Chapter 7).

A break in the middle of the experiment offers an opportunity to investigate significant priming explanations (Section 4.1.2). If participants were using a strategy to perform the priming task, then they might read targets in the related prime condition faster in the second half of the experiment. Practice with a successful strategy over the course of the experiment session may improve performance, resulting in greater priming in the second half of the experiment session (Wheeldon & Monsell, 1992). An analysis of session halves (split-half analysis) can be undertaken to identify whether performance improved over the session.

4.1.1.5 Summary of design changes

In Experiment 2, high frequency regular and exception words will be used as a target word types in addition to the low frequency types that were used in Experiment 1. The number of filler items between prime picture and word target will be reduced to one, with the time between prime and target onset a minimum 8.6 seconds. The one intervening filler item will be in word modality. Triplets will also be blocked together by target word type. These design aspects may increase the likelihood of finding a significant semantic priming effect.

4.1.2 Predictions

Given the trend towards priming in Experiment 1, the findings of past research (Plaut et al. 1996; Strain et al. 1995; 2002; Vitkovitch et al.; 2006), and the design amendments, it is expected that priming will be found in the low frequency word conditions of this experiment. Also of interest is whether any semantic priming effects are present for high frequency words. Though the triangle model of word reading favours finding a semantic contribution when orthography-to phonology is slow; a semantic contribution to high frequency word reading is not categorically eliminated. Secondly, Vitkovitch et al. (2006) found significant priming of target words, and though they were in the minority, these targets also included high frequency words. Therefore priming might also be found in high frequency target conditions.

If significant priming is found, then markers of non-semantic strategic processing using design and analysis “tools” will be considered. Design tools include the lack of association between primes and targets (Sections 3.2.3.2, and 4.2.2.2), the lack of category dominance of targets (Sections 3.2.3.1.3 and 4.2.2.2), and the covert nature of the priming design. Analysis tools include inspecting the responses to the post-experiment questionnaire to see whether there is verbal evidence of specific predictions of items or the priming design. Additionally, whether high frequency target priming *alone* is found will be considered (Section 4.1.1.1) (Tse & Neely, 2007). Also, a split-half analysis can be undertaken in order to investigate whether there is greater priming in the second half for each freq type (Section 4.1.1.4), as would be expected if participants were gaining practice over the session at actively predicting.

4.2 Methods

4.2.1 Participants

Forty-eight volunteers (38 female) from the University of East London, none of whom had taken part in any other investigation of this thesis, participated in this experiment and were entered into a prize draw for a gift voucher in exchange for their time. Participants were aged between 19 and 48 years old with an average age of 28 years. All reported speaking English as their first language and normal or corrected-to-normal vision.

4.2.2 Stimuli and design

The stimuli were used to create a picture-prime-to-word-target semantic priming paradigm similar to Experiment 1 of this thesis, though with some design changes. Experiment 2 triplets were presented in the order of: prime picture, filler word, target word (PWW). The target words were one of four types: high frequency regular, high frequency exception, low frequency regular, and low frequency exception. Each of these appeared in the related and unrelated priming conditions, counterbalanced across participant. Target words were locked to their filler item and were presented with the filler item whether the target was in the primed (related picture prime) or unprimed (unrelated picture prime) condition. Complete stimulus lists are available in Appendix C.

4.2.2.1 Target words

Target words were names of concrete items from a variety of semantic categories. In total there were 80 target words: 40 high frequency and 40 low frequency. Within each frequency type, half of the words were regular in spelling-to-sound correspondence and the other half were exception (Section 3.2.3.1.1). Values for the various stimulus measures were in the main provided by the MRC psycholinguistic database (Wilson, 1988), unless otherwise specified, and are provided in Table 4.1 with p-values for matching analyses.

4.2.2.1.1 High frequency target words

Similar to the target word stimuli of Experiment 1, high frequency target words were high in imageability and “picture-able”. Forty high frequency words, 20 of which were regular and 20 of which were exception were used in this experiment. Twenty was the maximum number of high frequency exception words with a category coordinate picture, and all of these items were included in Experiment 2.

High Frequency target words had a minimum Kucera and Francis (1967) written frequency count of 14 per million, average 105.97 per million and a minimum Celex lexical database count of 23 per million, average of 123.20, (Baayen et al., 1993 provided by MCWord, Medler & Binder, 2005). The high frequency regularity lists, i.e. exception and regular lists, were matched on Kucera and Francis frequency counts and on Celex frequency counts. The two high frequency lists were also matched on length in letters and syllables. The majority of words (35 out of 40) were one syllable and none were more than two syllables in length. Regular and exception high frequency target words were matched in terms of initial phoneme as far as was possible.

High frequency targets were retrospectively subjected to a post-hoc examination of spelling-to-sound consistency of the regular and exception target words lists. "Friends" and "enemies" of target stimuli were identified (Section 1.2.2.2, Section 3.2.3.1.2). An examination of the number and the summed frequency of friends and enemies was performed. High frequency regular target words on average had a greater number of friends than enemies and a greater summed frequency of friends than enemies. Fifteen of the 20 regular target words had no enemies. High frequency exception target words on average had more enemies than friends. Only one of the eligible "rule-breaking" exception words had no enemies.

4.2.2.1.2 Low frequency target words

From Experiment 1, 40 low frequency words, 20 of which were regular in their spelling-to-sound correspondences and 20 of which were exception in their spelling-to-sound were used in Experiment 2. The number of low frequency target words was dictated by the maximum number of words available in the high frequency exception word condition (20), which was the most difficult condition to recruit items for.

Therefore, the number of low frequency targets was limited in order that all conditions might an equal number of target words.

This necessary limitation meant that some low frequency prime-target pairs from Experiment 1 had to be eliminated. Pairs that had a higher number of errors in Experiment 1 were eliminated, including pairs in which there were a high number of prime errors. For example, the target word "salmon" was eliminated because participants struggled to name the related prime picture of an "eel". This also allowed for the chance to omit the target word "wool", which had a higher Celex frequency than other Experiment 1 target words. Low frequency regular and low frequency exception word lists were matched on Kucera and Francis and Celex frequency counts. They were also match on familiarity, imageability, and concreteness.

As in Experiment 1 and the high frequency condition of the current experiment, a *post-hoc* examination of "friends" and "enemies" of regular and exception low frequency target words was undertaken, examining the number and the sum-of-frequency of friends and enemies. In keeping with the description of the larger list (Section 3.2.3.1.2), this subset of low frequency regular target words on average had a greater number of friends than enemies and a greater summed frequency of friends than enemies. Low frequency exception target words had more enemies than friends and a greater summed frequency for enemies than friends.

4.2.2.1.3 Target stimulus lists

High and low frequency target word lists were matched on concreteness, length in letters, and number of syllables. The two frequency types did significantly differ on both Kucera and Francis and Celex frequency counts, as expected. Category norms were also used to inspect the category dominance of the target words. High frequency and low frequency target words neither differed in the amount they were produced in response to their given category, nor did they differ in the amount they were produced as the first response to their given category.

As noted in Chapter 1, the measures of frequency and regularity can be correlated with other variables. Controlling all correlated measures while also orthogonally manipulating frequency and regularity is difficult. As in Experiment 1 of this thesis, target word types were matched as far as was possible given the constraints on stimuli selection. The stimulus lists of Experiment 2, however, did statistically differed on some factors. Low frequency regular and low frequency exception lists differed on number of letters and on age-of-acquisition. High frequency regular and high frequency exception lists differed on age-of-acquisition, as well as other measures. Additionally high frequency and low frequency lists differed on age-of-acquisition and familiarity. There is independent evidence that age-of-acquisition, familiarity, and length are biased in such a way as to make orthography-to-phonology computation slower in the "more difficult" low frequency exception word condition (Section 1.2.3), e.g., low frequency exception words are longer in letters and higher (later/older) in age-of-acquisition than low frequency regular words and high frequency target words. Therefore "difficult" words chosen as low frequency exception words have a more difficult orthography-to-phonology than other words types, but this could be due to a longer length (though length effects are negligible in shorter words, Damian et al., 2010; Young and Ellis, 1985) or an older age-of-acquisition. Likewise the "easy" high frequency regular words may be easy because of their lower (earlier) age-of-acquisition. That the four word types may be easy or difficult due to factors other than frequency and spelling-to-sound correspondence is considered further in the discussion chapters of this thesis (Chapters 7 and 10). The four word types will still be referred to using the usual nomenclature, e.g. low frequency exception words. Though the various word types were not matched on all factors, target word reaction times are compared against themselves in related and unrelated conditions, so any significant priming effects within word type cannot be due to unmatched stimulus properties. However, as the stimuli are not matched across the four word type conditions on all measures, care must be taken when considering any difference in priming effects between the various word types.

Each regular and exception word list for high and low frequency types was divided to create eight lists of with an equal number of stimuli. There were two lists of each frequency by regularity with 10 target words in each list, i.e., two lists of 10 low frequency regular words, two lists of 10 low frequency exception words, two lists of 10 high frequency regular words, and two lists of 10 high frequency exception words.

	Target Word Type						
Measure	High Frequency Regular	High Frequency Exception		Low Frequency Regular	Low Frequency Exception		
Number of Friends	166	51		285	40		
Number of Enemies	10	86		7	60		
Average number of Friends per item	766.50	328.06		493.84	222.07		
Average number of Enemies per item	62.28	324.82		63.16	302.07		
Sum frequency of Friends	13797	5905		9383	3331		*Lf vs HF*
Sum frequency of Enemies	1121	5522	*p-values*	1200	4531	*p-values*	*p-values*
KF Frequency	98.47	113.10	*0.62*	6.47	7.81	*0.28*	*0.001*
Celex Frequency	127.06	119.35	*0.84*	6.81	6.27	*0.72*	*0.001*
p-values	*0.005*	*0.49*	*n/a*	*0.20*	*0.40*	*n/a*	*n/a*
Concrete	609.39	589.17	*0.01*	601.63	602.00	*0.57*	*0.34*
Imageability	609.39	586.60	*0.01*	577.63	591.07	*0.34*	*0.08*
Age-of-Acquisition	209.27	246.40	*0.04*	293.00	325.11	*0.00*	*0.001*
Familiarity	594.44	565.28	*0.02*	503.24	482.75	*0.42*	*0.00*
Number of letters	4.15	4.75	*0.08*	4.45	5.05	*0.01*	*0.1*
Number of syllables	1.11	1.15	*0.58*	1.10	1.25	*0.08*	*0.29*
Category dominance: Total produced	44%	38%	*0.19*	33%	35%	*0.80*	*0.30*
Category dominance: First produced	18%	7%	*0.11*	12%	11%	*0.80*	*0.48*

Table 4.1. Mean target stimulus values for Experiment 2 on variety of measures as a function of target word type with p-values from condition-matching analyses. See text for details.

4.2.2.2 Prime pictures

Related and unrelated picture primes for high frequency word were assigned in the same manner as in Experiment 1 (Section 3.2.3.2). High frequency target words were paired with a picture from the same semantic category to be its related prime. Unrelated primes were created by mixing the list of related prime pictures with unrelated target words.

Also, as far as was possible items were chosen so as not to be strongly associated. Low frequency target words remained with their related and unrelated primes from Experiment 1. The word association norms of Nelson, McEvoy, and Schreiber (2004) were used to explore whether targets were produced as associates when the prime items were given as a cue. Of the 76, out of 80, prime items that were listed in the database, 47 not did have their target words produced as an associate by participants. Of the primes that had their target produced as an associate, only four, two in each frequency type, were the dominant response.

4.2.2.3 Semantic similarity rating of prime-target pairs

As with the low frequency prime-target pairs in Experiment 1, semantic ratings were collected for high frequency related and unrelated prime-target pairs. Ten volunteers, unique from those participating in other studies of this thesis, rated each high frequency prime and target pair on a scale of one to seven for semantic similarity. Participants were asked to rate each pair in terms of how similar the two items were in meaning, with one being extremely dissimilar in meaning and seven being extremely similar in meaning.

The average rating for related high frequency exception prime-target pairs was 4.92 on the 7 point scale. The average rating for related high frequency regular prime-target pairs was 4.98. The two lists for each high frequency regularity condition did not significantly differ: regular lists $t(9)= .860$, SEM = 0.30, $p=0.41$, exception lists $t(9)= .626$, SEM = 0.42, $p=0.55$. Related pairs in the two regularity conditions did not differ in semantic similarity, $t(19) = .225$, SEM = 0.27, $p=0.82$. Average rating for unrelated high frequency exception pairs was 1.66 on the 7 point scale. The average rating for unrelated high frequency regular pairs was 1.73.[24] The two unrelated lists did not differ on semantic similarity rating, $t\ (17) =.301$, SEM = 0.25, $p=0.77$. As would be expected, related high frequency regular pairs significantly differed from unrelated high frequency regular pairs, $t(17) = 10.77$, SEM = 0.30, $p < 0.001$, as did the exception word related and unrelated pairs, $t(17) =10.15$, SEM = 0.32, $p < 0.001$.

As there were semantic similarity ratings for high and low frequency prime-target pairs, albeit from different samples of participants, analyses were performed to compare the semantic similarity ratings of the high frequency and low frequency prime- target pairs. Independent sample t-tests were performed on the ratings for the 20 low frequency pairs that were used in Experiment 2, but were collected for Experiment 1 (Section 3.2.3.3), and the ratings of the high frequency pairs collected

[24] All high frequency related pairs were presented for rating. However, due to changes in stimuli choice between the time of rating and the time of the data collection for Experiment 2 not all unrelated pairs were rated. Ratings were collected for 36 of the 40 unrelated pairs and information is reported for the 36 unrelated pairs used in Experiment 2.

for Experiment 2. High and low frequency words significantly differed in semantic similarity, $t(78) = 4.24$, SEM = 0.14, $p < 0.001$. Low frequency related prime-target pairs have a higher semantic similarity score on average than high frequency pairs, 5.7 and 4.95 out of 7, respectively. This difference in relation to the results is raised in the discussion (Section 4.4).

4.2.2.4 Intervening filler words

As noted in the introduction of this chapter, there was only one intervening filler item in word modality. Filler words were chosen and assigned according to the parameters described in Experiment 1 (Section 3.2.3.4). The frequency range of filler words using CELEX written frequency was 1.69 per million minimum to 255.48 per million with an average frequency of 35.04 per million. Using Kucera and Francis frequency measures, the frequency range for filler words was 0 per million to 406 per million, with an average of 36.04 per million. These numbers were provided using N-watch program (Davis, 2005). Using a strict rule-based judgement of spelling-to-sound regularity (Section 1.2.2.1), as provided by the N-watch program (Davis, 2005), 45 of the filler words were regular, and 35 of the filler words were exception.

4.2.2.5 Counterbalancing of materials

Each of the eight stimuli lists, two high frequency regular lists, two high frequency exception lists, two low frequency regular lists, and two low frequency exception lists, contained 10 prime-filler-target triplets. Participants received all eight lists with one list of a word type in the related condition and the other in the unrelated condition, as detailed in Section 3.2.3.5. Which of the two lists appeared in the related or the unrelated condition was counterbalanced across participants so that when data collection was completed the eight lists had been presented in the related and unrelated conditions an equal number of times. Participants were randomly assigned to one of 16 counterbalances (See Appendix D). The experiment was blocked together by word type triplets (high frequency regular, high frequency exception, low frequency regular, low frequency exception) and related and unrelated triplets were presented in a random order within each word type. Participants were not made aware of the blocks or word type manipulation until they had completed the experiment. The counterbalances were such that frequency was also blocked so that one frequency type was either the first or the second half of the experiment. This was also counterbalanced. All items were viewed only once by each participant.

4.2.3 Procedure

The procedure was the same as in Experiment 1 (Section 3.2.4) unless detailed here. The trial sequence was as follows: fixation cross (one second), stimulus (until response), blank screen (three seconds), and an illustration is provided in Figure 4.1. Practice trials consisted of four sets of three items always in a picture, word, and word order, and these items were unique to the practice. After naming half of the test items (40 triplets), participants received a break of five minutes before completing the second half the experiment. Each testing session lasted no longer than 40 minutes.

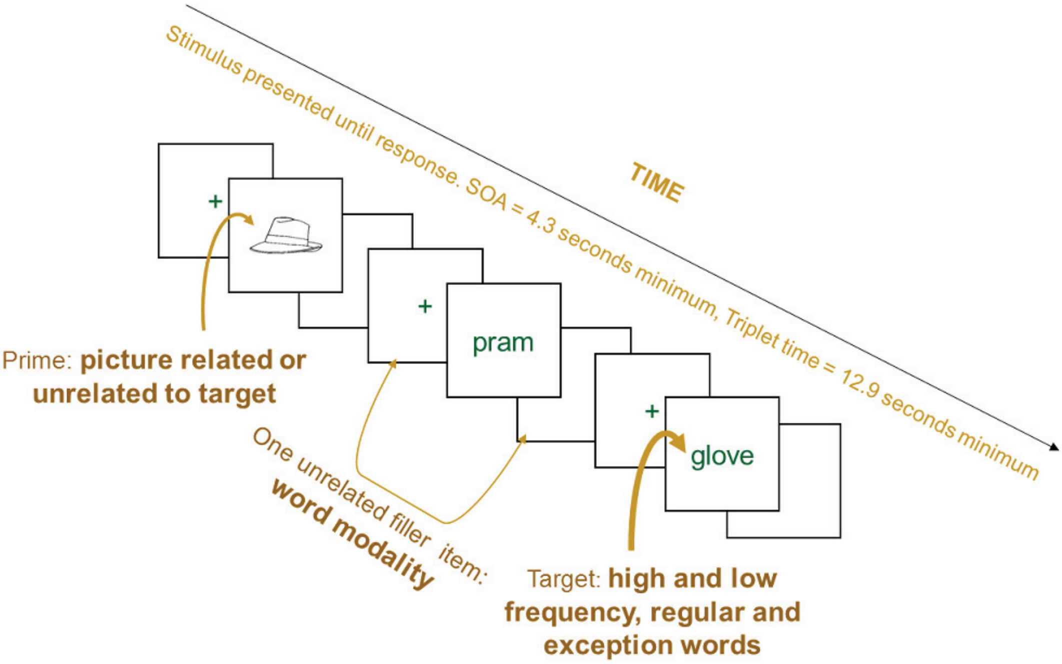

Figure 4.1. Illustrated example of the procedure for one triplet of stimuli from Experiment 2 (see text for details)[25].

4.3 Results

Analyses for Experiment 2 were carried out as in Experiment 1 (Section 3.3). Results are reported for 48 participants.

Naming times were removed from analysis when the prime was incorrectly named (8.1% of target responses), there was prime voice key error in the way of a stutter or microphone malfunction (2.4% of target responses), the target was incorrectly named (3% of target responses), and there was a voice key error when naming a target (1.2% of target responses). In total 14.7% of target responses were removed in total. A

[25] The image of the hat is from "A Standardised Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity," by J. G. Snodgrass and M. Vanderwart, 1980, Journal *of Experimental Psychology: Human, Learning, and Memory, 6* (2), p174-215. Copyright by Life Science Associates (LSA). Adapted with permission of the copyright holder.

threshold for mean naming times was implemented as in Experiment 1 and 1.3% of possible target response times were removed.

The percentage of target naming errors was calculated per condition. The number of misread or mispronounced target words per condition was divided by the total number of target words in that condition (10). The misreading of a target word composed 3% of total responses. There percentages were used in the errors analysis.

Two analyses were performed to determine whether the counterbalancing of materials affected the results. The blocking of triplets according to target word type and subsequent counterbalancing of materials meant that low frequency appeared before high and vice-versa. Therefore, the first analysis investigated whether receiving one frequency before the other affected the results. No significant interactions with frequency order were found. The second analysis checked that full counterbalancing of materials had no effect. An ANOVA was performed on reaction time means with counterbalance, of which there were sixteen, entered as a between subjects factor. There were no significant interactions with material order. These analyses were also performed in subsequent priming experiments of this thesis, and will only be reported if the results are significant to a $p = .05$ level.

	Target Word Condition											
	High Frequency Target Words						Low Frequency Target Words					
	Regular Target Words			Exception Target Words			Regular Target Words			Exception Target Words		
	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference
Mean RT	514.24	521.83	-7.59	542.07	553.51	-11.44	545.59	556.70	-11.11	561.28	571.19	-9.91
SD	89.67	97.02		93.69	95.47		99.08	98.05		106.06	111.40	
Median RT	509.66	522.90	-13.24	535.65	546.35	-10.70	537.73	553.75	-16.02	553.06	566.26	-13.2
SD	90.12	99.51		98.97	96.21		95.71	95.47		115.89	118.51	
Error %	0.0%	0.4%	-0.4%	3.8%	5.6%	-1.8%	0.8%	0.6%	+0.2%	6.0%	6.7%	-0.7%
SD	0.0%	2.0%		7.0%	7.0%		3.0%	2.0%		6.0%	8.0%	

Table 4.2. Experiment 2 results. Reaction times (RT) reported in ms, error rate (Error %) reported in percentages and standard deviations (SD) for target words (correct-only) as a function of target word type and priming (related/unrelated) conditions for Experiment 2.

4.3.1 Word reading times analyses

Within subject ANOVAs were carried out with frequency, regularity, and priming as the two (high, low) by two (regular, exception) by two (primed, unprimed) factors. The three main effects reached significance. There was a significant main effect of frequency with high frequency words being read faster than low frequency word targets, $F_p(1,47) = 16.65$, $MS_e = 3830.27$, $p < .001$, $F_i(1, 76) = 6.88$, $MS_e = 4317.41$, $p = .011$. There was also a significant regularity effect with regular words being read faster than exception words, $F_p(1,47) = 13.53$, $MS_e = 3567.48$, $p = .001$, $F_i(1, 76) = 7.49$, $MS_e = 4317.41$, $p = .008$. There was a significant priming effect with targets in the related condition being named faster than targets in the unrelated condition, $F_p(1,47) = 8.38$, $MS_e = 1149.06$, $p = .006$, $F_i(1,76) = 8.17$, $MS_e = 1070.03$, $p = .005$. No significant interactions were found, frequency x regularity $F_P(1,47) = 2.23$, $MS_e = 2314.27$, $p = .14$, $F_i(1, 76) = 0.28$, $MS_e = 4217.41$, $p = .60$, frequency x priming $F_p(1,47) = 0.02$, $MS_e = 1650.63$, $p = .90$, $F_i(1, 76) = 1.51$, $MS_e = 1070.03$, $p = .22$, regularity x priming $F_p(1,47) = 0.05$, $MS_e = 814.00$, $p = .82$, $F_i(1,76) = 0.17$, $MS_e = 1070.03$, $p = .68$, frequency x regularity x priming $F_p(1,47) = 0.11$, $MS_e = 1428.04$, $p = .75$, $F_i(1,76) = 0.05$, $MS_e = 1070.03$, $p = .83$. As the ANOVAs produced significant main effects of priming below a .05 level, planned comparisons of low frequency exception word targets were not performed.

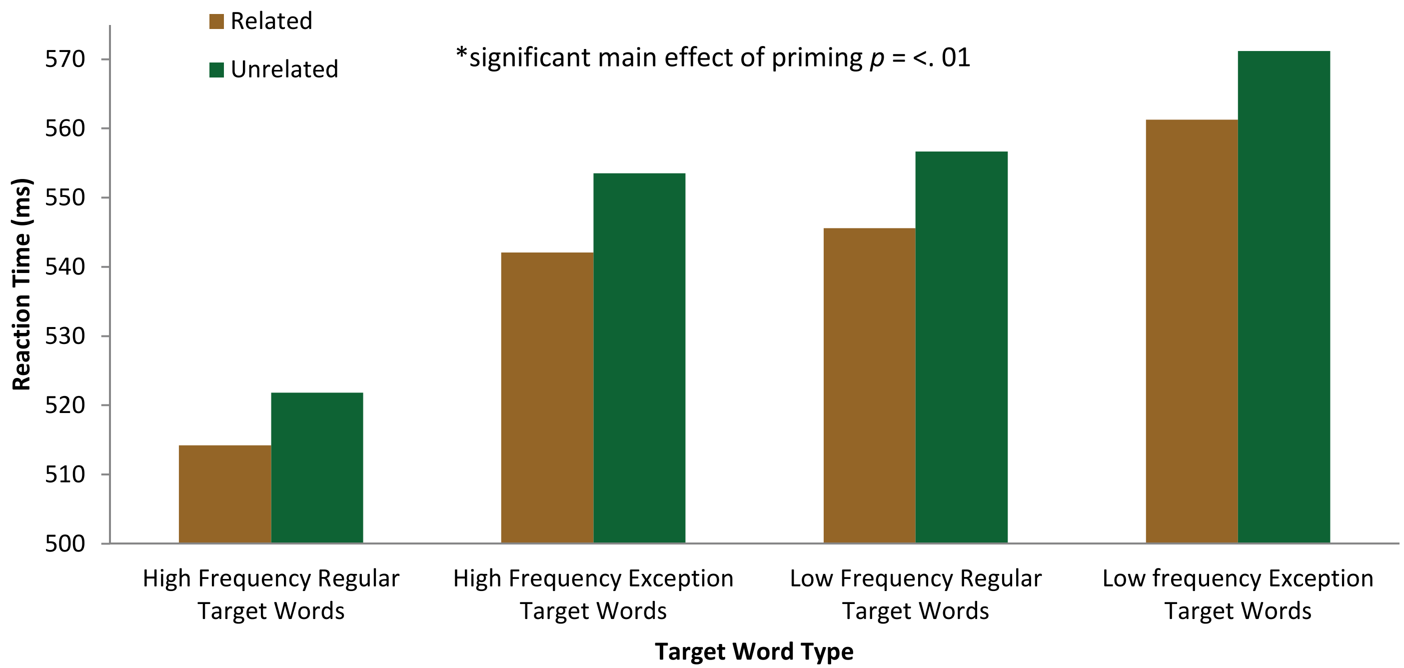

Figure 4.2. Mean participant reaction times of Experiment 2 as a function of target word type and priming (related/unrelated) conditions. There was a significant main effect of priming, as is noted in the figure (*).

4.3.2 Error rates analyses

Participants' mean error rates to targets were subjected to the same ANOVA with frequency, regularity, and priming as the two by two by two factors. There were more errors to exception than regular words, $F(1,47) = 77.43$, $MS_e = 0.003$, $p < 0.001$. There was a marginally significant trend for more errors to low frequency than high frequency words, $F(1,47) = 3.55$, $MS_e = 0.003$, $p = .07$. The main effect of priming, however, failed to reach significance, $F(1,47) = 2.78$, $MS_e = 0.002$, $p = 0.10$. The interactions also were not significant, frequency x regularity $F(1,47) = 1.05$, $MS_e = 0.003$, $p = .31$, frequency x priming $F(1,47) = 0.78$, $MS_e = 0.003$, $p = .38$, regularity x priming $F(1,47) = 2.66$, $MS_e = 0.001$, $p = .11$, frequency x regularity x priming $F(1,47) = 0.09$, $MS_e = 0.003$, $p = .77$.

4.4 Discussion

Experiment 2 aimed to replicate a word priming effect, as published by Vitkovitch et al. (2006), and to extend the design of Experiment 1 of this thesis to investigate four word difficulty types. The main effects of frequency, regularity, and, most importantly to the aim of this thesis, priming were significant.

The main effects of frequency and regularity demonstrate that participants were slower to read low frequency words than high frequency words and were slower to read exception words than regular words. The unrelated low frequency exception word condition had the slowest response times and the highest error rate, though this interaction of frequency and regularity did not reach significance, possibly due to the significant priming effect.

Additionally, there was significant semantic priming of all target word conditions, as revealed by the main effect of priming. Target words were read faster after naming a related prime picture and one intervening filler word, as compared to the unrelated prime picture condition.

The current experiment might have provided evidence of a semantic contribution to word reading through significant priming effects, and there was no clear evidence in the results of a difference in priming effects across the various word types. Significant priming was predicted for the low frequency target words, and not necessarily eliminated as a possibility for the high frequency target word condition. Though the word types were not matched on all measures, and these differences could be crucial when considering differing results across the word types (Becker et al., 1997; Hines et al., 1986; Joordens & Becker, 1997; Plaut, 1995), but no differences were found; the data from Experiment 2 revealed priming within all word types with a significant main effect.

It is possible that this design, which has removed one filler item, could be more prone to strategic effects than Experiment 2's design, even though there is still a long interval between prime and target (10 seconds), which would make automatic pre-activation of the target's orthography and phonology very unlikely. The nature of the significant priming effect, and whether it is non-semantic and strategic in nature or instead due to remaining semantic activation from the prime, can be assessed using a set of tools comprised of design aspects and analyses (Section 4.1.2). First, the assessment of priming in the two frequency types may provide information. Tse and Neely (2007) suggested that priming in a high frequency condition may indicate that effects are strategic in nature, as these words may be easier to predict than low frequency words

(Section 4.1.2). If participants were predicting, then priming might only be expected in the high frequency target words (Tse & Neely, 2007). In the current experiment, priming was found in the high frequency *and* low frequency word conditions. Secondly, the stimuli themselves, regardless of frequency, would have been difficult to predict. Primes and targets were not associates in word association norms (Section 4.2.2) (Nelson et al., 2004). High frequency and low frequency words did not differ in category dominance (Section 4.2.2.2) (Van Overschelde et al., 2004).

Post-experiments questionnaire data was examined to determine whether strategic pre-activation could explain the significant main effect. When questioned about related items in the experiment, 11 participants noticed categories of items, (20 participants reported incorrect relationships, e.g. associates or identical pictures and words). Fourteen reported noticing pattern, but this was a pattern of modality, and this identification was not correct, with seven reporting an order of WWP (the order was, in fact, PWW). Seven reported that they attempted to predict items, though of the five who elaborated on the prediction were not guessing targets or the priming pattern, but guessing modality or associates. Therefore any indication of priming is not likely due to strategic prediction as none of the participants reported correctly predicting any items (including targets), and none of the participants identified the design of the experiment, or primes and targets.

Finally, a split-half analysis was undertaken (Section 4.1.1.4). If participants were strategically guessing, then greater priming in the second half would be expected, because participants gain practice across the entirety of the session. Low frequency targets in the second half of the experiments were not primed more than low frequency targets in the first half; this was also true of high frequency word targets. As far as can be assessed using these design and analysis tools as markers, it is unlikely that the semantic priming effect in Experiment 2 is attributable to non-semantic strategic prediction by the participants.

Semantic priming in Experiment 2 is more than likely due to shared semantic information between the prime itself and target, yet, this was only found with one intervening item and no difference was found between word types. Priming effects over an increased lag may add weight to arguments in favour of a semantic interpretation of priming, as a design with two intervening filler items might have less strategically predictable design than Experiment 2. It may also provide more information about the four word types. Whether significant priming can be clearly captured over two intervening filler items, as it was in Vitkovitch et al. (2006), with these prime-target stimuli is also still of interest. Therefore in the next experiment uses the same prime-target pairs but returns the lag to two filler items. Before beginning a new investigation with a greater lag and four target word types, there is data available from Experiment 1 that may provide additional information.

4.4.1 Reanalysis of Experiment 1 data

As the 40 low frequency prime-target pairs used in Experiment 2 were a subset of those from Experiment 1, the data of Experiment 1 can be reanalysed to investigate whether priming is evident over a two item lag. Of interest is whether effects found in Experiment 2, priming of both regular and exception low frequency words can be clearly seen in Experiment 1 when the data from the same subset of stimuli are analysed. As a reminder, the results of Experiment 2 differed to those of Experiment 1. In Experiment 1, low frequency targets failed to show a clear priming effect; with low frequency exception words, the priming effect approached significance with participant medians, but the priming pattern in participant means was not significant. However, there were signs of variability in the data, as mean and median analyses produced different results. In Experiment 2, with similar stimuli, there was priming of all word types.

Analyses of Experiment 1 data were performed on the 40 low frequency word targets that were used in Experiment 2. Therefore, the reaction times to 16 target words, four from each stimulus list, were omitted from each Experiment 1 participant's data. T-test analyses were performed with participant means and medians and item means for regular and exception words. Because this is a further analysis of Experiment 1 data, critical alpha was adjusted; results will only be considered significant if $p \leq .025$. Priming of low frequency regular words failed to reach significance, $t_p(46) = 0.95$, SEM = 14.33, $p= 0.17$, $t_i(19) = 0.28$, SEM = 12.38 , $p= 0.78$. There was, however, a marginally significant priming of low frequency exception target words. In Experiment 1, for the subset of targets that were used in Experiment 2, when an exception target word is preceded two items earlier by a related prime picture, it is read faster, at a nearly significant level, than when in the unrelated prime condition, $t_p(46)=1.95$, SEM = 9.39, $p= .03$, $t_i(19)=1.5$, SEM = 10.22, $p= 0.08$. See Figure 4.3.

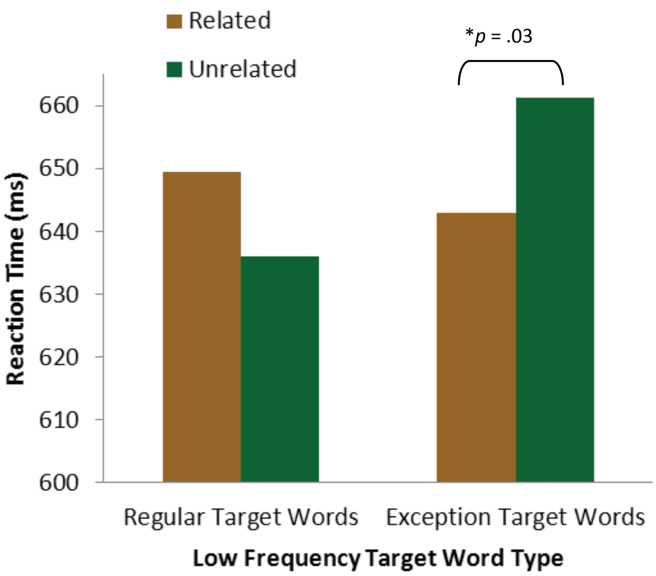

Figure 4.3. Mean participant reaction times of Experiment 1, including only stimuli used in subsequent experiments, presented as a function of target word type and priming (related/unrelated) conditions. The priming effect in the exception word condition was marginally significant, as is specified in the figure (*).

In keeping with the trend in the original analysis, a marginally significant low frequency exception word priming was found in Experiment 1 data when only the stimuli used in Experiment 2 were analysed and a were corrected for multiple comparisons. There may be variability in Experiment 1 data (Section 3.4). When removing certain prime-target pairs from the analysis, as in the current re-analysis, the original trend becomes a nearly significant effect. The subset of prime-target pairs used in Experiment 2 was chosen by eliminating pairs that were higher in errors. First, by excluding these pairs, targets contributing a lower number of reaction times to the mean calculation, due to a higher number of errors, were eliminated. Secondly, though only correctly named pairs were used in the target reaction time analyses, the prime picture of an error prone pairs may have been more difficult, possibly creating variability in reaction times for the quadruplet, including the target item. The data might show evidence of this. When examining the descriptive statistics for Experiment 1 exception words, the standard deviations are lower in the item reanalysis than in the original item analysis. Regardless of the reason, removing these pairs reduced the variability in the data, resulting in nearly significant priming effect in low frequency exception words in Experiment 1[26], even with a strict correction for multiple comparisons. Therefore, there is some evidence in the data of Experiment 1 that priming can be found over 2 intervening items with these prime-target pairs. However, there are still differences between the results of Experiments 1 and 2.

4.4.2 Discussion of Experiments 1 and 2

The results of Experiment 1 and Experiment 2 might demonstrate that semantic priming can be found with low frequency exception word targets. Following the reanalysis of Experiment 1 data, replicating the almost significant priming effect of low frequency exception words over two intervening filler trials in a further experiment is of interest. The priming results between Experiments 1 and 2 differed for low frequency regular word targets, even though the words were, in the main, identical; this also needs further investigation. Additionally, Experiment 2 showed priming of high frequency word types as well. As high frequency target words were not used in Experiment 1, whether high frequency word target priming can be found over two intervening items is still an empirical question. Therefore, priorities for the next experiment are replicating the priming effect for low frequency exception words, understanding the origin of the differing results for low frequency regular word targets, and exploring high frequency priming further. Examining the changes in design from Experiment 1 to Experiment 2 may explain the difference in the results between these two experiments and may help when designing the next experiment.

[26] However, as a reminder, in Chapter 3, t-test analyses were performed on log transformed low frequency exception word data, but the results did not differ from the non-transformed results.

Experiment 1 and Experiment 2 differed in a few ways including the inclusion of "easy" (high frequency regular) and "difficult" (exception word targets) words and the extension of the blocking of targets words by types by extending the blocking to also include regularity types not only frequency, as in Experiment 1. However, a crucial difference between Experiments 1 and 2 concerns what occurs between the prime and the target. The central differences involves inconsistencies in lag, including the modality of intervening filler item, and the length of lag, i.e., number of items and length of time between prime and target.

As discussed previously, two main ways to account for significant priming effects are either through shared semantic information, as argued for with this thesis' effects thus far, or through non-semantic pre-activation of the target word's orthography and phonology. Priming explained by either account could be affected by the differences in lag, the modality of filler, and the amount of time and of items between primes and targets, from Experiment 1 to 2.

The lag in Experiment 2 was shorter in number of items and length of time as compared to Experiment 1, and these are confounded in the experiments. There was one intervening filler word in Experiment 2, in contrast to the two fillers of either picture or word modality in Experiment 1. The SOA between prime and target in Experiment 2 was a minimum of 8.6 seconds, in contrast to in Experiment 1 this was a minimum of 15.9 seconds. These changes may have resulted in the main effect of priming in Experiment 2, as the design changes may have created a scenario in which either (1) the priming paradigm, including the targets, was more predictable, though this is not likely, as argued in a previous section, (2) the picture prime was more salient or (3) semantic activation could survive for the target, and these are discussed in turn.

4.4.2.1 Possible effects of differences in filler item modality

In Experiment 2, stimuli were presented in the repeated order of picture prime, word filler, word target, whereas in Experiment 1 stimuli were in PFFW order, where fillers could either words or pictures. Therefore one difference between Experiments 1 and 2 was lag modality. Lessening the number of filler items to only one *word* may have affected the predictability of this covert priming paradigm, though as argued in previously in this discussion, this is not likely. Amongst various design elements to lessen predictability, related and unrelated triplets were randomised with one another, and participants did not know the nature of the experiment; therefore, target words were no more explicitly predictable than in Experiment 1, so priming due to prediction is not likely.

The PWW design of Experiment 2, however, could have created a scenario in the picture prime was more salient as the only picture stimuli in Experiment 2 were the primes, and this could have a semantic benefit. It is possible that participants were unaware of primes, targets, and the priming design of the experiment, as is concluded using the "tools", but that the picture prime still may have seemed more relevant or more salient to participants, than pictures primes in Experiment 1 making participants focus on the picture. In over-thinking the prime picture, giving it weight may have somehow kept the prime's semantic activation alive for the target word to benefit from. Though this account would be semantic in nature, separating this from a non-semantic strategic pre-activation explanation would be difficult. Therefore eliminating picture saliency as a factor would add weight to a semantic explanation of significant priming results.

4.4.2.2 Possible effects of differences in lag length

The lag, in number of items, in Experiment 2 was purposefully shortened from that in Experiment 1 in order to maximise the opportunity to find a significant priming effect. The crucial difference when considering the effects of lag in these long SOA experiments might be the number of filler items, not the length of time (Section 2.4.4). In distributed representation models of semantic memory, priming can last over a lag (Becker et al., 1997; Hinton & Shallice, 1991; Joordens & Becker, 1997; Joordens & Besner, 1992; Masson, 1991, 1995; Plaut, 1995), but may be dependent on filler item(s) not replacing the pattern of semantic activation of the related prime (Masson, 1991, 1995; Plaut, 1995). Two filler items in the design of Experiment 1 may have replaced the prime's semantic activation, for some prime-target pairs, resulting in only a weak priming effect. In Experiment 2, when there was only one filler, this filler item, according to some distributed representation models, may not have replaced the primes' activation as in Experiment 1. However, as Vitkovitch et al. (2006) demonstrate that it is possible to find clear priming effects over two intervening items, differences in priming effects between Experiments 1 and 2 cannot be due to number of items alone. It is possible that differences in prime-target pair semantic strength are responsible as well. Semantic priming can occur over many filler items, if the semantic relationship strong, such as with the current low frequency pairs (Becker et al., 1997; Joordens & Becker, 1997; McRae & Boisvert, 1998). If the prime and target relationship is weaker, then multiple fillers may be more likely to replace the all of the activation, or learning, that occurs when the prime is named, thereby leaving little activation for the target.

4.4.3 A subsequent study

There has been some indication that priming of word targets over two intervening items can be found with a task of target word reading (Vitkovitch et al., 2006), though Vitkovitch et al. did not systematically vary target word type . The significant priming results of Experiment 2 with one intervening filler item and word reading are clear and present for all word types, yet priming is suggestively found for low frequency exception words only in Experiment 1 with two intervening filler items. A major difference is the shorter lag design created by the removal of one filler item from Experiment 1 to 2. This could create an environment in which picture primes are salient. It could also promote non-semantic strategic processes, though there is no clear evidence for this. It is also possible that the shorter lag creates an environment in which prime semantic activation can survive. In the next experiment, presented in the subsequent chapter, another filler item will be introduced between prime and target. Of interest is whether significant priming of all word types, including low frequency regular and high frequency word types, can be found over two intervening items with the current stimulus pairs, extending the effect of the current Experiment 2 and making it possible to replicate the implicated priming effects of Experiment 1.

Chapter 5

Semantic Priming Experiment 3

5.1 Introduction

The next semantic priming experiment of this thesis, Experiment 3, is described in this chapter. Results from Experiment 1 (Chapter 3) might indicate priming of low frequency exception word targets only. Experiment 1's original analysis was less than clear; the original analysis of medians approached significance for a priming effect in low frequency exception word condition, similarly participant means showed a pattern for a small priming effect, though this was not significant. Additional results confirmed variance in Experiment 1's data, and in the reanalysis of Experiment 1 data, using the set of low frequency words carried forward to other experiments, a clear priming effect was found with low frequency exception words (Section 4.4.1). Experiment 2 (Chapter 5), which reduced the number of filler items to one intervening word, produced a reliable, significant main effect of priming across four target word types, including low frequency exception words, but also with high frequency regular, high frequency exception, and low frequency regular words. Therefore, the results of Experiments 1 and 2 differed for low frequency regular words.

Experiment 3 (this chapter) investigates the clearest effect, those of Experiment 2, further, extending the design with the addition of a filler word and a manipulation of filler category membership, as will be described. This allows for the possible replication of the low frequency exception word priming effect from Experiment 1, the exploration of the differing low frequency regular word priming results between Experiments 1 and 2, and whether priming of low frequency regular target words, and, in particular, of high frequency target words can be found with two intervening filler items. Furthermore, these results will ultimately provide comment on the central aim of this thesis, i.e., whether semantic information contributes to orthography-to-phonology computation. Investigating whether semantic priming can be captured with the same set of prime-targets pairs over two intervening filler items is the next logical investigation.

5.1.1 Design consistency with Experiment 2

The design of Experiment 3, as far as was possible, was consistent with the design of Experiment 2. An intervening filler item between the prime and target will be added, but this filler item will be of word modality in keeping with the word only filler of Experiment 2. By keeping elements such as the presentation time of stimuli, target word types, and blocking consistent between Experiment 2 and 3, it may make interpreting the results of Experiment 3 easier. Therefore if the results of Experiment 3 differ from Experiment 2, then they may not differ because of these design elements as they were kept constant.

5.1.2 Filler design in Experiment 3

Experiment 3 included an additional intervening word filler item, as compared to the design of Experiment 2, making the stimulus order for Experiment 3 picture prime, word filler, word filler, word target (PWWW). Therefore the only picture was the prime item. As both filler items appeared in word modality this created a similar design to Experiment 2 in which the picture prime was salient. This offered the opportunity to investigate possible picture prime saliency that might have occurred with Experiment 2.

A manipulation of filler category membership was also included in the design of Experiment 3 - that is either two miscellaneous fillers separated prime and target or a filler from a definitive category, followed by a miscellaneous filler, separated prime and target-. This allowed for a preliminary and additional analysis into the nature of filler items, which, according to distributed representation models' account of priming, may be relevant to finding priming over a long lag (Masson, 1995; Plaut, 1995). Within these models, for target priming to occur with two intervening filler items, the filler items' semantic representation must not fully replace the prime's activated semantic representation (Section 2.4.4). It is possible that either a filler from a definitive category or a miscellaneous filler may be more likely interrupt priming that the other filler type.

5.1.3 Predictions

Considering the results from Experiment 1's original analysis and reanalysis and the results from Experiment 2, low frequency exception word priming might be expected in Experiment 3. If, in Experiment 3, only low frequency exception words are primed, as in Experiment 1, and there is no priming of high frequency target words, then the priming in Experiment 2 may not be dependent on picture saliency, as Experiment 3 recreated the modality and picture salient context of Experiment 2 by having filler items in word modality only. However, it is possible that the results of Experiments 1 and 3 might differ from the results of Experiment 2 because of the number of intervening filler items or amount of time between prime and target.

Whether significant priming of low frequency regular word targets and high frequency word targets (regular and exception) will be found is of interest. It is possible that the results from Experiment 3 will replicate those from Experiment 2, i.e., there will be a main effect of priming with all word types. If the results of Experiment 3 replicate those of Experiment 2, then this might implicate prime picture saliency and the stimuli modality context, as in Experiment 3 the picture was again made salient, as it was in Experiment 2. Therefore priming could be due to "over-processing" the picture, which could be semantic in nature, but this is difficult to untangle from possible strategic pre-activation of the target's orthography and phonology. However, the replication of Experiment 2's results would not eliminate the possibility of priming due to shared semantic activation between the prime and target. The "tools" will be used to assess whether strategic pre-activation might be a cause of any significant priming (Section 4.1.2).

Experiment 3 also includes a manipulation of first-filler category membership, and whether this has a significant effect will be monitored. As far as is known this specific manipulation is novel to the experiments of this thesis, and it is not entirely clear how the manipulation may affect the priming results. As filler items are unrelated to primes and targets, but still "picture-able" they all may have a semantic richness, e.g. a number of semantic features that could be listed, unlike abstract words, for example. It is possible that fillers from definitive categories, though unrelated, may have more semantic features in common with the primes and targets, which are also from definitive categories. Therefore, fillers from definitive categories may be less likely to interrupt the activation of the prime's semantic information. Miscellaneous fillers may be more likely to interrupt activated semantic information as they may share less semantic information with the primes and targets as their definitive category counterparts. When a miscellaneous filler are named, they may move semantic activation to other representations. Therefore, priming effect might be smaller with miscellaneous filler items. It is expected that some priming effects might still be found with miscellaneous fillers as this is the filler types that has been used previously when significant priming was found. Any filler effects will be interpreted in consideration of semantic memory models, such as distributed representations models.

5.2 Methods

5.2.1 Participants

Forty-eight volunteers (43 female) from the University of East London participated in this experiment and were given course credit for their participation. These participants did not participate in any other experiment of this thesis. Participants were aged between 18 and 47 years, with an average age of 24 years. All reported English as their first language, with six reporting fluency in a second language, and all reported normal or corrected-to-normal vision.

5.2.2 Stimuli and design

The stimuli created a picture prime-to-word target, two filler item, semantic priming paradigm similar to that of Experiment 1 of this thesis. The stimuli of the current experiment, Experiment 3, were presented in the order of prime picture, filler word, filler word, target word (PWWW). Related and unrelated prime picture-target word pairs taken from Experiment 2 of this thesis. Therefore, there was a target word manipulation of word type. Targets were one of the four word types: low frequency regular words, low frequency exception words, high frequency regular words, or high frequency exception words. The ways in which this experiment differed from Experiment 2 are detailed in the following subsections. Elements not presented are the same as in Experiment 2. See Appendix E for Experiment 3 stimulus lists.

5.2.2.1 Intervening filler words

Two intervening filler words, presented sequentially, separated the prime picture and the target word. The filler words were the names of concrete items, such as objects and animals. Information on the filler items was provided using N-watch program (Davis, 2005). The words ranged in frequency from zero to 1,068.25 per million, with an average of 33.47 million in CELEX written frequency (Baayen et al.,1995), and zero to 1,004 per million, with an average of 35.55 per million in Kucera and Francis (1967) frequency measures. Using a strict rule-based judgement of spelling-to-sound regularity (Section 1.2.2.1), 78 of the filler words were regular in their spelling-to-sound correspondence, and 82 of the filler words were exception in their spelling-to-sound correspondence.

There were some changes to filler word selection and assignment criteria, as compared to Experiments 1 and 2, as a large number of items was now needed for the design (160 fillers in total), and a filler manipulation of category membership was being included.

The selection criteria used for filler words in Experiment 3 is detailed here. Item names were used as filler words even when there was no object picture available. This is in contrast to Experiments 1 and 2; the filler word's object picture was available, though never used. Additionally, as a manipulation of filler category membership was included in the design, names of items from a definitive category were used as filler words, even if that category was being used in the experiment already. However, a filler word from a definitive category was never used in a frequency condition where that particular category was already being used as primes or targets. (As a reminder frequency conditions were in separate halves of the experiment.) For example, unused insect names, e.g. "flea", "spider", were used as filler words in the high frequency target word condition, but not in the low frequency target words condition "fly" was a picture prime and "ant" was a target word. In previous experiments, items were not selected as fillers if they were from the same category of any prime or target item within the experiment. Fillers in other experiments of this thesis were mainly miscellaneous items.

The manipulation of filler category membership meant that in this experiment there were two filler "types". There were filler words that represented items from specific categories, as detailed above, and there were filler words that represented miscellaneous items, e.g. "robot". With this filler manipulation, primes and target were either separated by two miscellaneous fillers, or by a filler from a definitive category and a miscellaneous filler. Half of the quadruplets in each list (five) had a first filler word from a definitive category. The other half of the quadruplets from each list (five) had a first filler word that was miscellaneous. The second filler word of each prime-target pair was miscellaneous, which was pseudo-randomly assigned using the criteria of previous experiments. Therefore it was the first filler word of each list that included a manipulation of category membership while maintaining a miscellaneous filler item before the target, as in previous experiments.

5.2.2.2 Counterbalancing of materials

Materials were counterbalanced and randomised as in Experiment 2. Each stimulus list, eight in total, contained 10 quadruplets of prime picture, filler word, filler word, target word, and items were presented in this order. See Appendix D for counterbalance orders.

5.2.3 Procedure

The procedural details were the same as described in Experiment 2. The trial sequence was as follows: fixation cross (one second), stimulus (until response), blank screen for experimenter response (three seconds). There was, therefore, four seconds between each stimulus, with each prime, filler, filler, target quadruplet lasted a minimum of 17.2 seconds, though the precise time is dependent on the speed of participants' response. Each testing session lasted approximately 40 minutes.

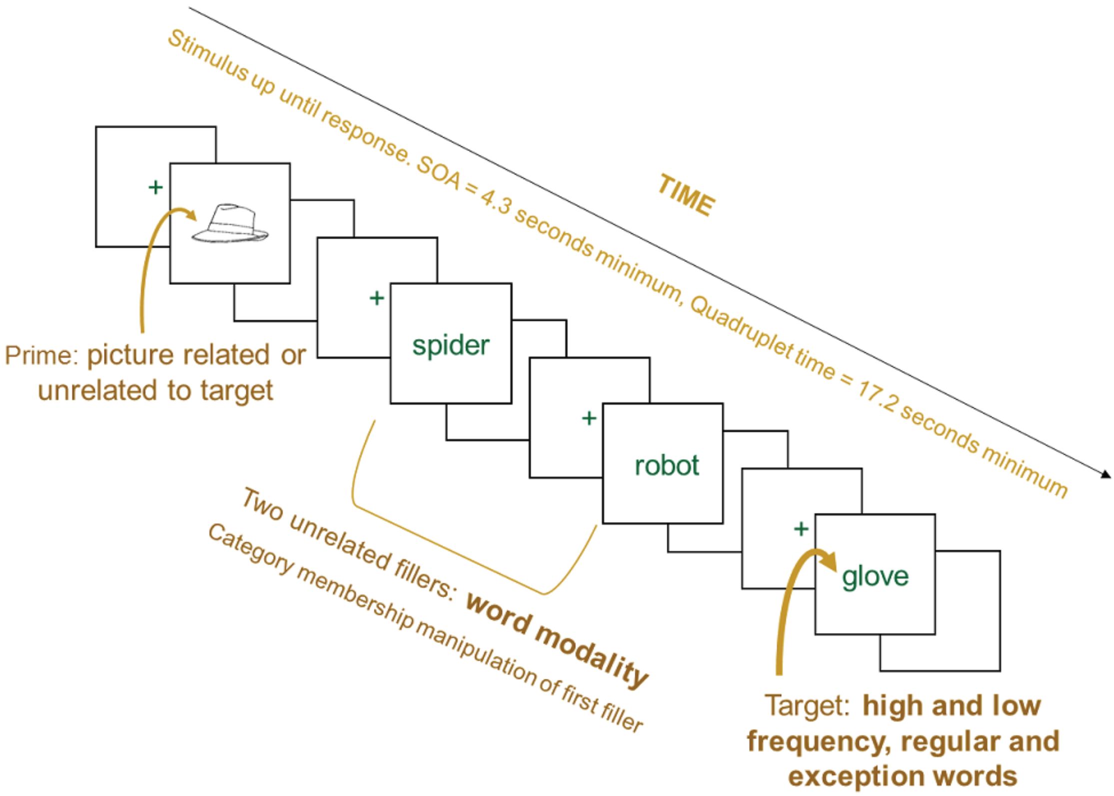

Figure 5.1. Illustrated example of the procedure for one quadruplet of stimuli from Experiment 3 (see text for details)[27].

5.3 Results

Analyses were performed in keeping with the two previous priming experiments of the thesis. Results are reported for 48 participants.

Target naming times were removed from analysis when the target was incorrectly named (4.4 % of target responses), or there was a voice key error when naming a target (1.4 % of target responses). Target naming times were also removed from the analysis when the prime was incorrectly named (9.8 % of target responses) or there was prime voice key error (2.2 % of target responses). In total 17.7 % of target responses were

[27] The image of the hat is from "A Standardised Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity," by J. G. Snodgrass and M. Vanderwart, 1980, Journal *of Experimental Psychology: Human, Learning, and Memory, 6* (2), p174-215. Copyright by Life Science Associates (LSA). Adapted with permission of the copyright holder.

removed due to incorrect names or voice key errors to primes and targets. As in previous experiments, a threshold for mean reaction times was implemented. Response times that fell outside lower- (300 ms) and upper- bound (1500 ms) thresholds were removed (less than 0.5% of possible target response times).

To create target naming error percentages per condition per participant, the total number of misread or mispronounced target words per participant per condition was divided by the total number of target words in that condition (10) and multiplied by 100. It was these percentages that were analysed in the errors analysis. In total, target name errors composed less than 5% of errors.

Table 5.1. shows means, median, and standard deviations of reading times, and mean and standard deviation error percentages for word target reading in Experiment 3. The pattern of reaction times differs from patterns in the previous two experiments. There is a pattern of inhibitory priming (slowing of response) in three of the four word types. High frequency exception, low frequency regular, and low frequency exception words show a pattern of inhibitory priming, whereas high frequency regular words show patterns of facilitatory priming.

	Target Word Condition											
	High Frequency Target Words						Low Frequency Target Words					
	Regular Target Words			Exception Target Words			Regular Target Words			Exception Target Words		
	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference
Mean RT	545.99	558.52	-12.52	574.01	569.08	+4.93	581.23	569.44	+11.79	606.36	594.09	+12.27
SD	93.76	102.47		102.28	93.92		105.91	101.95		131.42	113.69	
Median RT	538.11	546.17	-8.06	558.85	557.53	+1.32	581.24	559.64	+21.6	590.79	572.56	+18.23
SD	82.78	93.78		101.63	98.66		113.00	88.63		123.44	99.85	
Error %	0.2%	0%	+0.20%	5%	6%	-1.00%	2%	2%	0%	11%	9%	+2.00%
SD	14%	0%		7%	8%		6%	4%		9%	8%	

Table 5.1. Experiment 3 results by participant. Reaction times (RT) reported in ms, error rate (Error %) reported in percentages, and standard deviations (SD) for target words (correct-only) as a function of target word type and priming (related/unrelated) conditions for Experiment 3.

5.3.1 Reading time analyses

Within subject ANOVAs were performed; frequency, regularity, and priming conditions created the two (high, low) by two (regular, exception) by two (primed, unprimed) factors. The main effects frequency and regularity were significant by subject and item means analyses, frequency, F_p (1,47) = 20.56, MS_e = 3128.40, p < .001, F_i (1, 76) = 4.43, MS_e = 5041.00, p = .04; regularity, F_p (1,47) = 16.36, MS_e= 2867.91, p < .001, F_i (1, 76) = 7.17, MS_e = 5041.00, p = .01. High frequency target words (561.90 ms) were read more quickly than low frequency target words (587.78 ms). Target words with regular spelling-to-sound correspondences (563.80 ms) were read more quickly than exception target words (585.88 ms).

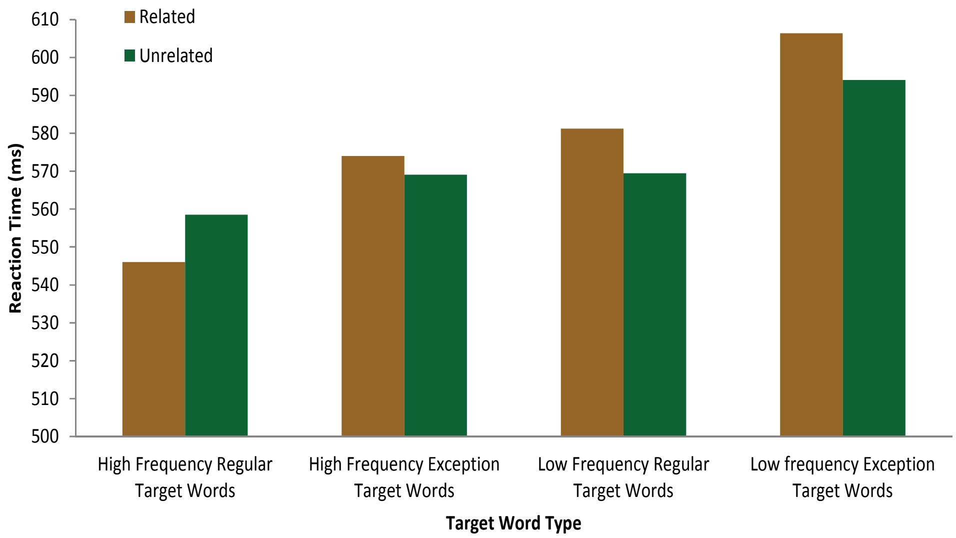

Figure 5.2. Mean participant reaction times of Experiment 3 as a function of target word type and priming (related/unrelated) conditions.

Participant means per target word type and priming condition are depicted graphically in Figure 5.2. The main effect of priming was not significant, F_p (1,47) = 0.70, MS_e = 2329.30, p = .41, F_i (1,76) = 0.04, MS_e= 1726.04, p = .84, nor were most of the interactions, frequency x regularity, F_p (1,47) = .21, MS_e = 3658.22, p = .65, F_i (1, 76) = 0.89, MS_e = 5041.00, p = .35; regularity x priming, F_p (1,47) = .68, MS_e = 2824.68, p

= .41, F_i (1, 76) = 0.82, MS_e = 1726.04, p = .37, frequency x regularity x priming, F_p (1,47) = .91, MS_e= 1899.89, p = .35, F_i (1, 76) = 0 .03, MS_e = 1726.04, p = .87. The interaction of frequency and priming with subject and item mean reaction times was not significant, F_p (1,47) = 2.63, MS_e = 2289.86, p = .112, F_i (1,76) = 0.69, MS_e = 1726.04, p =.41, nor was it significant with median subject reaction times, frequency x priming F_p (1,47)= 4.59, MS_e=2836.38, p=.04, as this is a further analysis using the same data the critical alpha was adjusted (also see Section 3.3., p 94).

As the main effect and interactions with priming were not significant in the ANOVA, planned comparisons of low frequency exception word reaction times were conducted (Section 3.3).The differences between related and unrelated conditions were not significant with subjects means, t_p(47) = 1.09, SEM = 11.28, p = .14, the median reading time analysis only appear to approach marginal significance, t_p (47) = 1.6, SEM = 11.61, p =.06. However, consistent with the simple effect analyses above, the pattern in the data is opposite to the relationship predicted, i.e. targets in the primed condition are read more slowly than targets in the unprimed condition. For the purposes of this thesis, one tailed t-tests were used to explore whether low frequency exception target words are primed. Because this is the specific difference of interest, i.e., that primed words might be read faster than unprimed target words, any indication of a difference in the opposite direction of this prediction cannot be considered as significant or nearing marginal significance (Howell, 1992).

5.3.2 Error rate analyses

Participants' mean errors to targets were subjected to the same ANOVA with frequency x regularity x priming as the two by two by two factors. There were a greater number of errors to exception target words than to regular target words, $F_p(1,47) = 124.53$, $MS_e = 0.004$, $p < 0.001$. There were more errors to low frequency than high frequency target words, $F_p(1,47) = 16.99$, $MS_e = 0.005$, $p < 0.001$. There was significant interaction of frequency and regularity, $F_p(1,47) = 6.25$, $MS_e = 0.003$, $p < .02$. Examination of the means shows that low frequency exception word targets had the highest proportion of errors, creating a difference between the low frequency regularity conditions that was not present between the high frequency regularity conditions. The main effect of priming failed to reach significance, $F_p(1,47) = .000$, $MS_e = .003$, $p = 0.99$. No other interactions with priming were significant, frequency x priming, $F_p(1,47) = 0.71$, $MS_e = 0.004$, $p = .40$, regularity x priming, $F_p(1,47) = 0.00$, $MS_e = 0.003$,b $p = 0.99$, frequency x regularity x priming, $F_p(1,47) = 1.14$, $MS_e = 0.004$, $p = .29$.

5.3.3 Filler manipulation analyses

Analyses were undertaken to explore the effect of the filler category membership manipulation. Response times to target words that received two miscellaneous fillers before them were compared to response times to target words that received one filler word from a definitive category and another filler from a miscellaneous category before them. There were a maximum of five data points for each condition. A two (regularity) by two (primed) by two (filler manipulation) ANOVA was performed using participant means for each of the frequency types, high and low.

There was no main effect of filler type in the high frequency target word condition, $F_p(1,47) = .10$, $MS_e = 3780.19$, $p = .75$. There was a main effect of regularity, $F_p(1,47) = 12.56$, $MS_e = 3899.16$, $p = .001$, but not of priming, $F_p(1,47) = 0.28$, $MS_e = 2668.80$, $p = .60$. There were also no significant interactions of filler type with the other factors, regularity x filler, F_p (1,47) = 2.25, MS_e = 3675.32, p = .14, priming x filler: F_p (1,47) = 1.09, MS_e = 3179.09, p = .30, regularity x priming x filler, F_p (1,47) = .22, MS_e = 4032.44, p = .64.

There was no main effect of filler type within the low frequency condition with means or medians, F_p (1,47) = 1.22, MS_e = 4090.76, p = .28. As with high frequency words, there was a main effect of regularity, F_p (1,47) = 7.21, MS_e= 8439.52, p = .01, but not of priming, F_p (1,47) = 1.60, MS_e = 7920.27, p = .21. The majority of the interactions of filler type and other factors did not reach significance with means or medians, regularity x filler, F_p (1,47) = .97, MS_e = 5487.31, p = .33, priming x filler, F_p (1,47) = .41, MS_e = 3650.56, p = .52. However, a trend for an interaction of filler x regularity x priming factors emerges in the analysis of mean reaction times, F_p (1,47) = 3.66, MS_e = 2837.13, p = .06. Means for this interaction are provided in Table 5.2. This did not approach significance with median reaction times, filler x regularity x priming, F_p (1,47) = 2.02, MS_e = 33367.71, p =.16.

		Filler Type					
		Specific Category			Miscellaneous		
Target Word Type		Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference
Regular							
	Mean	598.48	602.393	-3.913	612.773	587.654	+25.119
	SD	19.765	17.727		20.172	19.309	
Exception							
	Mean	577.024	557.933	+19.091	585.326	579.247	+6.079
	SD	17.954	14.816		15.342	16.816	

Table 5.2. Participant reading time means for filler type conditions as a function of target word regularity type and priming condition for Experiment 3.

As there is specific interest in any possible effects of filler type on priming, the means were examined and simple effect analyses of subject mean reaction times were performed. None of these patterns were significant (reported two-tailed, critical $p = .0125$), exception word definitive category, t_p (47) = 0.64, SEM = 8.56, p =.52, exception word miscellaneous, t_p (47) = 1.12, SEM = 11.69, p =.27, regular word definitive category, t_p (47) = .76, SEM = 7.62, p =.45, regular word miscellaneous, t_p (47) = 1.16, SEM = 9.61, p =.25.

5.4 Discussion

Experiment 3 investigated whether significant semantic priming could be found with target words over two intervening filler words, further exploring the significant main effect of priming with four word target types from Experiment 2. Experiment 3 also provided preliminary investigations into possible effects of filler type.

There were significant main effects of spelling-to-sound regularity and frequency, with low frequency and exception words having the slowest reaction times, though an interaction of the factors was only significant in error rates. It demonstrates that there was power to detect some effects in the data. Word type conditions, however, did not interact with priming.

Experiment 3 failed to find a main effect of priming with reading times or error rates. The analyses of median reaction times, but not means, revealed a significant interaction of frequency and priming. The pattern of effects, however, was unexpected and caution is needed when interpreting these. High frequency target word showed a pattern of facilitatory priming, though this was not significant, and low frequency targets showed a pattern opposite to the predicted priming pattern. Low frequency targets showed inhibition- that is, slower reaction times for primed targets-, but this also was not significant in planned comparisons or simple effect analyses. Analyses of the filler category membership manipulation also indicated some inhibition effects, and these too appeared to differ between frequency types. Though there were no significant filler effects in the high frequency word condition, in the low frequency target condition, the interaction of regularity, filler type, and priming approached significance ($p = .06$). This interaction is indicative of some influence of filler type. Low frequency regular target words were slowed (inhibitory priming) in the miscellaneous filler conditions, but not in the definitive category condition where there was a small non-significant facilitatory priming pattern. Low frequency exception words were slowed regardless of filler type. However, this pattern of filler effects was not significant in simple effect analyses. Overall, the inhibitory priming of low frequency words, was not part of any predicted pattern, e.g., priming of low frequency words, but not high frequency words, or the priming of all word types. Furthermore, caution is needed when considering the follow-up analyses and marginal results. Significant effects of the filler category membership manipulation warrants future investigations into filler qualities, such as category membership and modality, and their effect on capturing significant priming with word targets. Interpretation of interference effects will be readdressed in the priming discussion chapter following further investigations (Chapter 7).

Post-experiment questionnaire responses have the potential to provide some insight into the unanticipated pattern of results in Experiment 3; though the pattern of response for this experiment was similar to previous experiments. When questioned about related items in the experiment, 10 participants noticed categories of items, (22 participants reported incorrect relationships, e.g. associates or identical pictures and words). Eleven reported noticing pattern, but this was a pattern of modality, and identification was not correct, with six reporting an order of WWWP (the order was, in fact, PWWW). Only one participant reported predicting items. Therefore, as far as the participants could report, their knowledge of the experiment's design, the quadruplets, or primes and targets was no greater that the participant reports from the previous experiments.

5.4.1 Implications for subsequent studies

Experiment 3 failed to replicate the main effect of priming found in Experiment 2, though there was some indication of differing results for the two frequency types with a significant interaction in median reaction times. The design of Experiment 3 created a long lag priming design, in which the prime picture was potentially salient, as it was Experiment 2, with an additional word filler item. Had the main effect of priming from Experiment 2 been apparent here, then it is possible that strategic pre-activation of the target word's orthography and phonology could not have been eliminated as explanation for significant priming effects; automatic pre-activation in the paradigm is highly unlikely due to the long SOA between prime and target (Sections 2.4.1. and 2.4.4).

Experiment 3 also failed to replicate the low frequency exception target word priming effect of Experiment 1, as the planned comparison was not significant. If the result of Experiment 3 replicated the result of Experiment 1, then this may have indicated that lag length, both in number of items and time, was crucial when investigating word target priming, as these results from a two lag experiment would differ from a one lag experiment. A replication could offer firm evidence that the number of items (and amount of time) between prime and target might be responsible for whether or not significant priming is found, as the filler items might replace the remaining semantic activation from the prime. There is, however, as indication of filler effects. In low frequency targets, the filler manipulation appeared to have some effect. In fact, the results indicate that there is an unanticipated interference effect for low frequency words, and low frequency exception words in particular. This result, as it is unanticipated and opposite to previous results, is difficult to understand. Inhibitory effects and priming effects may occur in different systems, inhibition in semantic-to-phonology connections and priming in semantic memory (Section 2.4.5.2; Damian & Als, 2005; Howard et al., 2006; Oppenheim et al., 2010) and perhaps this result reflects this. It possible that picture primes create changes in both areas, i.e., semantic memory and the connections between semantics and phonology (Section 2.4.5.2). Within Experiment 3, only the activation in the connections between semantic and phonology survived, resulting in inhibition, whereas activation within semantic memory itself did not survive in this experiment with only word fillers, unlike previous priming experiments. Instead of over interpreting, there will be a further investigation, which may provide the opportunity to replicate these results, revealing additional information. However, there is some indication that what occurs between prime and target matters when investigating long term semantic effects. These ideas are discussed further in Chapter 7 once further investigations have taken place.

Of interest is why the pattern for low frequency target words has changed from Experiment 1 to Experiment 3, even though the lag was similar in number of items and length. Additionally, the results of the filler manipulation of Experiment 3 suggests that filler type (in the sense of category membership versus miscellaneous) does not affect both low frequency word regularity types, as low frequency regular words showed trends for an effect but low frequency exception words did not. A crucial factor when accounting for the difference in result between Experiments 1 and 3 could be the modality of filler items, as Experiment 1 used two filler items that were either in picture or word modality, and Experiment 3 used two filler words. Additional analysis of switching costs (Section 4.1.1.3) may suggest this too. Therefore, it could be the extra filler word that is creating the difference in results. It is therefore relevant for the next experiment to return to the filler modalities of Experiment 1 (PFFW), i.e., including picture filler items, to investigate whether facilitatory priming is again found for low frequency words, and low frequency exception words particularly.

In the high frequency conditions of Experiment 3 there was some suggestion of facilitatory priming in the means (Table 5.1), but the pattern was not significant in simple effects analyses. Additionally there was no indication that the manipulation of filler types affected the results of high frequency target words. It is, therefore, still unclear whether priming occurs with high frequency word targets over two filler items. As Experiment 4 will use two intervening filler items again, though with the inclusion of pictures fillers, priming of high frequency words over two items can again be investigated.

5.4.2 A subsequent study

It is relevant to now investigate whether facilitatory priming returns for low frequency exception words when the filler modality returns to that of Experiment 1; also of interest is whether facilitatory priming can be observed for high frequency target words, as this word type was not originally included in Experiment 1. Therefore the next experiment of this thesis, Experiment 4, includes filler items in picture modality, while keeping all other features of Experiment 3 constant, including the filler category membership manipulation.

Chapter 6

Semantic Priming Experiment 4

6.1 Introduction

Experiments 1, 2, and 3 of this thesis have produced inconsistent results, and final conclusions about a semantic contribution to word reading have been difficult to make. Results from the original analysis and the re-analysis of Experiment 1 data suggest that semantic priming occurs over two filler items with low frequency exception words, but not with low frequency regular words (high frequency words were not included in this experiment). In Experiment 2 a robust main effect of priming was found, with one less filler item than Experiment 1, using four word target types ranging from difficult (high frequency regular) to easy (low frequency exception) and the two word types in between (high frequency exception and low frequency regular). Experiment 3 investigated whether a main effect of priming could be found with an additional intervening filler word between prime and target, returning to two filler items, both in word modality. It also had preliminary investigation into whether filler type, in the form of category membership, might affect the priming results.

Experiment 3 failed to replicate the results of either of the previous priming experiments of this thesis. There was an indication of differing priming effects for the two target frequency types, but the pattern of results was unexpected. There were non-significant trends for facilitatory priming of high frequency target words types and for inhibitory priming of low frequency target word types. The previous marginally significant and significant patterns of low frequency target word facilitatory priming in Experiments 1 and 2, respectively, was now reversed; in Experiment 3 responses to low frequency target words were slowed in the primed condition, though this was not significant. Additionally, low frequency targets showed indications of filler category membership effects, whereas high frequency targets did not.

Considering the results of the various experiments, along with details of their designs, can provide more information. In a distributed model of semantic memory, priming of a target item over intervening items can be dependent on whether the activation of the prime representation is replaced by the activation of the representations of the target (Masson, 1995; Plaut, 1995). However, if only the number of filler items was important, then experiments with the same number of items (Experiments 1 and 3) should produce similar priming results. However, as the results from Experiments 1 and 3, were not similar, another design difference must be responsible for the inconsistent results.

The design of Experiment 1 differed from the design of Experiments 2 and 3 in the filler modality. In Experiment 1, three-quarters of stimuli quadruplets included a filler in the picture modality. In Experiments 2 and 3, no triplets or quadruplets, respectively, contained fillers in the picture modality. As Experiment 3 failed to replicate the (marginal) results of Experiment 1, an experiment with a similar intervening lag in number of items and amount of time, but differing filler modalities, there is an indication that filler modality may have an effect on priming, in addition to filler category membership effects. Discovering whether priming is dependent on the presence of picture fillers might help in understanding the effects found in Experiments 1, 2, and 3. It is possible that including a picture filler, as in Experiment 1, may see the return of facilitatory priming of low frequency target word types, and this is of interest in Experiment 4.

There are various ways in which a picture filler could affect a semantic priming result. As noted in Section 2.4.5.2, correct picture naming involves the activation of semantic information in order to name the specific picture (Humphreys et al., 1988; Lambon-Ralph et al., 2001; Levelt, et al., 1991; Wheeldon & Monsell, 1992, 1994). It is possible that naming a picture between a picture prime and word target affects whether semantic priming of a word target is found. Firstly, cross-modal priming data has suggested that mixing words and pictures may encourage a semantic route to word reading (Bajo, 1988). Since naming a picture involves activating phonology, processing may simply continue to use this route when reading words within the same paradigm (of course, words would use an *orthography*-to-semantics-to-phonology) (Humphreys et al., 1988; Humphreys, Lamote, & Lloyd-Jones, 1995; Lambon- Ralph et al., 2001; Levelt, et al., 1991; Wheeldon & Monsell, 1992). However, there is no evidence for this in the current data. If semantically-biased word reading was occurring, then priming should be found in all experiments of this thesis and with the four word target types, but it was not. Additionally, if pictures bias word reading to semantic memory, then priming of all of the word types might especially be expected in the experiment with picture fillers, i.e., Experiment 1, however, this was not found. Therefore a biasing of word reading via semantic memory cannot account for the results of the experiments of this thesis.

Though it is highly unlikely that priming in this thesis is non-semantic in nature (Section 4.1.2), including a picture filler might make the pattern of stimuli, the priming paradigm, and the relationship of prime and target less predictable, or less prone to guesses. Therefore, the inclusion of picture fillers can be used as an additional a design aspect in the “tools” making strategic prediction even less likely than in the previous (Experiment 2 and 3) designs.

Finally, picture and word filler modalities might make the experiment more difficult, as hinted at by the task switching analysis of Chapter 4 (Section 4.1.1.3). In making target word reading more difficult, it may provide time for semantic information to contribute, as the triangle model suggests semantic information has a chance to contribute when word reading is less efficient (Plaut et al. 1996) (Section 1.3.3). However, faster mean reaction times have not eliminated priming. Experiment 2 had, on average, faster reaction times than Experiment 1, but priming was still found. Regardless of the specific reason, it is possible that picture fillers could affect the priming of target words.

6.1.1 Aims

The results from the previous experiments encouraged a further semantic priming experiment that reinstated a picture filler item between prime and target, while keeping other design elements consistent with Experiment 3. Reinstating a picture filler allowed for investigation into whether modality of filler item affects priming results, as suggested by the opposing non-significant data patterns of Experiments 1 and 3. This also provides the opportunity to directly replicate the nearly significant results of Experiment 1 of this thesis with the shortened stimulus lists. As a reminder the reanalysis of Experiment 1 data, though marginally significant, confirmed a pattern of facilitatory priming of low frequency exception words in the data. This, however, was an analysis of a subset of data and only marginally significant; so, replication is needed.

Therefore, Experiment 4 returned to a design similar in modality to Experiment 1 of this thesis and Experiment 2 of Vitkovitch et al. (2006). Using the four word types as targets also provides the opportunity to further investigate a semantic contribution to reading the various types. Of interest is whether facilitatory priming of low frequency exception target words might return with the return of a filler picture, and whether high frequency priming can be found over two intervening items. It extends previous picture-prime-to-word-target two filler designs by including a manipulation of filler category membership as in Experiment 3 providing a chance to replicate the results from the previous experiment.

6.1.2 Predictions

As differences are minimal[28] between the designs of Experiment 4 and the reanalysis of Experiment 1 that used a shortened stimulus list, it is predicted that priming will be found in the low frequency exception target word condition. This result would indicate that the modality of the filler items effects priming, as they differ from Experiment 3's results and Experiments 3 and 4 only differed in the modality of filler items.

It is still unclear whether significant priming can be found for high frequency word types over two intervening filler items. High frequency words were not included in Experiment 1, and though there were patterns of high frequency word priming in Experiment 3, the differences were not significant to the critical $p = .05$ level. Therefore Experiment 4 continues to investigate possible high frequency word target priming.

[28] The precise differences are one second less of SOA per item, the precise filler items, and a slightly longer stimulus list in Experiment 1. The proportion of words and pictures is the same in Experiments 1 and 4.

It is possible that high frequency priming might occur. If the results of the current experiment are similar to Experiment 3, with trends for facilitatory priming of high frequency word conditions and for inhibition in low frequency target word conditions, then this may indicate that modality of filler item is not important, as the designs of Experiments 3 and 4 were identical with exception of the modality of filler items, and their results would differ from Experiment 1. However, this type of result may indicate that the specific filler item used is important, as Experiments 3 and 4 have identical filler items, though not identical filler modalities, and would also have identical results. Yet, these results would be in contrast to Experiment 1 which had different specific filler items, but the same modality of fillers as Experiment 4.

Filler category membership also remains under investigation in Experiment 4. There was a suggestion of filler category membership effects in the results of Experiment 3, for the low frequency word condition, which had a different pattern of priming results than expected (inhibitive priming). The results, however, were not significant to the critical p-value levels. Of interest is whether filler category membership effects will be present when the modality of filler items includes pictures and whether the effects will be similar to those of Experiment 3.

6.2 Methods

6.2.1 Participants

Forty-eight volunteers (37 female) from the University of East London participated in this experiment and were given course credit for their participation. These participants did not take part in any other experiment of this thesis. Participants were aged between

19 and 48 years, with an average age of 28 years. All reported English as their first language, with seven reporting fluency in a second language. All participants reported normal or corrected-to-normal vision.

6.2.2 Stimuli and design

Stimuli and design details, including prime and target information, are identical to Experiment 3, except those detailed below (Sections 4.2.2 and 5.2.2). A picture prime-to-word target, two filler item, semantic priming paradigm was used in this experiment. This design was similar to that used in Experiments 1 and 3 of this thesis. The stimuli of Experiment 4 were presented in the order of: prime picture, filler item, filler item, target word (PFFW). Related and unrelated prime picture - target word pairs were the same as those used in Experiments 2 and 3 of this thesis. As before, target words were one of the four word types: high frequency regular, high frequency exception, low frequency regular, or low frequency exception. Stimulus lists for Experiment 4 can be found in Appendix F.

6.2.2.1 Intervening filler items

Prime-filler-filler-target assignments were brought forward from Experiment 3 to Experiment 4 as far as was possible. As in Experiments 1, 2, and 3, the filler items, here, two, were linked with the target word whether it appeared in the primed or unprimed prime-picture condition. As in Experiment 3, Experiment 4 also included the manipulation of filler category membership of the first filler item; the filler item was either from a specific category or was miscellaneous. The second filler item was always miscellaneous. Specific filler item assignments from Experiment 3 to Experiment 4 were only changed if there was not a picture available for the assigned filler item. See paragraph below for modality details. In total only eight of the 160 filler items were changed. One was changed in the high frequency exception condition, two were changed in the high frequency regular condition, one was changed in the low frequency exception condition, and four were changed in low frequency regular condition, though one of these moved from second position to first filler position and remained with the same target.

The parameters of filler modality were similar to those of Experiment 1. Half of the 160 filler items were of word modality and the other half were of picture modality. Filler pairs were assigned a modality pattern to create the final stimuli quadruplets. Filler item pairs were in one of four modality orders: word-word (two per list), word-picture (two per list), picture-picture (three per list), or picture-word (three per list). Therefore, half of the target words per list were preceded by a picture filler and the other half of the target words were preceded by a word filler. Filler pictures were taken from the IPNP (Szekely et al., 2004).

6.2.2.2 Counterbalancing and randomisation

Counter balancing and randomisation were as in Experiment 3 of this thesis (Section 5.2.2.2). See the Appendix D for the counterbalance table.

6.2.3 Procedure

As in previous experiments of this thesis, participants named a set of practice trials before beginning the test items. Practice trials consisted of four sets of four items in the order of: picture prime, filler item, filler item, and target word. Filler item modality replicated that of the experiment. These items were unique to the practice. All other procedural details were identical to those of Experiment 3 (Section 5.2.3.).

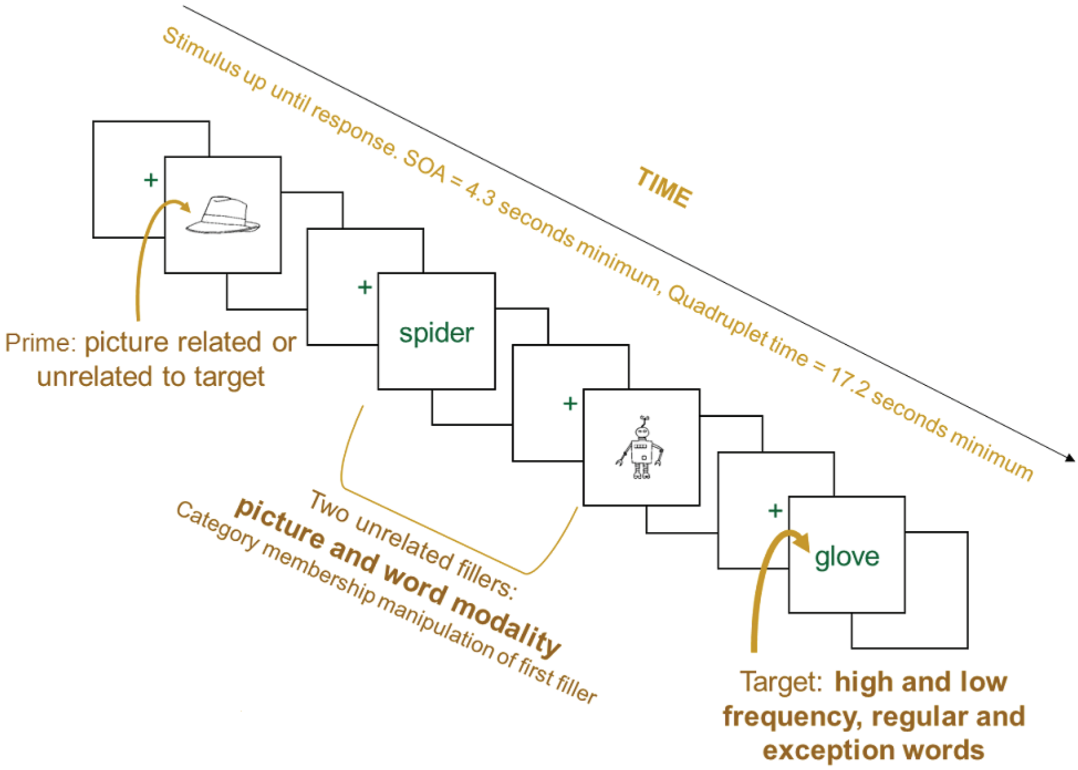

Figure 6.1. Illustrated example of the procedure for one quadruplet of stimuli from Experiment 4 (see text for details)[29].

6.3 Results

Analyses were performed in keeping with the three previous priming experiments of the thesis. Results are reported for 48 participants. Reaction times were excluded using the same criteria as the previous experiments in this thesis. Target naming times were excluded from analysis when the target word was read incorrectly (2.9 % of target

[29] The image of the hat is from "A Standardised Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity," by J. G. Snodgrass and M. Vanderwart, 1980, Journal *of Experimental Psychology: Human, Learning, and Memory, 6* (2), p174-215. Copyright by Life Science Associates (LSA). Adapted with permission of the copyright holder.

responses), or there was a voice key error when reading the target word (1.2 % of target responses). Target naming times were also removed from the analysis when the prime picture was named incorrectly (5.9% of target responses), or there was voice key error when naming the prime (1.5 % of target responses). Response times that were lower than 300ms or higher than 1500ms were also removed (0.3% and 1.6% of responses, respectively). Box plots of participant mean naming times did not reveal any outliers.

The number of misread or mispronounced target words per condition was taken as a percentage of the total number of target words in that condition (10). It was these percentages that were analysed. Target name errors (the misreading of a target word) composed less than 3% of total errors.

Table 6.1. shows means, median, and standard deviations of reading times, and mean and standard deviation error percentages for word target reading in Experiment 4. The pattern of reaction times is now opposite to that of Experiment 3. There is a pattern of facilitatory priming of low frequency targets similar to effects in the first two experiments of this thesis, whereas high frequency regular words show patterns of inhibitory priming. Patterns with high frequency exception target words depend on which the central tendency used; in means these words show a small pattern facilitatory priming, in medians they show a small pattern of inhibition.

	Target Word Condition											
	High Frequency Target Words						Low Frequency Target Words					
	Regular Target Words			Exception Target Words			Regular Target Words			Exception Target Words		
	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference	Related	Unrelated	Primed Difference
Mean RT	562.30	550.43	+11.87	597.76	603.56	-5.8	583.52	598.47	-14.95	603.82	613.64	-9.82
SD	88.27	81.78		113.92	121.35		102.43	106.30		104.48	113.94	
Median RT	555.03	550.84	+4.19	586.45	581.66	+4.79	577.20	586.51	-9.31	589.85	591.66	-1.81
SD	83.60	78.91		107.13	113.25		99.29	102.01		101.18	114.12	
Error %	<1%	<1%	0.0%	5%	4%	+1.0%	<1%	1%	0.0%	5%	7%	-2.0%
SD	1%	1%		5%	8%		3%	3%		8%	8%	

Table 6.1. Experiment 4 results by participant. Reaction times (RT) reported in ms, error rate (Error %) reported in percentages, and standard deviations (SD) for target words (correct-only) as a function of target word type and priming (related/unrelated) conditions for Experiment 4.

6.3.1 Reading time analyses

A within-subject ANOVA was performed with frequency, regularity, and priming conditions creating the two (high frequency, low frequency) by two (regular, exception) by two (primed, unprimed) factors. The two main effects of word type were significant in subject and item analyses, frequency: F_p (1,47) = 9.92, MS_e = 4410.75, p = .003, F_i (1,76) = 3.14, MS_e = 5408.37, p = .08, regularity F_p (1,47) = 24.36, MS_e = 3791.35, p < .001, F_i (1,76) = 7.29, MS_e = 5408.37, p = .009. High frequency target words (578.51ms) were read more quickly than low frequency target words (599.86ms). Target word with regular spelling-to-sound correspondences (573.68ms) were read more quickly than exception target words (604.70ms). Frequency and regularity significantly interacted in the subject mean analysis, F_p (1,47) = 6.58, MS_e = 2575.53, p = .01, though not by items, F_i (1,76) = 1.31, MS_e = 5408.37, p = .26. Observations of participant reaction time means show that high frequency regular words are read the quickest, and low frequency exception words were read the slowest.

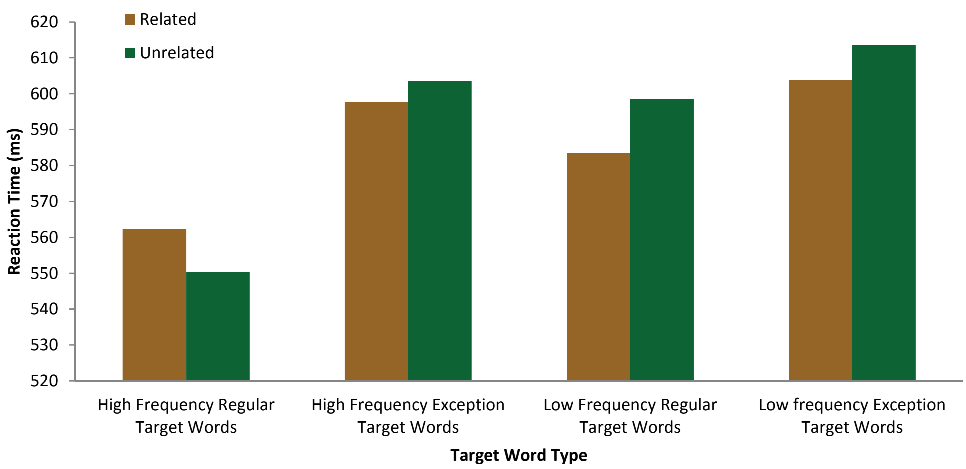

Figure 6.2. Mean participant reaction times of Experiment 4 as a function of target word type and priming (related/unrelated) conditions.

The mean target word reaction times for each word type in primed and unprimed conditions for Experiment 4 are dipicted graphically in Figure 6.2. There were no significant priming effects, though there is a trend for primed conditions to be faster in low frequency exception, low frequency regular, and high frequency exception target word conditions, but not for high frequency regular words. High frequency regular words show an opposite trend with primed target words being read more slowly than unprimed target words.

The main effect of priming was not significant, F_p (1,47) = 1.28, MS_e = 1641.63, p = .26, F_i (1, 76) = .705, MS_e = 1061.60, p = .40. No interactions with priming were significant, frequency x priming F_p (1,47) = 2.62, MS_e = 2179.17, p = .11, F_i (1,76)= 2.51, MS_e = 1061.60, p=.12, regularity x priming F_p (1,47) = 0.48, MS_e=1978.82, p =.49, F_i (1,76) = 0.58, MS_e = 1061.60, p = .45, frequency x regularity x priming F_p (1,47) = 1.78, MS_e = 1750.60, p = .19, F_i (1,76) = 0.94, MS_e = 1061.60, p = .34. See Figure 6.1.

Planned comparisons of low frequency exception target reading times in related and unrelated conditions were performed, yet no significant difference was found, t_p(47) = 0.55, SEM = 8.69, p = .19. The priming difference in mean reaction times was nearly 10ms.

6.3.2 Error rate analyses

Participants' mean error percentage to targets were also analysed in an ANOVA with frequency x regularity x priming as the two by two by two factors. Error values as function of word type and priming conditions are provided in Table 6.1.

There were more errors to exception words than to regular words, $F(1,47) = 43.02$, $MS_e = 0.005$, $p < .001$, and there were significantly more errors to low frequency than high frequency words, $F(1,47) = 5.40$, $MS_e = 0.002$, $p = .03$. The main effect of priming was not significant, $F(1,47) = 0.36$, $MS_e = 0.002$, $p = .55$, and the interactions also were not significant, frequency x regularity $F(1,47) = 0.38$, $MS_e = 0.002$, $p = .54$, frequency x priming $F(1,47) = 1.59$, $MS_e = 0.002$, $p = .21$, regularity x priming $F(1,47) = .12$, $MS_e=.002$, $p =.74$, frequency x regularity x priming $F(1,47) = .87$, $MS_e = 0.002$, $p = .36$. However, of note is that there was a pattern of priming in the error percentage with low frequency exception words with the primed condition having less errors than low frequency exception target words in the unprimed condition.

6.3.3 Filler manipulation analyses

As in Experiment 3, analyses were undertaken to explore the effect of the manipulation of filler category membership. As a reminder, the first filler of each quadruplet was either the name of an item from a specific category or the name of a miscellaneous item. A two (regular/exception) by two (primed/unprimed) by two (specific category filler/miscellaneous filler) was conducted for each frequency type.

There was no main effect of filler type within the high frequency condition, F_p $(1,47) = 0.97$, $MS_e = 3460.76$, $p = .33$. As in the main analysis of reaction times, with high frequency targets alone there was a significant main effect of regularity, F_p $(1,47) = 36.68$, $MS_e = 5362.93$, $p < .001$, but not of priming, $F_p(1,47)= 0.34$, $MS_e = 42.19.78$, $p = .56$. Filler type did not significantly interact with priming, filler x priming: F_p $(1,47) = 0.25$, $MS_e = 4725.36$, $p = .62$, regularity x priming x filler: F_p $(1,47) = 2.79$, $MS_e = 4074.44$, $p = .10$.

There was no main effect of filler type within the low frequency condition, F_p (1,47) = 1.68, MS_e = 3205.06, p = .20. When low frequency targets are analysed alone in the filler analysis, means reveal a significant main effect of regularity, F_p (1,47) = 3.99, MS_e = 7056.08, p = .05. The main effect of priming trended towards significance, F_p (1,47) = 3.09, MS_e= 4069.53, p =.09, confirming a pattern in the means. Filler type did not significantly interact with priming, filler type x priming F_p(1,47) = 2.53, MS_e = 3427.55, p = .19, regularity x priming x filler F_p (1,47) = 0.26, MS_e = 3488.06, p = .61.

6.4 Discussion

Experiment 4 (PFFW) investigated whether semantic priming of a word target is possible over two intervening filler items when some filler items were in picture modality, extending Experiment 3's design (PWWW) to include picture fillers. Whether filler modality affects semantic priming from a picture prime to a word target is of interest. Specifically, whether returning picture fillers to the design also returns the priming of low frequency target words is of interest, as well as whether priming of high frequency target words is possible in a design with two filler items. Experiment 4 also further investigated possible effects of filler category membership on target priming.

Results of Experiment 4 revealed significant word type factors with a main effect of regularity and of frequency and a significant interaction of these two factors. Low frequency exception words were the slowest, and had the highest error rate, demonstrating that these words were the most difficult to read. This word type also showed a priming pattern in error rate. High frequency regular words were read more quickly than the other three word types and they also had the lowest error percentage and the smallest standard deviations, demonstrating that these words were the easiest of the four to read. However, there was also some indication that the pattern of high frequency target word reading was different in this experiment than in the previous experiments.

There are no significant priming effects in Experiment 4, though reaction time means show a trend for priming in both low frequency word regularity type conditions (regular and exception). There was no clear pattern in the reaction times for high frequency targets, as high frequency exception words seemed to show non-significant facilitatory priming effects (with means only), whereas high frequency regular words seemed to show non-significant inhibitive priming effects; this is opposite to the pattern of priming effects in the other three word types. Priming was also not affected by the filler category membership manipulation. Therefore the (non-significant) priming patterns are more similar to Experiments 1 and 2 than Experiment 3.

Results from Experiment 4 (PFFW) did not replicate those of Experiment 3 (PWWW), which used the same stimulus lists, had the same number of fillers, and the same amount of time between prime and target, but only used word modality. Experiment 4's (PFFW) results showed non-significant priming patterns that were opposite to the data patterns of Experiment 3 (PWWW) in which non-significant patterns of high frequency word target facilitatory priming and of low frequency word target inhibition were found. Also, results of Experiment 3 (PWWW) suggested that category membership of filler items might affect low frequency target word priming, but similar indications were not found in the results of Experiment 4 (PFFW). The differing results between Experiments 3 (PWWW) and 4 (PFFW) suggest that filler modality is important when investigating the semantic priming of target words, as it was the only difference between these two experiments. Therefore whether a picture filler is presented between primes and targets may affect priming results.

Though filler modality seems to play a role in whether semantic priming is found, it is possible that this may not be the only filler aspect that affects priming results. As the results of Experiment 4 (PFFW) were not significant and did not clearly replicate the suggestive results of Experiment 1 (PFFW), this could indicate that the specific filler items used could also affect whether significant priming is found, as these items differed between the two experiments.

Indications from Experiment 4 are that filler modality and specific filler item affect whether priming is found. Therefore aspects of filler items affect whether semantic activation from a prime is still available for a target word. An effect of fillers on priming could be accommodated within some distributed representation models of semantic memory (Masson, 1995; Plaut, 1995). This possibility is explored in the next chapter, which offers additional analyses and a general discussion of these four semantic priming experiments.

The four priming experiments of this thesis produced various patterns of priming results across the target word types. Though marginally significant and significant priming occurs in Experiments 1 and 2, respectively, and this is argued as being due to shared semantic information between prime and target. To discover any further evidence of a semantic contribution to word reading in these four experiments, as marked by semantic priming, further discussion and analysis is needed. The subsequent chapter addresses these points.

Chapter 7

Further Analyses and Final Discussion of Semantic Priming Experiments 1-4

7.1 Introduction

The central aim of this thesis is to investigate whether semantic information contributes to orthography-to-phonology computation, and this was investigated in the first four experiments of this thesis using long lag semantic priming paradigms with a manipulation of target word type. In these priming experiments, if a target word was read faster after a related prime picture than when in an unrelated prime picture condition (and pre-activation arguments could be arguably eliminated), then this facilitatory priming effect could be used as evidence of a semantic contribution to word reading (Section 2.4.5.2; Damian & Als, 2005). The results of the four priming experiments of this thesis have offered some evidence of significant facilitatory priming that may be the result of shared semantic information between prime picture and word target, but the results have been inconsistent.

The four experiments of this thesis revealed some evidence of target word priming, yet none of the results were precisely replicated across experiments within chapter analyses, and some patterns of priming were not significant. The designs that were used across the four experiments varied, principally in number and type of intervening filler item. Experiment 1 (PFFW), which only used low frequency word types, showed some evidence of significant low frequency exception word priming, though this was clearer in a re-analysis with a shorter stimulus list, and high frequency target word types were not included. Experiments 2, 3, 4 used four word target types: high frequency regular, high frequency exception, low frequency regular, and low frequency exception. Experiment 2 (PWW) showed a main effect of priming with all four word types, while Experiments 3 (PWWW) and 4 (PFFW) were less clear. The former (Experiment3) showed non-significant patterns of high frequency target word priming and of low frequency inhibitory priming, while the latter (Experiment 4) showed a small non-significant trend for high frequency target word inhibitory priming and non-significant facilitatory priming of low frequency word target types. The priority now is to further examine these results through hypothesis driven additional analysis and final discussion in order to draw some conclusions as to whether there may be evidence of a semantic contribution to word reading as revealed by semantic priming.

7.2 Further analyses of the priming data

The four experiments of this thesis, thus far, have primarily been analysed on their own. However, if the experiments' data are judiciously analysed in conjunction with one another, where appropriate, then further information concerning semantic priming during word target reading may be revealed. As each priming experiment was derived from its predecessor, some experiments may naturally combine with one another to allow for limited, useful analysis. In particular, Experiments 1 and 4 could benefit from combined inspection, as they are very similar in design, and further examination of the data from the PWWW order of Experiment 1 and the data from Experiment 3 could also reveal further information.

Moreover, throughout this thesis there has been special interest in whether low frequency exception word conditions, in particular, are primed. This word type is of interest because previous research has demonstrated that this word type may be less efficiently processed than other word types thus allowing time for semantic information to contribute (Plaut et al. 1996; Shibahara et al., 2003; Strain et al. 1995, 2002; Woollams, 2005). Additionally, there is some evidence from the results of Experiments 1 and 2 that low frequency exception word priming occurs. Therefore a limited number of hypothesis driven additional analyses, which combine data from certain experiments, with a focus on low frequency exception words, are presented here.

7.2.1 Further analysis of data from Experiments 1 and 4

Experiment 4 (PFFW) and Experiment 1 (PFFW) of this thesis were the most similar in design. Experiments 1 and 4 had the same number of intervening filler items and these items were of the same picture and word modalities[30]. Prime-target pairs were also similar between the two experiments as Experiment 1 used low frequency target words, and a subset of these (20 out of 28 per word type) were used in Experiment 4's low frequency target word condition. Whether these similar experiments, both in design and prime-target stimuli, produced similar results is still not clear, especially as the results of Experiment 4 were not significant. Experiment 1's results indicated nearly significant priming of low frequency exception target words, and not of low frequency

[30] The precise filler items were different, but the modality patterns were the same. Experiment 1 and 4 were the only experiments of this thesis to use two intervening filler items of both modalities.

regular words for the subset of stimuli used in Experiment 4, while Experiment 4 results showed non-significant patterns of priming of both low frequency regular and exception target words. The data of these two studies are explored further here since there is a specific hypothesis driven interest in low frequency exception words.

One aim of Experiment 4 was to replicate the significant low frequency exception word priming of Experiment 1 (Section 6.1.1), and, as mentioned above, low frequency exception words are of special interest to this thesis. By analysing the reaction times of Experiment 4's low frequency exception word targets in conjunction with the same data from Experiment 1 power will be increased. Additionally, whether Experiments 1 and 4 show similar patterns of priming in the low frequency exception word conditions can be explored.

The design of Experiment 4 allows for low frequency target word data that was collected in a similar circumstance to Experiment 1 to be extracted and analysed. Experiment 4, which included high and low frequency word types as targets, was blocked by target word type (see Section 4.1.1.4). Therefore low frequency words appeared in a separate half from the high frequency target words. Because of the counterbalancing, low frequency words appeared in the first half of the experiment for half of the participants, making this part of Experiment 4 very similar to Experiment 1. Therefore the first analysis of this chapter was performed on low frequency exception target word data from Experiment 1 and the low frequency exception target word data from the first block of Experiment 4 that was collected in the most similar circumstance to Experiment 1.

A two factor (primed, unprimed) ANOVA was performed on the aforementioned data of Experiments 1 and 4 with a between subjects factor of Experiment. The analysis revealed a significant main effect of priming, F_p (1,69) = 464, MS_e = 2757.99, p = .04[31]. Primed low frequency exception words were read more quickly than unprimed ones in Experiment 1 and in the first half of Experiment 4. The main effect of experiment was not significance, F_p (1,69) = 0.65, MS_e = 32412.29, p = 0.42. The interaction between experiment and priming was not significant, F_p (1,69) = 0.03, MS_e = 2757.99, p = .85.

This analysis indicated that the pattern of priming effects in Experiment 4 were similar to those suggested in the marginally significant results of Experiment 1. The pattern in the original data approached significance when only low frequency target words were analysed in the filler type analysis (Section 6.3.3). Furthermore the first block of Experiment 4 may be the most similar to Experiment 1, as a high frequency block did not precede it. There may be significant priming in the first block of Experiment 4 with both low frequency word types. Therefore an analysis of Experiment 4 data from participants who received low frequency target words as the first block was performed.

A two (regular, exception) by two (primed, unprimed) ANOVA of low frequency target word participant mean reaction times from the first half of Experiment 4 was performed. The main effect of priming was significant, F_p (1,47) = 5.87, MS_e = 2428.87, p = .02. The main effect of regularity was not significant, F_p (1,47) = .39, MS_e = 4257.96, p = .54. There was no significant interaction of the two factors, F_p (1,47) = .05, MS_e = 3239.37, p = .83. Although priming was significant using low frequency reaction times from the first half of Experiment 4, an ANOVA of low frequency target word data from

[31] A similar analysis of low frequency exception word reaction times from Experiment 1 and from the first half of Experiment 4 was performed in which data was collapsed across experiment, i.e. no between subject factor was included. This revealed a similar significant priming result, F_p (1,70) = 4.97, MS_e = 2719.93, p = .03.

the second half of Experiment 4 showed no significant priming, F_p (1,23) = 0.003, MS_e = 1260.29, p = .96.

In the original analysis of Experiment 4 data in Chapter 6 (Section 6.3.1) there were non-significant patterns of priming for low frequency target words; however, in the ANOVA analysis and the planned comparison these patterns failed to reach significance at p = .05 level. The current additional analyses in this chapter revealed that the pattern of facilitatory priming in Experiment 1 and the first half of Experiment 4 for low frequency exception words is similar and significant. Further analyses of Experiment 4 data alone showed significant priming of low frequency word type conditions, both regular and exception, with low frequency words from the first half of the experiment session only, when this low frequency block was not preceded by another word type. However, in Experiment 4, a low frequency priming effect is not present in the data from the second half of the experiment when the low frequency target word block was preceded by high frequency word types in the first half. Therefore Experiment 4 shows previously unobserved block effect.

With this knowledge Experiments 2 and 3, the other experiments with high and low frequency word types counterbalanced in different halves were also analysed for this block effect. There is no indication of priming differences between experiment halves in the other experiments with a similar design. There is also no indication of priming differences between experiment halves for the high frequency target words of Experiment 4. Therefore the block effect is only limited to Experiment 4 low frequency conditions.

Because Experiments 2-4 ‘block’ the priming sequences by word target type, it could be possible that the previously reviewed effect of block order in Experiment 4 is due to a ‘block effect’, however this is unlikely, as will be described here. A block effect is, for example, when healthy participants read aloud pure unmixed blocks of easy high frequency stimuli faster than pure unmixed blocks of difficult low frequency stimuli; yet when these high frequency and low frequency stimuli are presented in one mixed block, participants show slower responses to easier high frequency words and quicker responses to more difficult low frequency words, resulting in uniform responses times for both word types (Lupker, Brown, & Colombo, 1997; Monsell, Patterson, Graham, Hughes, & Milroy, 1992).

Effects of a preceding block of stimuli on the naming times of subsequent stimulus blocks could also be considered within ‘block effects’. Therefore the unanticipated block effect in Experiment 4 would be explained in the following way. The low frequency blocks in the first half in Experiment 4 that was not preceded by any other word type would have been named at a normal, slower, speed, and the target words in the related prime condition could be speeded because reaction times for this word type were not at ceiling. However, with the low frequency blocks in the second half that were preceded by the high frequency blocks, their reading times would have been sped up, therefore no priming could occur in the related prime condition because responses were already at ceiling. Yet, a block effect explanation is not appropriate for the data of this thesis. First, the ‘blocks’ of stimuli in the priming experiment of this thesis were actually mixed blocks, though the target words were all of one type, the non-target items in the sequence (primes and fillers) could have been of any frequency or word type.

Therefore all 'blocks' of this thesis were more similar to a mixed block condition[32] . Also, no effects of block order were found in Experiment 2 or 3, which used similar (Experiment 2) if not identical (Experiment 3) stimulus items and design, specifically, block counterbalancing. Nor was an effect of block order found in Experiment 1, in which the stimuli modality was the same. Therefore, this unexpected effect is difficult to explain and cannot be accounted for by the traditional 'block effect'. This effect could instead be due to the limited size of the effect.

Keeping these effects of block in mind, data from the first half of Experiment 4 is the arguably clearest, as it may be less affected by blocking of frequency type, category repetition in the filler items, and participant fatigue. Additionally, data from the first half of an experiment may be most like Experiment 1, which included low frequency target word regularity types only. It is in the analysis of first-half low frequency target words data that priming was found in Experiment 4.

7.2.2 Further analysis of data from Experiments 1 and 3

Though there were indications of similar data patterns for low frequency exception words in Experiments 1 and 4 that have been confirmed through additional analysis, there are other indications that the data of Experiment 3 might have a dissimilar pattern of results from the other experiments, including the results of low frequency exception words. The next analysis seeks to investigate this seemingly different pattern as facilitatory priming patterns are of direct relevance to this thesis's central aim.

[32] An account which extends the block effect to individual sequential item, i.e., a preceding item effects the speed at which the subsequent item is read (Chateau & Lupker, 2003) is also inappropriate here because the between subjects design of Experiment 4 meant that precisely the same items were seen in the first and second half for different participants, and there was a priming effect in when low frequency items were in the first half.

Experiment 3 (PWWW) investigated whether priming of all word types was possible with two intervening filler items in word modality while also exploring possible effects of picture prime saliency. Results of Experiment 3 revealed some non-significant evidence of an unexpected interference effect (slowing of response in the related condition) with low frequency target words, including low frequency exception words. This was an opposite effect to the facilitatory priming pattern in reaction times in the low frequency conditions of the other priming experiments of this thesis. Experiment 3 was not the only experiment, however, to have PWWW orders in the design. A few targets in each condition of Experiment 1 appeared in PWWW modality[33], and Experiment 1 was similar in design to Experiment 3 including the same prime-target and quadruplets of stimuli. However, the results of Experiment 1 showed (marginally significant) patterns of facilitatory priming of low frequency exception target words, whereas Experiment 3 did not. Investigating the general pattern of results for targets in the PWWW order of Experiment 1 is of interest, as this perhaps could be different to the reported pattern with all quadruplet modality sequences. Additionally, discovering whether the pattern of PWWW target reaction times in Experiment 1 is similar to the results of Experiment 3 is of worth as it can provide more information about possible filler modality effects in the priming designs of this thesis. An analysis was undertaken in order to investigate the apparent differing pattern of results for low frequency exception target word between Experiments 1 and 3.

Low frequency exception target word data from Experiment 3 and from the PWWW orders of the 10-item shortened stimulus list of Experiment 1, i.e. the stimuli of the

[33] Experiments 4 also contained quadruplets in this modality order. The experiment with the most number of items in PWWW order and also with a significant priming effect prior to the additional analyses was chosen for analysis. Experiment 4 has a maximum of two PWWW quadruplets per condition, whereas Experiment 1 had a maximum of 3. Experiment 4 also had non-significant effects until this chapter's analyses. Therefore further analysis was performed with Experiment 1 data.

reanalysis (Section 4.4.1), were used in this analysis. An ANOVA using experiment number as a between subjects factor was performed using the participant means of the conditions described above from related and unrelated conditions. The main effect of priming was not significant, F_p (1,93) = 0.17, MS_e = 2568.09, p = .68. There was a significant main effect of experiment number, with Experiment 3 reaction times being significantly faster than those of Experiment 1, F_p (1,93) = 4.01, MS_e = 31915.45, p = .048. There was also a significant interaction of priming and experiment, F_p (1,93) = 4.33, MS_e = 2568.09, p = .04. Analyses investigating this interaction revealed, as would be expected, significant priming of low frequency exception word targets in Experiment 1, $t_P(46)$ = 2.00, SEM = 9.39, p = .03, but not in Experiment 3, $t_P(46)$ =0.42, SEM = 14.44, p = 0.34. See Figure 7.1.

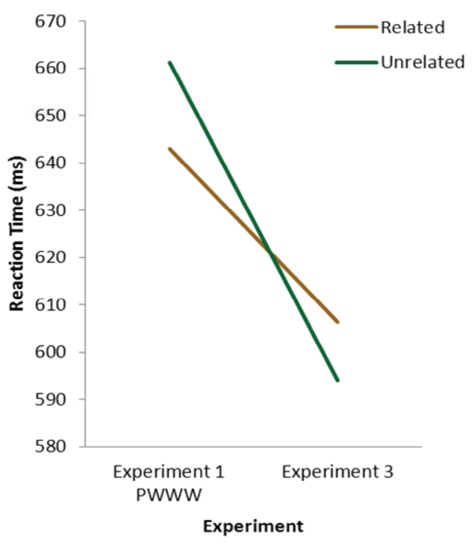

Figure 7.1. Participant mean reaction times for low frequency exception target words of Experiment 3 and the PWWW orders of Experiment 1 as a function of priming condition (related/unrelated). There is a significant interaction of priming condition and experiment.

Though caution should be taken when interpreting these results, as Experiment 1 means included a maximum of 3 items and this is an additional analysis, information is worth noting from these analyses. The pattern of priming results for low frequency exception target words significantly differed between Experiment 3 (PWWW) and the PWWW targets of Experiment 1. With low frequency exception word targets, when PWWW targets were the only modality order in the experiment, as in Experiment 3, the pattern of results differed (i.e., no significant facilitatory priming effects) from the results of an experiment with PWWW orders mixed with other modality sequences, as in Experiment 1. Inline with previously revealed patterns of priming in Experiment 1 data, this analysis, with a subset of the original data, considering targets from the PWWW order, also found significant priming of low frequency exception words. This analysis further highlights the odd pattern of results in Experiment 3. Therefore PWWW orders do not seem to produce interference results when in a mixed-modality sequence context, as in Experiment 1, but they do when they are presented as the only modality sequence. There is something about the design of Experiment 3 as a whole that has produced the unexpected trend for inhibitive priming in Experiment 3, and it could be the consistency of an experiments' modality pattern. It is worth noting that this result may highlight the importance of filler modality and filler modality patterns across experiments as a whole. This is explored further in the subsequent final discussion section of the priming experiments.

7.3 Final discussion of Experiments 1-4

In this thesis four long lag cross modal semantic priming experiments have been conducted in order to investigate whether semantic information contributes to the orthographic-to-phonological computation in healthy adult word reading. These four experiments have produced inconsistent results, but further analyses have provided additional information. The remainder of this chapter summarises the priming results of all analyses, which are also provided in Table 7.1; this is followed by a discussion of the results in relation to each other, the aim of this thesis, and the literature. This chapter concludes with an introduction next investigations.

Target Word Type	**Experiment 1** Prime Picture Picture Filler Word Filler Target Word	**Experiment1-Short List** Prime Picture Picture Filler Word Filler Target Word	**Experiment 2** Prime Picture Word Filler Target Word	**Experiment 3** Prime Picture Word Filler Word Filler Target Word	**Experiment 4** Prime Picture Picture Filler Word Filler Target Word
Low Frequency Exception (LFE) (glove)	× priming main effect n.s. ~✓ LFE planned comparisons approached significance *p*=.08 with median	~✓re-analysis with shortened stimulus list ✓Main effect of priming when LFE only combined with LFE from 1st half of Exp 4 ✓Simple effects analysis with LFE PWWW orders in analyses with Exp 3	✓ Main effect of priming	× priming main effect n.s. × planned comparison of LFE n.s × n.s. patterns of high frequency facilitatory priming and low frequency inhibition × n.s. facilitatory priming of Exp 3 LFE when analysed with PWWW of Exp 1	× priming main effect n.s. × planned comparison of LFE n.s. ✓Main effect with LFE of Exp 1 and 1st half of Exp 4 ✓Main effect with only LF in 1st half of Exp 4
Low Frequency Regular (LFR) (vest)	× priming main effect n.s.	× priming main effect n.s.	✓ Main effect of priming	× priming main effect n.s. × LF Simple effects analysis of pattern of LF inhibition n.s.	× priming main effect n.s. ✓Main effect with only LF in 1st half of Exp 4
High Frequency Exception (HFE) (blood)	n.a.	n.a.	✓ Main effect of priming	× priming main effect n.s. × HF Simple effects analysis of pattern of HF priming n.s.	× priming main effect n.s.
High Frequency Regular (HFR) (table)	n.a.	n.a.	✓ Main effect of priming	× priming main effect n.s. × HF Simple effects analysis of pattern of HF priming n.s.	× priming main effect n.s.

Table 7.1. Summary table of priming effects in participant reaction times for Experiments 1-7 as function of word type. ×= significant priming not found, ✓ = significant priming found, ~✓= priming approached significant, n.s. = not significant, Exp = experiment, n.a.= not applicable because word type not included, LF =low frequency

7.3.1 Results summary

In Experiment 1 (PFFW), which used only low frequency target word types, there was a trend for priming of low frequency exception target. Significant priming of low frequency exception words was obtained in subsequent re-analyses using only stimuli taken forward to subsequent experiments. This priming effect was then significant in new analyses of this chapter when Experiment 1 data was combined with first-half data of Experiment 4 and in simple effects analyses following a significant interaction with Experiment 3 data using this same stimuli subset. Experiments 2, 3, and 4 used four word target types: high frequency regular, high frequency exception, low frequency regular, and low frequency exception. Experiment 2 (PWW) provided the clearest result with a main effect of priming of all word types.

In Experiment 3 (PWWW) there was some weak evidence for different priming effects for high and low frequency words with high frequency targets showing a pattern of priming and low frequency targets showing a pattern of inhibition. A marginally significant interaction indicated possible category membership effects in the low frequency conditions in mean analyses. Low frequency regular target words were primed when preceded by a miscellaneous filler item, but inhibited when preceded by a specific category filler item. Low frequency exception target words were inhibited with both types of fillers. In Experiment 3, there is certainly no facilitatory priming of low frequency exception words, as was found in both of the previous experiments. New analyses in this chapter of low frequency exception words of Experiment 3 and the PWWW orders of Experiment 1 indicated that it is not the PWWW order per se that might change the pattern of priming of low frequency exception words. When PWWW is mixed with other filler patterns, such as PWPW, PPPW, or PPWW, as in Experiment

1, no inhibitory priming is evident. However, it may be the design of Experiment 3 as a whole that created the pattern of results that is opposite to those of other experiments.

In Experiment 4 (PFFW), the main effect of priming was not significant, nor was the planned comparison with low frequency exception words, though the data showed a non-significant facilitatory priming, for low frequency word types, including low frequency exception words, and high frequency exception words in an opposite pattern to the data of Experiment 3. There was no evidence of an effect of filler category membership either, as analyses from this manipulation produced non-significant results. New analyses of means in this chapter further investigated the possibility of priming of low frequency exception target word reading in Experiment 4. Experiment 1 data in conjunction with Experiment 4 data from the first half of the session, showed significant facilitatory priming of low frequency exception words, and this pattern did not differ between Experiments 1 and 4. However, as previous planned comparisons of low frequency exception word conditions in Experiment 4 did not reach significance, the new analyses of Experiment 4 benefited from the increase in power gained from Experiment 1 data. Analyses of only Experiment 4 low frequency word types (regular and exception) from the first-half showed a significant main effect of priming. However, no priming was found when low frequency items were the second half (i.e. high frequency before low frequency). Therefore, there may have been a previously unidentified effect of frequency block order. However, considering Experiment 4 analyses together, there are indications of significant priming of low frequency word targets in this experiment.

Because the original results within each chapter did not consistently yield effects significant at conventional level, and many of the significant results come from analyses that combine data or use only a particular set of stimuli, the interpretation of the results must be treated with a degree of caution. These results are worth noting, however, as they were obtained from a small number of hypothesis driven analyses and point to some consistent, yet small, significant priming effects that were observed, but non-significant, in the original analyses. The limited number of subsequent analyses were performed to aid interpretation of the results, and provided power to uncover these patterns, where previous original analyses yield non-significant effects. If this line of inquiry were to be extended it would, of course, be valuable to employ designs with greater statistical power.

7.3.2 Relevant priming literature and the results

The four semantic priming experiments of this thesis give some evidence of significant semantic priming of target words from a picture prime over an intervening item lag. Facilitatory priming was evident in low frequency target word types within certain experiments of this thesis. The priming investigations of the thesis began by replicating and extending the work Vitkovitch et al. (2006). As a reminder, significant priming of word targets was found after naming a related picture and two intervening items in one condition of Experiments 1 and 2 of Vitkovitch et al. However their aim and designs were somewhat different to those of this thesis. The two priming experiments of this thesis that were most similar in design to Experiment 2 of Vitkovitch et al., i.e., Experiments 1 and 4, also produced results that were the most consistent with Vitkovitch et al.'s study, replicating those priming effects.

The priming experiments of this thesis have also extended the work of Vitkovitch et al. by manipulating target word type. Priming within this thesis was found more reliably with some word types than others. Though the results seem to be dependent on the stimuli and factors included in the analyses, there is an indication of low frequency word priming over one and two intervening filler items. The most reliably primed word type was low frequency exception words, as this word was primed in Experiments 1, 2, and 4. The results of Experiment 2 and 4 also indicate that low frequency regular words were primed, which is consistent with the results of Experiment 2 Vitkovitch et al. (2006) in which the majority of target words were of this type. However, low frequency regular word priming was not evident in Experiment 1 and priming of low frequency word types was not found in Experiment 3. In contrast in Experiment 3, the opposite pattern, that of inhibition, was shown. The design as a whole may have been responsible for these opposite effects. It is also possible that filler modality might be responsible for the difference in low frequency results; this and other factors affecting priming are discussed in Section 7.3.4. High frequency word priming was found over one intervening filler item (Experiment 2), but was not found over two intervening items (Experiments 3 and 4).

As noted in Chapter 1 there are other factors that are correlated with frequency and regularity and it can be difficult to control them. Therefore though the target words were manipulated on frequency and regularity to create the four word types, “easy” and “difficult” word types may not have been a pure manipulation of these two factors (Section 4.2.2.2). For example, high frequency regular words may have been easier to read because they also were lower (earlier) in age-of-acquisition, and therefore were in fact high frequency regular low age-of-acquisition words. Though the significant priming effects are not in question here as target words are compared against themselves in related and unrelated conditions, other factors accounting for “easy” and “difficult” word types are. The discussion of frequency and regularity, word types, and confounded measures is returned to in Chapter 10 once subsequent investigations have been performed.

Of interest, here, is why there are differing patterns of priming results between high frequency “easy” and low frequency “hard” conditions. Difference in priming *between* the target word types could be to unmatched measures. Hines et al. (1986) found that the more typical an item is of a specific category, the larger the priming effect. However, both high and low frequency targets in this thesis were matched for category dominance, which would therefore mean that category dominance could not account for the difference between frequency types (Section 4.2.2.2).

Ideally prime-target pairs would be matched on semantic similarity across frequency type. However, low frequency related prime-target pairs used in this thesis were rated more semantically similar than high frequency related prime-target pairs (Section 4.2.2.3)[34], and lower semantic similarity could be responsible for the lack of high frequency priming over two intervening filler items. It has been claimed that priming can be found over intervening items, but that finding the effect may be dependent on a strong semantic relationship between the prime and target (Becker et al., 1997; Joordens & Becker, 1997; McRae & Boisvert, 1998; Plaut et al., 1995). When there is a strong relationship between prime and target, for example with a high number of similar semantic features, then though intervening filler items’ semantic representation may replace some of the prime’s activated semantic features, the filler items would be unlikely to wholly replace the many semantic features that the prime shares with the

[34] Differences in semantic similarity of prime-target pairs between high and low frequency could be an artefact of the word type manipulation. For example, target words that are suitable for high frequency regularity manipulation and also have a category coordinate in picture form may also naturally differ in their category coordinate semantic similarity as compared to low frequency target words. Of note is that within this priming design words are compared against themselves in related and unrelated conditions. Therefore any priming effects within word type are not due to differences in semantic similarity.

target, leaving some semantic activation for the target. The greater the number of shared features to begin with, the less likely they will all be replaced by the filler items' semantic activation. Significant priming of high frequency prime-target pairs in the current experiments was only found over one intervening filler item. It is possible that if the prime-target semantic relationship was stronger than that of the current pairs of high frequency words, then priming might be found over two intervening filler items perhaps. The difference in priming effects over two intervening filler items between high frequency and low frequency conditions may be due to unmatched stimulus parameters, such as semantic similarity, and not due to differences in a semantic contribution to the reading of the various word types. Therefore, semantic priming of high frequency words, though not a reliable effect, cannot be eliminated as a possibility.

Though Vitkovitch et al. (2006) found significant priming from a related picture prime to a word target over a two filler items, word priming has not consistently been found with word target naming in the literature (Sections 2.4.5.1 and 2.4.5.2.1) nor was it consistently found here in this thesis. Several priming studies have indicated that word target priming from a word prime to a word target over an intervening item might be limited to semantically biased button press responses (Becker et al., 1997; Joordens & Becker, 1997; Schvaneveldt et al., 1976). Bajo and Canas (1989) found priming of word targets from a picture prime with an intervening item when word and pictures were mixed. Significant priming of target words was attributed to a semantically biased route for word reading when picture primes are involved. Therefore the literature might indicate that priming of word targets over intervening items is only possible when word reading is biased toward semantic memory in some way, such as by task (e.g., semantic decision) or by context (e.g., mixed with pictures).

Priming was found (though not consistently) in the experiments of this thesis, which used word reading, and it is not likely due to a semantic biasing of word reading. If word reading was biased to semantic processing when pictures were used, then priming should be evident in all related word target conditions when primed by a picture, including those of this thesis. However, priming was not found in all word target conditions or in all experiments of this thesis, as a main effect was not present in all experiments. Therefore significant semantic priming of a word targets in the experiments of this thesis was not simply due to semantically-biased processing of word target reading.

Joordens and Besner (1992) found word target priming with word reading and one intervening filler item between the word prime and word target. Moreover, though Masson (1991) did not find significant differences there was a trend for priming in the means. The results of this thesis are in keeping with priming effect with no intervening items and with word primes and word targets (Cortese et al., 1997; Tse & Neely, 2007). However, experiments with no lag may be susceptible to non-semantic explanations of priming. The results of priming experiments of this thesis have added to the literature by demonstrated the priming of word targets is possible and it is possible over a lag, and it occurs with a task of word reading. This and other design aspects of the current priming experiments can be used to argue that the significant priming of word targets in this thesis is likely due to shared semantic information, and this discussed further in the subsequent section.

7.3.3 Evidence of a semantic memory contribution to word reading

As mentioned in Section 2.4.5.2, pictures can have long-lasting effects. The results from this thesis' experiments may show evidence of these long term effects with long term effects within semantic memory evidenced by facilitatory priming effects, and also possibly long term effects in links between semantics and phonology evidenced by non-significant inhibitory priming patterns (Damian & Als, 2010; Howard et al., 2006, Oppenheim et al., 2010; Vitkovitch & Humphreys, 1991), if other pre-activation non-semantic (Section 2.4.1) explanations can be eliminated. As inhibitory effects likely occur outside of semantic memory (Vitkovitch & Humphreys, 1991) and therefore may not be indicative of shared semantic information between prime and target, inhibitory patterns priming within these studies are not considered further; there were also no inhibitory effects that were significant at critical *p* levels.

Crucial to the central aim of this thesis is understanding whether the facilitatory priming effects from Experiments 1, 2, and 4 are due to shared semantic information, as opposed to pre-activation explanations, as this would provide evidence of a semantic contribution to word reading. In the main, there may be two explanations for a significant semantic priming effect (Section 2.4.1). Priming could be due to shared semantic information between related prime-target pairs (Damian & Als, 2005). Priming arguably due to this explanation could be evidence for a semantic information contribution to word reading. In contrast, priming could be explained with non-semantic accounts, i.e., the target word benefits from the pre-activation of its non-semantic information, such as the pre-activation of orthography or phonology. This could occur by automatic spreading activation or through strategic expectancy by participants.

Within the priming experiments of this thesis there are design aspects and analysis tools that can be used to help identify whether the significant priming effects from this thesis' studies are due to shared semantic information (Sections 4.1.2)[35]. Though these tools were used to implicate shared semantic information in the significant priming of Experiment 2, they are also relevant to the discussion of all of the priming results together. Certain design aspects made non-semantic pre-activation through strategic prediction of the target words very difficult. Prime-target pairs were not associated, target words were not high in category dominance, and this was a covert priming design. The random inter-mixing of related and unrelated stimulus sequences, and the inclusion of the intervening filler items, sometimes in varying modalities, made detection of the priming design difficult. Non-semantic pre-activation of targets through spreading activation was also not likely with this design as the amount of time between prime and target onset was longer (between 8.6 and 15.9 seconds minimum), which within spreading activation models activation is too long for activation to survive. Therefore if a target's orthography and phonology had been as pre-activated from the prime, then when the target came to be read, this pre-activation would have already subsided. Analysis tools also do not reveal indications of priming due to non-semantic means through strategic pre-activation. Participants do not report evidence of strategy in the post-experiment questionnaires, there is not priming of only high frequency words in any of the Experiments (Tse & Neely, 2007), and there is not greater priming in the second half in any of the experiments. There is, if anything, less priming in the second half of Experiment 4 (Section 7.2.1). There is, therefore, arguably no evidence for a non-semantic explanation of priming within the experiments of this thesis.

[35] Within the priming literature there are ways to investigate strategic versus automatic priming. This includes manipulation of the number of related and unrelated trials, and tasks such as letter search (Tse & Neely, 2007). The aim of the priming experiments of this thesis was not to investigate automatic versus strategic mechanisms of priming, so a reasonable set of markers for semantic versus non-semantic explanations of priming was established using of the design of the current studies.

The significant semantic facilitatory priming found in this thesis is more than likely semantic in nature. As the significant priming is likely evidence of shared semantic information between prime pictures (themselves) and target words, it is interpreted here as offering evidence of a semantic information contribution to word reading. It at least, exposes the relevant pathways of orthography-to-semantics-to-phonology or orthography-to phonology-to-semantics-to-phonology as detailed by the triangle model (Section 1.3.3). Therefore the priming experiments of this thesis suggest that semantic information contributes to low frequency exception word reading, and it may also contribute to low frequency regular word reading. A semantic contribution to high frequency word reading, as measured by semantic priming, cannot be discounted. However, as not all results were significant in participant and item analyses (only Experiment 2 was significant by both, though the results of Experiment 1 with the shortened stimuli list was marginally significant by both), and the stimuli words were a constrained selection, there may be low external validity and caution should be taken when generalising results to a larger sample of words.

7.3.4 Factors that might affect priming

The priming studies of this thesis also provided information about factors affecting priming. Priming results with a word target experiment may be dependent upon the design elements of the paradigm used (McNamara, 2005; Neely, 1991), and the results of the priming designs of this thesis were not an exception. Factors addressed in the following paragraphs are the reaction times of the experiment, the filler qualities, and the blocking of experiments by target conditions.

The results across all four studies of this thesis suggest that significant priming in these experiments was not solely present when reaction times were slow. Results of Experiment 2, which had the fastest reaction times of the four experiments and also had main effect of priming, suggest that priming is not simply dependent on target word reading times being generally slower, such as in Experiment 1, which had the slowest reaction times of the four experiments and also had nearly significant priming of low frequency exception words.

These four experiments also give an indication that intervening filler items, in number, modality pattern, and modality of the item itself may affect semantic priming. It is not entirely clear whether filler category membership affects semantic priming. Firstly, there is an indication that the number of intervening filler items is important, as the main effect of priming of all four words types from the experiment with only one intervening item, i.e., Experiment 2, differs from the varying results of the experiments with two intervening items, i.e. Experiments 1, 3, and 4.

Secondly, there is an indication that filler modality pattern within the experiment as a whole may affect priming. A new analysis in this chapter of Experiment 3 (PWWW) and Experiment 1 PWWW orders indicated that if PWWW orders form the whole of an experiment, then priming results may differ from experiments with mixed modality orders. This is also demonstrated by the differing pattern of results from experiments with different filler modality patterns (PFFW) differ from the results from Experiment 3 (PWWW).

Thirdly, there is an indication that the modality of intervening filler items themselves is important as the non-significant pattern of data from Experiments 3 (PWWW) and 4

(PFFW) are different, and the only difference between these two experiments' designs is the modality of filler items. This is also supported by the similar results Experiments 1 and 4, which have similar filler modality patterns and differ from the results of Experiment 3 that only had PWWW orders. It is also possible that filler modality is also responsible for the difference in low frequency priming results as Experiment 1 and 4 used pictures, and Experiment 3 did not. It is possible that these extra pictures or a higher proportion of pictures in the experiment as a whole help semantic activation to remain active in the semantic system. There may also be some hint that the modality of stimuli in some way interacts with filler category membership, as this category membership manipulation neared significance in the low frequency condition of Experiment 3 (PWWW), but was not significant in Experiment 4 (PFFW) when the only difference was modality.

Fourth, there might be an indication that specific filler items themselves effect on whether priming is found. Though there is no difference in low frequency exception word facilitatory priming between Experiments 1 and 4, there may be some evidence of a difference in the pattern of data between low frequency regular word priming of Experiment 4, as a main effect of priming is found in the new analyses using first-half low frequency data from Experiment 4 (PFFW), and was not found in Experiment 1 (PFFW, low frequency targets only). One of the central differences between Experiments 1 and 4 designs was the precise filler items used[36]. Therefore the priming experiments of this thesis suggest that individual intervening filler items themselves effect whether significant priming is found.

[36] The designs of Experiments 1 and 4 also differed in the prime-target pairs used and the precise SOA between items. Experiments 2-4 used the majority of Experiment 1's low frequency prime-target pairs (40 out of 56); they also included high frequency target word triplets/quadruplets unlike Experiment 1 which did not include high frequency target words. In Experiment 4 there was one additional second SOA between each item as compared to Experiment 1.

Though results of the filler category membership manipulations of Experiments 3 and 4 are inconclusive, it is evident that other filler qualities, such as number, modality, and possibly even the specific items themselves could ultimately interrupt the sharing of semantic activation from the prime to the target, and thus affect whether priming is found, even though the filler items are not semantically related to the primes and targets. Additionally if prime-target relationships are not very strong, as could be the case with high frequency prime-target pairs (Section 7.3.2), then this combined with the fillers could result in a lack of significant priming. These effects are discussed further with semantic memory models in the next section.

There is an indication of a block order effect in Experiment 4 data, and this (block order) may be another factor that affects whether priming is found. Though there was no indication of greater priming in the second half, and there was no indication that counterbalance had a significant effect, new analyses of Experiment 4, presented in this chapter, suggest that block order might also affect whether priming is found. This effect may not have been found in the previous order analyses 1) because, since all 16 counterbalances were examined in the analyses, there may have been of a lack of power to detect this effect, or 2) because in counterbalance analysis, and ultimately in all half analyses of this thesis, including the split-half analyses, frequency and half analysis cannot be separated out; they are confounded. Additionally, the analyses of low frequency targets in the first half and low frequency targets in the second half are not "pure" as the targets in the second half are always preceded by the other word type, whereas items in the first half are not. Therefore the effect, with priming in the first half, but not the second half with low frequency target words of Experiment 4 may be due to a difference in low frequency priming in first and second halves, or it could be due to low frequency priming being affected by being read after a block of high frequency target words, which was presented in the first half of the session. However, as no other first half effects emerged for the other experiments when analysing them in the same way as the additional analyses of Experiment 4, evidence for block effects are inconsistent, and the four experiments together only suggest that block effects are occasionally present.

7.3.5 Models of semantic memory and the results

The priming results of this thesis provide information concerning the models of semantic memory and their accounts of priming (Sections 2.4.1 and 2.4.4). The results from the priming studies of this thesis do not support an automatic spreading activation account of priming. Spreading activation models can account for very short lived activation with the upper threshold being around one second, regardless of the number of intervening items (Anderson, 1993; Collins & Loftus, 1975; Masson, 1995; McNamara, 2005; Ratcliff & McKoon, 1988). Therefore evidence of word target priming in experiments with minimum SOAs of 8.6 or 12.9 seconds, i.e., Experiments 2, 1, and 4 respectively, cannot be accounted for through either remaining activation of the prime, or automatic shared semantic information to the targets information, nor can it explained by automatic pre-activation of the target word's orthography and phonology by spreading activation from semantic memory, as activation does not survive for this length of time. Also, in spreading activation models the number of items between prime and target is not important, but the research of the thesis indicates that this affects priming. Either the spreading activation models are incorrect concerning the short-lived nature of activation (and the effect of the number of intervening items), or another account must be considered.

Distributed representation models of semantic memory allow for shared semantic information accounts of priming over a lag, when pre-activation accounts (whether automatic or strategic) have been eliminated. In distributed models of semantic memory, significant priming with longer SOAs and intervening items between prime and target is possible; time may not affect semantic activation from a related prime that is available for a target, but filler items could. Priming over unrelated filler items might be possible, insofar that filler items do not fully disrupt the activation of the related prime's semantic representation that is distributed across a series of links. In this thesis, priming was found over intervening filler items (Experiment 1, 2, and 4), but was also affected by qualities of filler items (Section 7.3.4). Therefore this thesis's priming effects support distributed representation accounts of semantic priming. However, there is more than one type of distributed representation account, and which one would be more likely to account for the priming results of this thesis is considered below.

In distributed model accounts by Masson (1991, 1995) and Plaut (1995) filler items play an important role in whether priming is found as they can disrupt the semantic activation of the prime with their own activated semantic representation. These accounts initially seem to be able to account for the effects of filler qualities seen in this thesis (Section 7.3.4), as filler representation would replace the activation of the related prime's representation and perhaps some qualities, such as modality, may be more likely to replace activation than others. Moreover Plaut places importance on the semantic relationship between prime and target. If the prime and target have a strong semantic relationship, then the filler may be less likely to replace all of the prime's activation that remains for the target. However, these accounts have modelled priming over one intervening item, not two (Masson, 1995; Plaut, 1995), and priming over two intervening items was found in this thesis (Experiments 1 and 4 in the additional analyses). Within these models multiple representations cannot be activated simultaneously, therefore priming over an intervening items only occurs if the representation of the filler is not fully settled in to a final activated state. Yet, within the investigations of this thesis, presentation time (until response) and task (naming) were such that the filler representation was arguably fully settled. For these reasons, therefore, though promising, the accounts of Masson (1995) and Plaut (1995) could likely not account for the effects found in this thesis.

An alternate distributed model proposed by Becker and colleagues (1997; Joordens & Becker, 1997) can account for priming over a multi-item lag and may be able to account for the significant priming found in this thesis (Experiments 1, 2, and 4) over a one and two intervening filler items. In this model, naming a prime item changes the connectionist network by altering the weights of the links so that they favour the information of the prime item's representation. This results in learning, and the target item, as it is related and has a similar semantic representation to the prime, can benefit from these long term changes. Filler items create new learning, altering weight within the network and this may include learning that occurred when naming the prime. In this model, factors affecting the amount of learning in the system are the depth at which items are processed and the semantic relationship of the prime-target pair (in the form of semantic features). Deeper processing and a stronger relationship created greater learning, and greater the learning more likely to result in priming of a target over a longer lag.

With the Experiments of this thesis, the same task (naming) is performed with all items in this paradigm therefore, based on this factor, fillers could affect weights in the network as much as primes and targets. For each filler item new learning would occur altering the weights in the network. One filler items might be less likely to alter all of the weight changes that occurred when naming the prime, hence why a main effect of priming is found in Experiment 2 with one filler item, and not in the other experiments that had two filler items. Within this model it does seem possible that various stimulus properties, such as filler modality or the filler item itself can affect the weights within the network, thereby affect whether priming is found. That is, different filler characteristics could affect depth of processing, and filler identity, though unrelated to the prime and target, may (or may not) affect the weight changes instigated by the prime. It may also be able to account for possible effects of semantic relationship differences, as priming is more likely with a stronger relationship, which could explain the lack of high frequency priming over two intervening items.

In conclusion, the priming results of this thesis suggest significant priming effects over one and two intervening filler items and indicate that priming is affected by aspects such as semantic relationship of prime and target and qualities of filler items, including number, modality and the item itself. These results support a distributed representation account of priming and might be specifically accounted for by the distributed model of by Becker and colleagues (Becker et al., 1997; Joordens & Becker, 1997), which accounts for priming through "long-term" learning.

7.4 Conclusions

In summary, the evidence from the four priming studies of this thesis suggest that orthographic-to-phonological computation may involve a semantic memory contribution. The priming studies specifically offer evidence of a semantic contribution to reading of the most difficult word types, here labelled low frequency exception, when primed by a picture. There is also a suggestion of a semantic contribution to low frequency regular word reading. The results for the easiest target word type, high frequency regular words, and also high frequency exception words are less than clear. Though priming of high frequency words was not found over two intervening items, these words were primed over one intervening item. Lack of significant facilitatory priming over two filler items could be due to the weaker semantic relationship of prime and target, and the qualities of the filler items. Therefore a semantic contribution to word reading may be possible with high frequency word types. It is clear that whether a semantic priming effect is found is partially dependent on the specific intervening filler items and on qualities such as number and modality, and also, in Experiment 4 at least, frequency block order. These preliminary filler effects should be investigated in future. (See General Discussion for more on future investigations of filler type). The results of these priming studies are best accommodated by the distributed models of semantic memory of Becker and colleagues (Becker et al., 1997; Joordens & Becker, 1997).

Further investigations into the aim of this thesis are warranted as the results of these priming studies offer some evidence of this phenomenon. It is also of interest to see if this suggestion of a semantic contribution of orthography-to-phonology contribution in priming can be seen in single word reading. By moving to a regression design that does not use primes, fillers, or targets, the effect of filler items will not need to be considered. A regression design also provides the opportunity to further investigate whether semantic information contributes to word reading and whether this contribution is the same for various word types while statistically controlling for certain variables that may have been confounded across word types in the priming experiments. An alternate design also allows for the use of a greater number of word stimuli. For example, stimuli selection will no longer be constrained by the need to find words with category coordinates in picture form. The following chapter continues to investigate whether semantic memory contributes to word reading with healthy adults by using the complementary design of regression. While Experiments 1-4 have suggested semantic contribution effects during word reading, using pictures primes and word targets, and the results are consistent with relevant pathways to phonology, i.e., orthography-to-semantics-to-phonology, or orthography-to-phonology-to-semantics-to-phonology, the following chapter will now investigate whether a semantic effect can be obtained without the use of semantic priming designs.

Chapter 8

Regression Investigations

8.1 Introduction

The experiments of the previous chapters used semantic priming techniques to investigate a semantic contribution to orthography-to-phonology computation in healthy adults. These priming experiments, however, provided inconsistent results that suggest semantic information contributes to low frequency exception word reading, but its contribution to other word types is still unclear. Further investigations, therefore, are needed to investigate the central aim of this thesis. This chapter concentrates on regression analyses to explore a possible semantic contribution to word reading.

In the following introduction sections, there is (a) a brief review of the priming results and an introduction to regression methods for use in this chapter; (b) a review of literature that has used regression techniques to investigate a semantic contribution to word reading, as this was only briefly summarised in Chapter 2; (c) an introduction to elements important to the analyses of this chapter; and finally (d) aims and predictions. Sections for the methods, results, and discussion then follow.

8.1.1 A brief review of the priming results and a regression introduction

The priming experiments may have provided evidence of a semantic memory contribution during word reading. This evidence, however, was not consistent across priming experiments. Though the experiments give a suggestion of a semantic contribution to low frequency word reading, it is not clear from the results of the priming experiments whether this contribution occurs with all words, including high frequency words. Additionally, unmatched stimulus measures meant that care had to be taken when comparing priming results across word type; variables confounded with word type meant that priming effects found with one word type might be due to measures other than frequency and regularity. There were also elements of the priming paradigms, e.g., filler qualities also affected the priming results. Therefore, to aid in the interpretation of these priming results and to further investigate the central aim of this thesis, converging results are sought using regression design and analysis as a complementary method to explore whether there is a semantic contribution to orthography-to-phonology computation.

Factorial design is only one approach that can be used to study whether semantic information can contribute to word reading. Chapter 2 presented regression investigations that were used as complementary methods to factorial designs. There are a number of benefits, as pointed out by Balota et al. (2004), to using regression methods. Firstly, regression methods allow for the use of a large number of words, which naturally vary along the spectrums of frequency and spelling-to-sound consistency (Section 1.2 and subsections). Words are not divided into orthogonal categories of four word types within a regression design, as they are in a factorial design. Regression design offers the opportunity to utilise a majority of words without having to eliminate items that could not be easily categorised. Additionally, regression design also allows the opportunity to investigate many variables from previously published databases by statistically controlling for each factors' effects. This is unlike a factorial design, in which, when perfectly designed, one measure is manipulated while all other measures are controlled. Regression methods offer the opportunity to study the impact of variables on word reading while statistically controlling other variables. This includes variables such as imageability, semantic features, and age-of-acquisition. The impact of each factor on word reading can be investigated as the analysis can calculate the unique impact of each factor while accounting for others. Lastly, regression methods within this thesis also provide the opportunity to extend regression studies within the literature by investigating a newer semantic measure, that of semantic features, as will be described in more detail in subsequent sections. A selection of large regression analyses will be reviewed in the following section in order to clarify the approach.

8.1.2 A literature review of large regression studies

In Chapter 2, the literature review included research that used factorial and relatively smaller-scale regression analyses to investigate a semantic contribution to orthography-to-phonology translation. There are also investigations, as briefly mentioned in the same chapter, that have used large regression analyses alone to investigate measures that predict word reading times. The following section reviews large published regression analyses before introducing the regression analyses of this thesis.

Many of the regression studies reviewed here have used the reaction times from the English Lexicon Project (ELexicon). As briefly mentioned in Chapter 2, ELexicon is a large database of word reaction times, both reading times and lexical decision times, and some lexical measures (Balota et al., 2007). The ELexicon project used over 1,200 healthy adult participants with a mean age of 23 years. Within the ELexicon database, there are word reading times and lexical decisions times for over 40,000 words, with each participant only receiving a maximum of 2,531 words for response. To provide the data for all of the words, a Z-score was calculated for each of the responses (Z-score of the reaction times and the ratings) for each of the words, and for each word the Z-score reaction times (or ratings) were averaged together. ELexicon has been used by Balota et al. (2004) for analyses, which will be reviewed in detail in the following section. ELexicon information, including reaction times, has been made available for other researchers to use (Balota et al., 2007).To pre-empt, a small subset of these ELexicon reaction times are used in the analyses of this chapter.

8.1.2.1 Regression analyses, single-syllable words, and ELexicon

Regression analyses using ELexicon reaction times and lexical and semantic predictors have been performed and published by Balota and colleagues (2004), as was briefly summarised in Chapter 2; this research is reviewed in detail here. They investigated factors that influence lexical decision and word reading reaction times in younger (average age 20.5 years, similar in age to participants of the priming studies of this thesis) and older healthy adults (average age 73.6 years). From the ELexicon database, Balota and colleagues used the reaction times to over 2,400 single-syllable words (that varied in frequency and a number of factors) and these analyses included a large number of predictor variables. (Analyses of item mean reaction times with semantic level variables by Balota et al. (2004) used between 997 and 2,342 items, depending on the analysis). The focus of this review considers the item analyses of word reading reaction times and the impact of semantic variables because this is of direct relevance to the central aim of this thesis.

Hierarchical multiple regression analyses of word reading reaction times with three sets of predictors were performed (Balota et al., 2004). Each of the three sets contained many variables. Surface level variables, such as initial phonemes, were entered first, followed by the lexical level block, such as frequency. Finally, and most importantly to the central aim of this thesis, a block of semantic level variables was entered as the third block. Within this block there were three sets of semantic measures; therefore three analyses were run with only the semantic block differing between them. A total of six semantic measures were used as predictors across three analyses.

The first analysis used semantic set size (using the number of associates produced to an item in Nelson, McEvoy, Schreiber, 1998), imageability, and meaningfulness scores (the number of associates for the word and how strongly that item comes to mind, from Toglia and Battig, 1978) as the semantic block of predictors. The second semantic analysis used newer imageability ratings of Cortese and Fugett (2004) as the semantic block of predictors. The third analysis of semantic variables was composed of two measures of interconnectivity as the semantic block of predictors. Measures of interconnectivity were calculated by accounting for the number of connections *from* a word *to* other words, as well as how many connections there are *to* that word *from* other words. One measure of interconnectivity was the log of Nelson et al.'s connectivity measure, which is calculated based on the number of associates produced for a word. Associates may not share meaning, e.g., semantic features; for example, "mouse" and "cheese" are associated in that a mouse is traditionally thought to eat cheese. However as one is a small rodent and one is a dairy food, they do not share any common semantic meaning. Therefore, a second measure based on shared meaning was used in the regression analysis as well. The second measure of interconnectivity was the log of WordNet (Miller, 1990). WordNet is a large electronic database that clusters words together based on their meaning, linking them together by listing similar semantic sets when a word is searched for. Using this measure, it is, therefore, possible to calculate the number of words that are similar in meaning to a given word, and it is a log of this measure that was used by Balota et al.

In the regression analyses of Balota et al. (2004), semantic variables were a significant predictor of word reading times in all analyses after statistically accounting for surface and lexical variables in other stages of the analysis. Results differed somewhat between younger and older adults, with the semantic level variables predicting a higher percentage of variance in younger adults' word reading times. The effect of each semantic variable was also examined on its own. Both imageability ratings (Cortese & Fugett, 2004; Toglia & Battig, 1978) were significant predictors of the word reading times of young adults. For older adults' reading times only Cortese and Fugett's imageability and WordNet connectivity were significant predictors. Therefore, the semantic blocks were significant predictors of word reading times with Cortese and Fugett's imageability ratings being the most robust semantic measure of the six semantic predictors and WordNet connectivity being second. Because the semantic measures were significant predictors of word reading reaction times, the authors conclude that semantic information contributes to healthy word reading.

Since the results of Strain et al. (1995, 2002) (Section 2.3.2) and the simulations of the triangle model of word reading (Plaut et al., 1996) (Section 1.3.3.), as reviewed in Chapter 2 and Chapter 1 of this thesis respectively, argue that a semantic contribution is most likely with difficult English words, Balota and colleagues also studied the interaction of word type (frequency and regularity) and imageability to explore whether the semantic predictors in their analyses predicted variability in the reading times of one word type more than the other. The three-way interaction of imageability, frequency, and regularity, with low frequency exception word reading showing an imageability effect measures only approached significance. Recall, though, in Experiment 1 of Strain et al., the three-way interaction only approached significance. It was the two-way interaction of regularity and imageability, with an imageability effect only in low frequency exception word reading that was significant in the Strain et al. studies. When Balota et al. investigated this two-way interaction of semantic level variables and regularity in low frequency words, it also only approached significance. Balota et al.'s interpretation of the results of the analyses is presented in the next paragraph.

Because semantic measures significantly contributed to the regression models and because the semantic measures account for a significant and unique proportion of variance in word reading times, the authors conclude that the results from their large regression analyses offer support for a semantic contribution to word reading. However, because the results from the interaction analyses were not significant and therefore did not single out only the most difficult word type (low frequency exception words) as receiving a semantic contribution during word reading, Balota and colleagues (2004) concluded that semantic measures predict variance in word reaction times for all one-syllable words, not solely low frequency exception words. Since measures that capture semantic information affect word reading times in these regression analyses, this is taken as evidence of a semantic contribution to orthography-to-phonology computation. The authors' conclusions that semantic information contributes to the computation of phonology during word reading are in keeping with the conclusions of Strain et al.

(1995) and Plaut et al. (1996), who also argue that semantic information can contribute to the reading of all words. These results have not gone without criticism; other researchers, however, have concluded that other measures might account for these results, and this is presented in the following section.

8.1.2.1.2 Regression analyses and age-of-acquisition

As briefly noted in Chapter 2, other research that used regression analyses to investigate factors that influence word reading offered an alternative explanation of Balota et al.'s (2004) results (Cortese & Khanna, 2007). Because age-of-acquisition was not included in the regression analyses of Balota et al. (2004) and because critiques of Strain et al. (1995, 2002) showed that age-of-acquisition might affect word reading (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002), the regression analyses of Cortese and Khanna, reviewed here, investigate whether age-of-acquisition might be responsible for variance previously attributed to imageability.

First, a regression analysis was performed using age-of-acquisition as the dependent variable and surface level factors, lexical level factors (e.g.,frequency), and semantic factor (i.e., imageability) as predictors. The results revealed that lexical variables and imageability were significant predictors of the variability in age-of-acquisition ratings. Since variance in age-of-acquisition ratings are significantly and uniquely predicted by not only non-semantic measures, but also semantic measures, Cortese and Khanna (2007) concluded that the age-of-acquisition effects within their regression analyses were semantic in nature, though they cannot categorically eliminate the possibility that age-of-acquisition is a non-semantic variable.

Of significance is whether age-of-acquisition is itself semantic in nature. As mentioned in Chapter 2, age-of-acquisition could influence the strength of connections between orthographic and phonology, or representations within phonology (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002). If these were the only connections influenced, then age-of-acquisition would not be a semantic variable; this would then mean that age-of-acquisition effects are not evidence of a semantic contribution to word reading. However, age-of-acquisition might also capture connections between semantic memory and other systems, such as those between semantics-and-phonology, as, Cortese and Khanna argue, is evidenced by the semantic measure of imageability being a significant and unique predictor of the variance in age-of-acquisition ratings. In recent accounts age-of-acquisition characterises "arbitrary" connections, such as those between semantic memory and orthography, rather than "quasi-regular" links, such as those between orthography-to-phonology in English (Belke, Brysbaert, Meyer, Ghyselinck, 2005; Brysbaert, VanWijnendaele & DeDeyne, 2000; Brysbaert & Ghyselinck, 2006; Lambon-Ralph & Ehsan, 2006; Monaghan & Ellis, 2010; Steyvers &Tenenbaum, 2005). This might also be demonstrated by picture naming times being more affected by age-of-acquisition than word reading times. The discussion section of this chapter will expand on the locus of age-of-acquisition.

In the subsequent analyses of this series, Cortese and Khanna (2007) used the word reading reaction times of healthy adults as provided by the ELexicon database and the three sets of predictors: surface level factors, lexical level factors, and semantic factors, in regression analyses similar to those of Balota et al. (2004). In contrast to the three analyses with six semantic variables used by Balota et al., the regression analyses of Cortese and Khanna (2007) included the newly collected age-of-acquisition ratings and Cortese and Fugett's imageability ratings as the third level of semantic level variables.

When age-of-acquisition was excluded from the regression analysis, results were the same as in Balota et al. (2004), i.e., imageability was a significant predictor of word reading times. However, when age-of acquisition was included, it, not imageability, was a significant predictor of reading reaction times. Though age-of-acquisition, instead of imageability, is a significant predictor of word reading times, when other factors have been statistically controlled for, Cortese and Khanna conclude, similar to Balota et al. (2004), that their analyses provided evidence of a small semantic contribution to word reading. They claim this because they argue that their first analysis and the theories of age-of-acquisition show age-of-acquisition to be a semantic variable, and the analyses show age-of-acquisition as significant predictor or word reading times. Therefore the semantic contribution to word reading in their analyses is captured with age-of-acquisition ratings not imageability.

Another set of regression analyses also uncovered effects of age-of-acquisition on word reading times, while also including other semantic level factors (such as, imageability) as predictors in the analyses. Brown & Watson, 1987 central aim-to compare subjective familiarity ratings and objective word frequency counts- was quite different to that of this thesis, yet their analyses and results are relevant to the empirical investigations of this Chapter. In two multiple regression analyses, either familiarity ratings or word reading times, as provided by healthy adult readers, were used as dependent variables, respectively. Multiple predictor variables were entered into the analyses as well, including lexical measures (e.g. frequency), semantic measures: ambiguity, imageability, and concreteness, and age-of-acquisition ratings. The largest predictor of variance in both analyses was age-of-acquisition ratings. Additionally, the semantic factors of ambiguity, imageability, and concreteness were not significant predictors of either analyses, and did therefore not significantly predict any variance in familiarity ratings or word reading reaction times.

Because age-of-acquisition is a significant predictor of word reading times and ambiguity, imageability, and concreteness are not, of importance, as highlighted previously, is the locus of age-of-acquisition. If age-of-acquisition is interpreted as semantic in nature, then significant age-of-acquisition effects can indicate a semantic contribution to orthography-to-phonology formation. However, if age-of-acquisition is interpreted as measuring non-semantic information, then effects of age-of-acquisition on word reading times in the absence of any other know semantic effects, e.g., imageability effects, do not indicate a semantic contribution to orthography-to-phonology formation. Because age-of-acquisition significantly predicts the non-semantic variable of familiarity, Brown and Watson interpret their results as providing evidence for the phonological completeness hypothesis of age-of-acquisition in which age-of-acquisition is not semantic in nature, but instead measures phonological representations. With this interpretation of age-of-acquisition, Brown and Watson conclude that their results do not provide evidence of a semantic contribution to word reading, even though the results are similar to Cortese and Khanna (2007).

Closer inspection of the stimuli used in these analyses revealed that words were high in frequency. Words were limited to items present in the Gilhooly and Logie (1980) and Brown (1984) corpus, and were present with at least 100 instances per million in the norms of Thorndike and Lorge (1944). Therefore, this study may only provide information about variables that influence only high frequency word reading.

8.1.2.1.3 Conclusions concerning age-of-acquisition effects

From the regression investigations, the correlations between the variables (Brown & Watson, 1987; Cortese & Khanna, 2007) and previous word reading research (Monaghan & Ellis, 2002), as reviewed in this Chapter and in Chapter 2, it is evident that imageability and age-of-acquisition are often confounded. As a consequence, results that seem to initially provide evidence for a semantic contribution to word reading through imageability effects are often later accounted for with age-of-acquisition measures (Balota et al., 2004; Strain et al. 1995, 2002). Evidence of a semantic influence on word reading may then be dependent on the interpretation of the source of age-of-acquisition influence. Accounting for semantic level variables and age-of-acquisition could reveal information about these measures and information important to the aim of this thesis. The ELexicon regression analyses by Cortese and Khanna and the previous work by Brown and Watson that included age-of-acquisition improved upon other research that failed to include it as a predictor. It is clear from the research on Cortese and Khanna, Brown and Watson, and Ellis and Monaghan that including age-of-acquisition is important, especially when analyses also include other semantic variables, such as imageability.

8.1.2.2 Regression analyses, multi-syllable words, and ELexicon

Recent research has used ELexicon reading reaction times and regression analyses to investigate factors that influence not only single- but also multi- syllable word reading (Yap & Balota, 2009), extending the research of Balota et al. (2004). The analyses were similar to Balota et al. with surface, lexical, and, most importantly, semantic level sets being used as predictors. Within the analyses of Yap and Balota two semantic variables were included in the semantic level, but neither of them were imageability. The two semantic variables were the number of senses, which is the log of the number of meanings a word possesses, as counted in WordNet (Miller, 1990). This measure was the second best semantic predictor from the analyses of Balota et al. (2004). The other was semantic neighbourhood size, which is the number of neighbours within a certain distance as modelled within semantic space.

Both semantic variables were negatively correlated with reaction times; a higher number of senses or a higher number of neighbours resulted in words being read more quickly. Results of the regression analyses revealed that the semantic level predictors significantly contributed to the model, once other factors had been accounted for. Semantic factors predicted the variance in word reading times for all words and also for single- and multi- syllable words when analysed as individual sets. Drawing on the previously reviewed investigations of Strain and colleagues (1995) and Balota et al. (2004), further analyses were performed to investigate any possible interactions of word type (frequency by regularity) and semantics; however, no significant interaction was found.

Yap and Balota (2009), however, may lessen the impact of a claim for a semantic contribution to word reading with the computer model simulations. Two multi-syllable models, the junction model (Section 1.3.3; Kello, 2006) and the CDP++ (Section 1.3.2.2; Perry et al., 2010b) were used to attempt to simulate the results of the human performance data from Yap and Balota's analyses. The CDP++ model successfully simulated the human performance from the regression analysis, including the impact of semantic variables on word reading times; yet, this model does not actually have implemented semantic representations. Yap and Balota acknowledge two possibilities as to why a model with no semantic memory system could correctly reproduce semantic effects. First, semantic variables might contribute to word reading, and this word reading model (CDP++) captured this accidentally with differences in orthography-to-phonology computation. Plaut et al. (1996) demonstrated that is may be possible to capture semantic effects on the orthography-to-phonology pathway with no semantic system, though Plaut and colleagues directly 'lesioned' their model, by removing phonological connections and thereby changing the weights on the orthography-to-phonology pathway, when attempting to simulate semantic effects. Alternatively, as suggested by Yap & Balota (2009, p. 26) the semantic effects found in the human data may instead be due to an, as yet, unmeasured and unidentified confounding variable(s).

Because in the main regression analyses semantic variables are significant predictors of word reading reaction times, there is an indication of a semantic contribution to orthography-to-phonology computation. Analyses into whether there is a semantic contribution to only the most difficult word types (i.e., low frequency exception words) did not find significant results; therefore these investigations indicate that a semantic contribution might occur with all words regardless of type. However, though these results are initially promising, after considering simulation evidence from the computational models of word reading, whether the studies of Yap and Balota actually indicate a semantic contribution to orthography-to-phonology computation is unclear. It will be of interest to examine whether an alternative semantic variable, such as semantic features, may provide a clearer indication of the involvement of semantic information in word reading, especially when age-of-acquisition is controlled.

8.1.2.3 Regression analyses and number of semantic features

Regression methods have also been used to investigate whether the semantic measure of number of semantic features affects word reading times (Pexman et al., 2002), and these analyses can provide information relevant to the aim of this thesis. Also, shortly, the number of semantic features will be shown to be a relevant variable in view of Experiments 1- 4 of this thesis. Pexman and colleagues used the semantic features of McRae and Cree (2002) as the predictor of interest, and this variable can be considered as measuring semantic information as will be described here. Semantic feature norms are a compendium of lists of features produced by a population for a set of items. For example, participants asked to list defining features for the item “duck” may include statements such as “is a bird”, “is an animal”, “flies”, “lays eggs”, “has wings”, “has a beak”, “has feathers”, “lives in water”, “is edible”, etc. (McRae, Cree, Seidenberg & McNorgan, 2005). The number of semantic features for a word may capture its semantic richness because it is a verbal expression of the semantic information in semantic memory about an item.

Semantic features are positively, though only lowly, correlated with the word’s concreteness and imageability. As indicated Section 2.3.2, semantic richness of a high imageability word may be partially due its number of semantic features (Strain et al. 1995). Therefore potentially there is some overlap in the two measures (semantic features and imageability), and this is explored in the empirical investigations of this Chapter and the next Chapter. Of direct relevance to this thesis and the investigations of this Chapter is Experiment 1C of Pexman et al. (2002). Regression methods were used with lexical level predictors and semantic features as the semantic level predictors and word reading times as the criterion variable. Once the influence of lexical variables was removed, the number of semantic features was a significant and unique predictor of word reading reaction times. Because semantic features are considered a semantic measure that captures semantic information and it impacts on word reading times, the results of this analysis suggests a semantic contribution to orthography-to-phonology computation. It is worth noting that Pexman et al. (2002) did not include imageability or age-of-acquisition in their analyses, so whether semantic features capture information unique from these other variables is unknown. It is relevant, therefore, to investigate whether semantic features account for a different proportion of variance than other known predictors of word reading time, such as imageability and age-of-acquisition, using ELexicon's published naming times for words.

8.1.3 Elements of the regression analyses of this thesis

Analyses presented in this chapter will use regression methods to explore whether semantic variables affect single word reading times. The next sections discuss further the important elements of the current analyses, including additional information about a central semantic measure, that of semantic features, in relevance to the priming Experiments 1- 4 of this thesis, and the criterion variable (ELexicon word reading times). This is followed by sections on aims and predictions.

8.1.3.1 McRae et al. (2005) semantic feature production norms

A set of semantic feature production norms has recently been published by McRae and colleagues (2005), for over 500 living and nonliving items up to four syllables in length. The items used in these norms were similar to the target words from priming Experiments 1-4 of this thesis and the semantic features measure itself may capture the relationship between related prime-target pairs. Concepts used in the McRae et al. (2005) semantic feature norms are similar to the target words of Experiments 1-4 in this thesis. Target words were object names in word form and therefore, represented highly picture-able items. The concepts used in the semantic feature norms were also names of highly picture-able items. In this way, the items of the McRae semantic feature norms are similar to the target words of Experiment 1-4.

Semantic features may also capture relationship between the related prime-target pairs of Experiments 1-4, as McRae and colleagues (2005) claim semantic features, generally, can be used to explain semantic priming results. The prime-target pairs of Experiments 1-4 were chosen to be conceptually similar and have shared category membership. A prime-target pair that were both animals, such as “sheep” and “goat” may also have a visual similarity, though this is not true for every related pair. There were also conceptually similar prime-target pairs that shared category membership, but that do not share visual similarity, such as clothing category pair “hat” and “glove”. Conceptual similarity and shared category membership can be measured by semantic feature norms, such as those published by McRae et al. The results from the priming studies can be briefly considered in these terms.

Semantic priming results of Experiments 1-4 might be the result of remaining activation in shared semantic features between the prime- target pairs. For example, within a distributed model[37], if the target word “pear” is primed by the prime picture “apple”, then naming the picture “apple” activates semantic information associated with that item, including the representations associated with its semantic features, such as “is a

[37] It is worth noting that the connectionist triangle model of word reading uses semantic features to capture the semantic memory component of their model (Harm & Seidenberg, 2004), and describe semantic effects in practical terms using semantic features (Plaut et al., 1996).

fruit", "has seeds", "is round", "grows on a tree", "domestic", "is edible", etc. The related target word "pear" may benefit from the activation remaining in the shared semantic representation, as many (but not all) of the semantic features of the two items are the same. If semantic features are shared between related primes and targets, and to the extent that priming is accounted for by shared semantic information, as opposed to non-semantic accounts, then it might be reasonable to expect that words with a greater number of semantic features will benefit over words with a lower number of semantic features, as suggested Pexman et al. (2002). Regression analyses may provide converging evidence for such an account. Semantic features therefore will be used to investigate the aim of this thesis. Using the feature norms of McRae et al. (2005) will also allow for the investigation of a new semantic measure.

8.1.3.2 Criterion variable

The regression analyses of this chapter require three pieces of information: (1) stimulus words of interest, which are provided in the form of a subset of concept words used by McRae et al. (2005), (2) predictor measures for these words, of which one central variable is the McRae semantic feature norms, (3) and a dependent variable, which is the ELexicon word reading times (Balota et al., 2007), as introduced in Chapter 2 (Section 2.3.3) and in this chapter (Sections 8.1.2). ELexicon word reading reaction times, as well as some predictor variables, are available for a subset of concepts from McRae et al., and it is this information that is used in the analyses of this chapter

8.1.4 Aims and predictions

The analyses presented in this chapter use a complementary method of regression to investigate the aim of this thesis. This study also extends the regression analyses of the literature (Balota et al., 2004; Cortese & Khanna, 2007; Ellis & Monaghan, 2002; Pexman et al. 2002; Strain et al., 2002; Yap & Balota, 2009) by including a new set of semantic feature norms (McRae et al., 2005) and other known predictors, namely imageability and age-of-acquisition. It includes one and two syllable words, which extends previous regression analyses that used only single syllable words in the main (Balota et al. 2004; Cortese & Khanna et al., 2007; Ellis & Monaghan, 2002; Monaghan and Ellis, 2002; Strain et al., 1995, 2002). The hierarchical regression analyses were performed to be comparable to those of Balota et al. (2004), Cortese & Khanna (2007), and Yap & Balota (2009) as these are large regression analyses that have set a standard when using ELexicon word reading times. The investigations seek to determine whether imageability, age-of-acquisition, and semantic features account for different and unique proportions of the variance in word reading times in a series of analyses.

These regression analyses are concerned with all words, regardless of type, and the factors that influence their reading reaction time. Using the foregoing review as a guide, McRae semantic features, imageability, and age-of-acquisition are force entered as a final third level of variables in series of hierarchical regressions, assuming there is no multi-collinearity between them. If this block of variables is a significant predictor of word reading time variance, after accounting for other variables, then this may indicate that semantic information is contributing to word reading. The individual influence of each of the semantic level factors (imageability, semantic features, and age-of-acquisition) will also be investigated. For example, if imageability and semantic features are unique predictors then this could indicate that they measure different aspects of semantic knowledge. If they do not remain unique predictors when entered together, then they may be measuring a similar aspect of semantic knowledge.

Previous analyses indicate that age-of-acquisition is a significant predictor of word reading times, and that it can eliminate imageability as a significant predictor (Brown & Watson, 1987; Cortese & Khanna, 2007). If age-of-acquisition subsumes the effects of the other semantic predictors, then whether semantic information contributes to word reading would then be subject to interpretation since age-of-acquisition can be interpreted as having non-semantic and/or semantic locus (Brown & Watson, 1987; Brysbaert et al., 2000; Zevin & Seidenberg, 2002); further consideration of age-of-acquisition will be important. If both or either of the known semantic variables remain a significant predictor of word reading times after the effect of age-of-acquisition has been accounted for, then this would lend stronger support to an argument for a semantic contribution to word reading than age-of-acquisition alone being a significant predictor.

8.2 Methods

8.2.1 Stimulus words

The 427 one and two syllable word stimuli for the regression analyses reported in this chapter had available semantic feature values from McRae et al. (2005) and ELexicon word reading times. The words used in the analyses of this chapter were the names of objects from various categories. Examples are “boat”, “peach”, “eagle”, “pencil”. One and two syllable words were chosen, as measures for words longer than two syllables are rarely available, e.g., imageability and familiarity. Additionally, the semantic priming studies of this thesis also used one and some two syllable words as target word stimuli. Therefore, the words were also selected in keeping with the target stimuli used in priming experiments 2-4. These selection criteria, however, may limit the extent to which the results can be generalised, and this is discussed in Chapter 10.

All words (100%) were provided with a reaction time, and measures of length, frequency, and Coltheart's N. Nearly 100% (98%) of words were provided with a measure of familiarity and number of semantic features. For imageability and age-of-acquisition, 71% and 77% of words, respectively, were provided with measures. Table 8.1 provides the percentage of words that were provided with each measure and the average of that measure, amongst other information.

Measure	N	%	Avg	SD
Word Reading Times (ELexicon)	427	100%	631.11	57.24
Length (letters)	427	100%	5.34	1.56
Frequency (Celex)	427	100%	20.26	49.54
Familiarity (McRae)	418	98%	5.79	1.979
Coltheart's N	427	100%	4.51	5.48
Age-of-Acquisition	329	77%	3.44	1.15
Imageability (MRC)	304	71%	583.53	33.76
Semantic Features (McRae)	418	98%	13.46	3.58

Table 8.1. Descriptive information for measures used in the regression analyses. The number (N) and percentage (%) of words that are provided with a measure along with the average (avg) and standard deviation (SD) for that measure across the stimulus words.

Semantic feature values were available for nearly 50% of the target words from Experiments 2-4, and the shortened list of Experiment 1 of this thesis. Twenty-five percent of the high frequency exception word targets, 55% of the high frequency regular word targets, 45% of the low frequency exception word targets, and 65% of low frequency regular word targets from the priming studies were contained in the database. These target words were included in the current regression analyses.

Overall, there were not an equal number of the four word types in the 427 words of this analysis. Less than ten percent of the words were exception words. Words of all frequencies were included, with a Celex written frequency range of zero to 558 per million. Around 20% of the words were high frequency using the high frequency Celex threshold from the semantic priming experiments (23 per million). The words had an average Celex written frequency of 20 instances per million. Though the four word types were represented, the majority of the words were low frequency (less than 14 per million) and regular in spelling-to-sound correspondence.

8.2.2 Predictor variables

A large number of predictor level variables from various sources were used, and these are detailed in the respective sub-sections. Predictor variables used in the regression analysis literature (Balota et al., 2004; Cortese & Khanna, 2007; Ellis & Monaghan, 2002; Pexman et al. 2002; Strain et al., 2002; Yap & Balota, 2009) were included, along with other variables made available, such as those from the MRC Psycholinguistic database (Wilson, 1988) and ELexicon (Balota et al., 2007). If multiple versions of a predictor variable were available, such as multiple imageability ratings, then the norm that provided ratings for the most number of words (the highest number of completed cells) was chosen for use in the analysis. Highly correlated variables resulted in one representative of this group being chosen for the analysis. For example, Coltheart's N is strongly and significantly correlated with both orthographic and phonological neighbourhoods, and Coltheart's N was used in the analyses. The individual predictor level variables are described in detail below in the order that they were entered into the hierarchical regression analyses.

8.2.2.1 Surface level predictors

The initial phoneme of a word significantly impacts on the reading reaction time recorded for that particular item (Rastle & Davis, 2002). The ELexicon word reading reaction times, the dependent variable of these regression analyses, were collected by Balota et al. (2007) using a voice-key. This electronic voice-activated key registered when the vocal response reached a threshold. This threshold was met at different points for various initial letters depending on where this phoneme was articulated in the mouth. The initial phoneme of a word may also be a predictor of word reading times because of phonemic encoding differences between various letters (Balota et al., 2004; Rastle, Croot, Harrington, & Coltheart, 2005).

The initial phoneme for each word was dummy-coded as follows. Each word was coded dichotomously for the presence or absence of 13 phoneme-articulation features, as was the method used by Balota et al. 2004. These features were: voiced, stop, fricative, affricative, nasal, liquid, bilabial, labiodental, linguadental, alveolar, palatal, velar, and glottal. This set of known surface level predictors was simultaneously entered at the first level of hierarchical regression (Balota et al., 2004; Cortese & Khanna, 2007; Strain et al., 2002; Yap & Balota, 2009).

8.2.2.2 Lexical level predictors

The lexical level variables are known predictors of word reading time that are not semantic in their locus (Balota et al., 2004; Cortese & Khanna, 2007; Strain et al., 2002; Yap & Balota, 2009). Lexical variables measure orthographic or phonological information of the word. There were four in these analyses: length (in number of letters), frequency, familiarity, and Coltheart's N.

Length of the words was quantified in the number of letters within each word, and has previously been found as a significant predictor of word reading times (Balota et al., 2004; Cortese & Khanna, 2007; Ellis & Monaghan, 2002; Strain et al., 2002; Yap & Balota, 2009).

Objective frequency of a word was included in the database using Celex written frequency (Baayen et al., 1993). Kucera and Francis (1967) written frequency was an alternate option for this measure. These two frequencies were significantly correlated at .93 ($p < .01$). Since these measure are highly correlated with one another, only one, the Celex measure, was used in the analysis as it provided 100% of the words with frequency values; in contrast, only 78% of words were provided with frequency values when using Kucera and Francis measures[38].

Familiarity ratings were provided by McRae et al. (2005) as they provided the highest number of words (97.9%) with ratings as compared to the other sources. Other available familiarity ratings investigated were the Bristol norms (Stadthagen-Gonzalez & Davis,

---

[38] Analyses were performed separately using both frequency measures. Results from the analyses using Kucera and Francis frequency measure did not differ from the results of the analyses presented in this chapter.

2006), which only provided ratings for 17.9% of words, and the MRC Psycholinguistic database (Wilson, 1988), which provided information for 73.8% of the words. The McRae ratings and the MRC ratings were positively and significantly correlated at .75 ($p < .01$).

Coltheart's N is a known predictor of word reading times (Spieler & Balota, 2000). Coltheart's N, as provided by the N-watch programme (Davis, 2005), is a measure of orthographic neighbourhood and is correlated with phonological neighbourhood. Taking a given word, Coltheart's N is a count of the number of alternate words of the same length that are formed by only changing one letter of the original word (Coltheart, Davelaar, Jonasson, & Besner, 1977). For example, as provided by Davis (2005), "calm" has 13 orthographic neighbours, such as "palm", "call", and "calf". Phonological neighbourhood is the number of alternate words that can be made by changing one phonological sound of the original word, e.g., "bull" and "shawl" are both phonological neighbours of "ball". Coltheart's N is significantly correlated with phonological neighbourhood at .76 and with orthographic neighbourhood at .97, as provided by the ELexicon database ($p < .01$) (Balota et al., 2007). Coltheart's N is included in the analyses and is representative of the variance that may be attributed to neighbourhood measures.

8.2.2.3 Semantic level predictors, including age-of-acquisition

The third set of variables entered into this regression was semantic level predictors of which there were three: the known semantic measures of semantic features and imageability, and the measure age-of-acquisition. Of interest is whether this set has any predictive power when surface and lexical measures have been accounted for.

8.2.2.3.1 Imageability ratings

The first semantic measure was imageability, which is a known predictor of word reading time within regression (Balota et al., 2004; Strain & Herdman, 1999; Strain et al., 2002). Imageability is also used as a measure of semantic information within factorial designs, as discussed in Sections 2.3.1 and 2.3.2 (Strain et al., 1995; Woollams, 2005), and imageability may have a relationship with the other known semantic measure of these analyses, that of semantic features (Strain et al., 1995).

Previously published regression analyses that included imageability used only one syllable words, as until recently research focused on monosyllabic words (Balota et al., 2004; Cortese & Khanna, 2007; Strain & Herdman, 1999; Strain et al.,2002). Also, the majority of available imageability ratings are only for one-syllable words, e.g., the imageability ratings of Cortese and Fugett (2004). The current analyses included two syllable words as well.

In the current analyses, imageability ratings were provided by the MRC Psycholinguistic Database (Wilson, 1988), as this provided the most number of words with ratings (71%), and also provided a measure of imageability that is indicative of the other imageability ratings. The other sources of imageability ratings did not provide ratings for a majority of the words. Cortese and Fugett (2004) provide imageability measures for the majority of one-syllable words, but this only provided ratings for 48% of words. Less than 40% of words were provided with ratings when using either the Bristol norms (Stadthagen-Gonzalez & Davis, 2006), i.e., 17%, the ratings of Bird et al (2001), i.e., 10%, or the ratings of Morrison et al. (1997), i.e., 36%.

The regression analyses of Balota et al. (2004), which included six semantic measures, found that the imageability ratings of Cortese and Fugett (2004) were the only significant predictor of reading times in both young and old adults. However, the ratings of Cortese and Fugett did not provide a sufficient number of words in the current analyses with imageability ratings. The MRC ratings used in these analyses are highly, significantly correlated with the ratings of Cortese and Fugett, i.e., .75 ($p < .01$). The MRC ratings were also highly, i.e. above .80, and significantly, $p < .01$, correlated with the other imageability ratings, i.e. the Bristol norms (Stadthagen et al., 2006), the Bird et al. (2001) ratings, and the Morrison et al. (1997) ratings.

8.2.2.3.2 McRae et al. (2005) semantic features

Of special interest in these analyses is whether McRae et al. (2005) semantic features are a significant and unique predictor of word reading times, after other known variables have been accounted for.

McRae et al. (2005) present several measures of the number of semantic features, including number of unique features, and number of semantic features with and without taxonomic features. Number of semantic features including taxonomic features was chosen for these analyses. It may best capture the relationship between the related prime-target pairs in the semantic priming experiments of this thesis because prime-target pairs may not only share similar conceptual features, but also share super-ordinate category features as well.

8.2.2.3.3 Age-of-acquisition ratings

Age-of-acquisition ratings are available from a number of sources, e.g. Bird et al. (2001), Cortese & Fugett (2004), Morrison et al. (1997); Stadthagen-Gonzalez and Davis (2006); MRC Psycholoinguistic Database (Wilson, 1988). The ratings of Bird et al., Cortese and Fugett, Stadthagen-Gonzalez and Davis –the Bristol norms-, Morrison et al., and the MRC Psycholinguistic Database were all investigated as possible sources of age-of-acquisition ratings for the present investigation. None of them individually provided over 50% of words with ratings with Cortese and Fugett providing the most ratings, i.e., 46% (the majority of the one-syllable words). Less than 40% of words were provided with ratings when using Morrison et al., i.e., 36%, the MRC Psycholinguistic Database, i.e., 30.2%. the Bristol norms (Stadthagen-Gonzalez & Davis, 2006), i.e., 17%, or the ratings of Bird et al., i.e., 17%. These five ratings were strongly (most ranged between .98 and .76, though two of the correlations with the Bird et al. ratings were lowly, .5), and significantly ($p < .05$) correlated with each other.

When creating the ELexicon database, Balota et al. (2007) combined reaction times and ratings from various data collection sessions with various samples. Using this as an example, the age-of-acquisition ratings were combined from these five sources, i.e. Bird et al. (2001), Cortese & Fugett (2004), Morrison (1997); Stadthagen-Gonzalez and Davis (2006); MRC Psycholoinguistic Database (Wilson, 1988), in the same way. This provided a sufficient number of words with ratings; 77% of words were provided with ratings. Importantly, the age-of-acquisition ratings of the named sources were all collected using the same standardised method and rating scale as introduced by Gilhooly and Logie (1980) from random samples of similar populations on random samples of words. Following Balota et al.'s (2007) procedure that was used when combining reaction times and scores for the ELexicon project, each age-of-acquisition rating was transformed into a standardised Z-score. The average of the Z-scores from the five sources was used as the age-of-acquisition ratings in the analyses[39],[40]. As would be expected, the age of acquisition rating used in these analyses are strongly (all greater than .92), significantly ($p < .01$) correlated with the ratings from five sources, i.e. Bird et al., Cortese & Fugett , Morrison et al., Stadthagen-Gonzalez and Davis, and MRC Psycholinguistic Database.

8.3. Results

To determine whether semantic level variables can account for a significant and unique proportion of variance in word reading times, hierarchical multiple regression analyses of word mean reaction times were performed. Following the analyses of previous investigations(Balota et al., 2004 ; Cortese & Khanna, 2007), in addition to known surface and lexical level predictors, which were entered as block 1 and block 2 of the analyses respectively, semantic variables and age-of-acquisition were also entered as block 3 of the analyses. Each of the regression analyses used three blocks; the 13 phonological onset variables were simultaneously entered as block 1; the four lexical level variables, i.e., length, objective frequency, familiarity, and Coltheart's N, were simultaneously entered as block 2, and finally the third block of semantic features and imageability, and age-of-acquisition[41] were entered. These three blocks of predictor

[39] The average z-score was only taken if there was more than one age-of-acquisition rating for that word.

[40] Alternate analyses were also performed with Z-scores. Instead of averaging Z-scores, scores from certain sets provided the score for the stimulus words. For all one-syllable words the Z-scores of the Cortese and Fugett (2004) ratings were used. For two syllable words, Z-scores were provided by the other sources. When two ratings were available for a two-syllable word, the MRC rating was used over others. These analyses did not differ from the analyses presented in this chapter that used average Z-scores.

[41] The specific variables entered in block 3 were determined by the analysis being performed, as is explained in the subsequent paragraph. Phonemic and lexical level variables were entered identically in all analyses.

variables were regressed onto the criterion variable of ELexicon mean word reaction times.

Four regression analyses are presented. The first analysis included only semantic features as block 3. This was carried out in order to determine whether semantic feature measures were a significant and unique predictor of word reading time, once other known surface and lexical predictors had been accounted for. The second included only imageability as block 3. This analysis was performed in order to determine whether imageability measures were a unique significant predictor of word reading time, once surface and lexical predictors had been accounted for. This analysis was also performed as previous analyses have found imageability to be a significant predictor of word reading times (Balota at al. 2004; Strain et al. 2002). The third analysis included semantic features and imageability together as block 3. Of interest is whether McRae's semantic features can uniquely predict variance that is not accounted for by imageability. Finally, the fourth and final analysis included semantic features, imageability, and age-of-acquisition as block 3. Of interest is whether imageability and semantic features are unique and significant predictors of word reading times when age-of-acquisition is also included in the analysis.

Words with missing data were deleted list-wise in the analyses, as recommended by Howell (1992). There were a total of 427 words prepared for use in these analyses. As can be seen in Table 8.1, not all sources provided 100% of the words with measures. When a measure with less than 100% was used in an analysis, any word with a missing value on any of the measures was omitted from the analysis. Therefore, the various analyses contain different words depending on which measure was being used in the analysis; this is reflected in the degrees of freedom, and the specific number of words used in each analysis is reported in the respective section.

Continuous variables were mean centred by calculating the mean of each variable and subtracting it from every individual score for that particular variable. Continuous variables were centred with a view to performing interaction analyses, such as imageability by frequency. Mean centering does not affect the correlations between predictor variables. It does reduce the possibility of multicollinearity with interaction terms, such as those mentioned in the previous sentence, which can be caused by differences in the predictors' scales of measurement by centring all scales around zero (Tabachnick & Fidell, 2007, p.158). Continuous variables were also centred as to be in line with the other standard Z-Scores (standard deviation centred score) of the ELexicon word reaction times and age-of-acquisition scores, as the data was collected from more than one population (Balota et al., 2004).

The following analyses met the assumptions of regression analysis (Field, 2000). Collinearity statistics were examined to ensure predictor variable were not strongly correlated. Tolerance values indicate whether predictors are highly correlated with each other (Howell, 1992, p. 512-513). Tolerance values range from zero to one with a value of zero being less than desirable and indicating a perfect correlation (Howell, 1992, p. 512-513). For all regression analyses reported in this Chapter, the tolerance value ranged from 0.40 (minimum) to 0.97 (maximum) with an average tolerance value of 0.67. These higher tolerance values are in keeping with those suggested by Howell (Howell, 1992, p. 512-513) and well above the minimum of 0.20 as advocated by Field (2000, p. 153) and indicate that there is not high correlation between predictor variables. Conversely, lower VIF values are desirable when examining the data for multicollinearity, as this indicates lower standard error and more reliable regression coefficients. Howell (1992, p 513) advocates "lower" VIF values and Field (2000, p. 153) advocates that no single VIF value be above 10. VIF values for all regression analyses reported in this Chapter ranged between 1.04 (minimum) and 2.48 (maximum) with an average of 1.59. Therefore the VIF values do not indicate multicollinearity amoung the predictor variables (Howell, 1992, p 513; Field, 2000, p. 153). Predictors and the dependent variable were normally distributed, as would be expected when using centred values. The data was also examined to ensure there were not outliers present in the data. Mahalanobis' distance was examined, and was higher than recommended (Field, 2000). This analysis does, however, have a high number of predictors (a maximum of 20) and a large sample (427 words with reaction times) therefore these values may mean there is not a violation of an assumption. Cook's distance was, therefore, also examined. All values were less than one, indicating that one score is not overly influencing the data. In line with Tabachnick and Fidell (2007), the ratio of cases to the number of predictor variables was acceptable. The maximum number of predictors used in the analyses was 20 and the minimum was 18. If looking at individual predictors, then a minimum of 124 words would be needed; if looking at correlations and the regression model, then a minimum of 210 words would be needed. In the analyses of this chapter, the minimum number of words used for an analysis was 244 (Section 8.3.5).

8.3.1 Correlations

Correlations are presented in Table 8.2. Correlations between most measures were weak, around zero or 0.1, with only a few exceptions. ELexicon mean word reaction times are weakly and significantly correlated with several surface level variables and lexical level variables, at less than 0.30. Word reading times are also weakly, significantly, and negatively correlated with semantic features, at -0.13, and weakly, significantly, and positively correlated with age-of-acquisition, at 0.14. Imageability is not significantly correlated with the dependent variable. This lack of significant correlation between imageability and reaction times, therefore, is not unique to the current analyses. Balota and colleagues (2004) report a similar finding using Tolia and Battig (1978) imageability, which partly make up the MRC imageability ratings that are used in the current analyses. This lack of original bivariate correlation is explored further in the discussion. The three block 3 variables, imageability, semantic features, and age-of-acquisition, are significantly and weakly correlated with each other, at less than 0.33. Imageability and semantic features are positively correlated with each other, i.e., the higher the imageability rating the greater number of semantic features a word has, and negatively correlated with age-of-acquisition, i.e., the higher the imageability rating and the greater the number of semantic features, the lower (younger) the age at which the word was acquired.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
(1) ELexicon Z RT	--	.164**	-.129**	.204**	0.076	-0.063	-0.079	-.160**	0.022	0.058	.154**	-0.068	-0.017	-0.045	.243**	.198**	-.199**	-.265**	.137*	-0.029	-.128**
(2) Initial Phoneme Voiced		--	-.345**	.188**	-0.070	-0.079	-0.061	-.145**	-0.074	-0.016	0.083	.379**	-.145**	-0.066	.224**	-.108*	-.113*	-.198**	0.008	-0.041	-0.040
(3) IP Manner Stop			--	-.204**	-0.019	-0.021	-0.016	-0.066	-0.020	-0.004	0.038	.137**	-0.038	-0.017	-.154**	0.053	0.037	.138**	0.007	-0.029	0.031
(4) IP Manner Fricative				--	-.116*	-.132**	-.102*	-.362**	.396**	0.087	.236**	0.072	-.239**	.353**	.095*	-0.040	-0.073	-0.070	0.068	-0.011	-0.016
(5) IP Manner Affricative					--	-0.050	-0.038	-.136**	-0.046	-0.010	.309**	-0.066	-0.090	-0.041	0.027	0.025	0.030	-0.059	0.036	0.075	0.028
(6) IP Manner Nasal						--	-0.044	.252**	-0.052	-0.011	-0.049	-0.075	-.102*	-0.047	0.046	-0.055	0.020	-0.026	.123*	-0.034	-0.024
(7) IP Manner Liquid							--	.225**	-0.040	-0.009	-.124*	0.033	-0.079	-0.036	-0.058	-0.014	-0.078	-0.011	0.033	-0.015	-0.065
(8) IP Place Bilabial								--	-.143**	-0.031	-.439**	-.205**	-.279**	-.128**	-0.036	0.007	0.074	0.094	-0.014	-0.005	-0.044
(9) IP Place Labiodental									--	-0.011	-.149**	-0.070	-.095*	-0.043	-0.050	-0.048	-0.009	0.009	0.106	0.006	0.067
(10) IP Place Linguadental										--	-0.033	-0.015	-0.021	-0.009	0.051	-0.020	-0.048	-0.040	0.068	-0.023	-0.034
(11) IP Place Alveolar											--	-.213**	-.290**	-.133**	.120*	-0.011	0.060	-.121*	-0.053	0.048	0.007
(12) IP Place Palatal												--	-.135**	-0.062	-0.025	-0.029	-0.049	0.010	-0.022	-0.027	-0.013
(13) IP Place Velar													--	-0.084	0.002	-0.016	0.009	0.068	-0.031	0.002	0.052
(14) IP Place Glottal														--	-0.021	.108*	-0.085	0.059	0.018	-0.034	-0.036
(15) Length															--	-.203**	-.106*	-.693**	.185**	0.014	-0.046
(16) Frequency (CELEX)																--	.313**	.184**	-.374**	0.104	0.070
(17) Familiarity																	--	.170**	-.542**	.207**	.143**
(18) Colheart's N																		--	-.290**	-0.017	0.040
(19) Age-of-Acquisition (Z)																			--	-.322**	-.286**
(20) Imageability																				--	.307**
(21) Semantic Features (McRae et al., 2005)																					--

Table 8.2 Correlation matrix for the dependant measures and predictor variable from the regression analyses of Section 8.3.5, N = 243. IP = Initial Phoneme.
** = Correlation is significant at the p = 0.01 level (2-tailed); *= Correlation is significant at the *p*= 0.05 level (2-tailed).

8.3.2 Regression analysis with semantic features

Predictor Variable	β	
Block 1: Surface Predictors R^2adj		0.24***
voiced	-0.03	
stop	-0.04	
fricative	0.42	***
affricative	0.14	***
nasal	-0.05	
liquid	-0.09	*
bilabial	-0.05	
labiodental	-0.13	***
linguadental	0.01	
palatal	-0.14	**
velar	0.04	
glottal	-0.25	***
Block 2: Lexical Predictors R^2adj		0.50***
length	0.28	***
freqeuncy	-0.03	
familiarity	-0.29	***
Colthearts N	-0.11	*
Block 3: Semantic Predictors R^2adj		0.51***
Semantic Features	-0.13	***

Table 8.3. Standardised regression coefficients and R^2_{adj} from blocks 1, 2 and 3 of the regression analysis with only semantic features in block 3. * = p < .05, ** = p < .01, *** = p < .001.

Semantic feature measures alone were entered as block 3 after known surface and lexical predictors in blocks 1 and 2 in this first regression analysis. Standardised beta coefficients and R^2 for blocks 1, 2 and 3 are reported in Table 8.3. The overall model, at block 3, was significant, $F(17, 377) = 25.56$, $p < .001$ and accounted for 51% of the variance in reading times, $R^2_{adj} = .51$. Blocks 1, 2, and 3 within the model were significant. Semantic features, which composed block 3, accounted for a small and significant portion, i.e., $sr^2 = R^2_{change=}$ 1.6%, of variance found in word reading times. The negative significant correlation of semantic features with ELexicon word reading times indicates that words with a higher number of semantic features are read more quickly than words low in semantic features. Of note, is that though small, this contribution was significant, $t(377) = 3.57$, $p < .001$.

8.3.3 Regression analysis with imageability

Predictor Variable	β	
Block 1: Surface Predictors R^2adj		0.26***
voiced	-0.08	
stop	-0.12	*
fricative	0.44	***
affricative	0.11	*
nasal	-0.07	
liquid	-0.13	**
bilabial	-0.02	
labiodental	-0.17	***
linguadental	0.02	
palatal	-0.18	***
velar	-0.01	
glottal	-0.33	***
Block 2: Lexical Predictors R^2adj		0.52***
length	0.27	***
frequency	-0.02	
familiarity	-0.32	***
Colthearts N	-0.12	
Block 3: Semantic Predictors R^2adj		0.53***
Imageability	-0.12	*

Table 8.4. Standardised regression coefficients and R^2_{adj} from blocks 1, 2 and 3 of the regression analysis with only imageability in block 3. * = p < .05, ** = p < .01, *** = p < .001.

Imageability alone was entered as block 3 after known surface and lexical predictors in blocks 1 and 2 in this regression analysis. This analysis used a slightly different set of words than previous analyses, as can be seen in the difference in degrees of freedom. Standardised beta coefficients and R^2 for blocks 1, 2, and 3 are reported in Table 8.4. The model was significant, $F(17, 262) = 19.356$, $p < .001$ and accounted for 53% of the variance in reading times, $R^2_{adj} = .53$. Blocks 1, 2, and 3 within the model were significant. Imageability, which composed block 3, accounted for a small and significant portion, i.e., $sr^2 = R^2_{change=}$ 1.1%, of unique variance found in reading times. The Beta value indicates that words higher in imageability are read more quickly. Of note, is that though small, this contribution was a significant contributor to the model, $t(262) = 2.53, p = .01$.

8.3.4 Regression analysis with imageability and semantic features

Predictor Variable	β	
Block 1: Surface Predictors R^2adj		0.26***
voiced	-0.08	
stop	-0.12	*
fricative	0.44	***
affricative	0.11	*
nasal	-0.07	
liquid	-0.14	**
bilabial	-0.02	
labiodental	-0.16	***
linguadental	0.02	
palatal	-0.17	***
velar	0.00	
glottal	-0.33	***
Block 2: Lexical Predictors R^2adj		0.52***
length	0.26	***
frequency	-0.02	
familiarity	-0.31	***
Colthearts N	-0.13	*
Block 3: Semantic Predictors R^2adj		0.53***
Imageability	-0.09	*
Semantic Features	-0.07	

Table 8.5. Standardised regression coefficients and R^2_{adj} from blocks 1, 2 and 3 of the regression analysis with imageability and semantic features in block 3. * = p < .05, ** = p < .01, *** = p < .001.

Thus far, regression analyses have produced results that indicate imageability and semantic features when entered on their own, after surface and lexical level variables, are significant predictors of word reading time. Also of interest is whether McRae's semantic features can uniquely predict variance that is not accounted for by imageability, especially since semantic features are significantly correlated with ELexicon reaction times, and imageability is not (Table 8.2). In this third analysis, both semantic variables are simultaneously force-entered as the third block of a regression analysis. This analysis used slightly different items (280 words) to the analysis in 8.3.2, but the same items as those in the analysis presented in 8.3.3, due to the availability of ratings.

Standardised beta coefficients and R^2 for blocks 1, 2, and 3 are reported in Table 8.5. The model was significant, $F(18, 261) = 18.50$, $p < .001$ and accounted for 53% of the variance in reading times, $R^2_{adj} = .53$. Blocks 1, 2, and 3 within the model were significant. The two semantic factors in block 3 accounted for a small and significant portion, i.e., $sr^2 = R^2_{change=}$ 1.5%, of unique variance found in reading times (Table 8.5). Within this block only imageability was a significant contributor to the model, $t(261) = 2.03$, $p = .04$. Semantic features failed to contribute significantly to this model, $t(261) = 1.51$, $p = .13$.

8.3.5 Regression analysis imageability, semantic features, and age-of-acquisition

Predictor Variable	β	
Block 1: Surface Predictors R^2adj		0.30***
voiced	-0.07	
stop	-0.09	
fricative	0.51 ***	
affricative	0.12	
nasal	-0.05	
liquid	-0.10	
bilabial	-0.03	
labiodental	-0.21 ***	
linguadental	0.02	
palatal	-0.19 ***	
velar	0.00	
glottal	-0.34 ***	
Block 2: Lexical Predictors R^2adj		0.53***
length	0.24 ***	
frequency	0.03	
familiarity	-0.24 ***	
Colthearts N	-0.10	
Block 3: Semantic Predictors R^2adj		0.56***
Age-of-Acquisition	0.15 *	
Imageability	-0.09 *	
Semantic Features	-0.02	

Table 8.6. Standardised regression coefficients and R^2_{adj} from blocks 1, 2 and 3 of the regression analysis with imageability, semantic features, and age-of-acquisition in block 3. * = p < .05, ** = p < .01, *** = p < .001.

Though the regression analyses have produced results indicating a significant contribution from semantic variables imageability and semantic features, age-of-acquisition has yet to be accounted for. In this analysis, age-of-acquisition and both semantic variables were simultaneously force-entered in the third block of a regression analysis. This analysis used slightly different items (244 words) to the previous analyses.

Standardised beta coefficients and R^2 for blocks 1, 2, and 3 are reported in Table 8.6. The model was significant, $F(19,224) = 16.93$, $p < .001$ and accounted for 56% of the variance in reading times, $R^2_{adj} = .56$. Blocks 1, 2, and 3 within the model were significant. The semantic level factors accounted for a small and significant portion, i.e., $sr^2 = R^2_{change}$= 2.9%, of unique variance found in reading times. Within block 3, age-of-acquisition was significant contributor to the model, $t(224) = 2.42$, $p = .02$. Imageability within this block was also a significant contributor to the model, $t(224) = 1.99$, $p = .048$. Semantic features failed to significantly contribute to this model, $t(224) = .39$, $p = .70$. As detailed at the beginning of Section 8.3, in the paragraph directly prior to section 8.3.1, the minimum number of words needed for examining the regression model with 20 predictors is 210, and there were 244 words used in this analysis (Tabachnick & Fidell, 2007, p 123). Therefore, though this analysis included fewer items than the previous analyses of this Chapter, the results should still be robust as an adequate number of stimuli were included in the analysis.

8.4 Discussion

The impact of semantic factors on ELexicon word reading times was investigated using a large set of words and regression analyses. The final model of the current analyses accounted for 56% of the variance in word reading times, which is a slightly larger proportion (6%) of variance than was accounted for by the regression analyses of Balota et al. (2004). After phonemic and lexical factors were statistically controlled for in the first two blocks of each regression analysis, block 3- the semantic factors-, which included semantic measures semantic features and/or imageability, was a significant predictor of word reading times in all analyses. Additionally once age-of-acquisition was included imageability was still a significant predictor. Block 3 with semantic features, imageability, and age-of-acquisition together accounted for a unique and significant 3% of the variance in word reading times.

8.4.1 Surface and lexical level effects

Each of the current regression analyses included known surface and lexical predictors in the first two blocks (Balota et al. 2004; Cortese & Khanna, 2007; Ellis & Monaghan, 2002; Strain et al., 2002; Yap & Balota, 2009). As would be expected with word reading, which involves the computation orthography-to-phonology, the current analyses, using only a subset of ELexicon words and reading times, indicated that measures of surface level features and measures of orthographic and phonological properties of words, as captured by the lexical variables, have a significant relationship with word reading times. Surface level variables of initial phonemes significantly influence the reaction times, accounting for 30% of the variance. This suggests that manner and placement of articulation is crucial when using a voice-key microphone, as would be expected (Section 8.2.2.1) (Balota et al., 2004; Rastle & Davis 2002, Rastle et al., 2005). In the item-level regression analyses of Balota et al. that included imageability as a semantic level predictor and used reaction times of young participants (most similar to the final analysis of the Chapter, Section 8.3.5, Table 8.6), surface level variables account for 35% (in the analysis with Cortese and Fugett's imageability) and 39% (in the analyses with Nelson's imageability) of the variance. Therefore, Balota and colleagues found that surface level variables predicted slightly more variance than in the analyses of this Chapter. The difference, however, in the current results and those of Balota et al. was only small, less than 10%, and likely due to the differences in the word sample used in the analyses.

After accounting for the surface level measures, lexical level variables that measure orthographic and phonological properties of words were entered into the analyses of this Chapter. They accounted for an additional 23% of unique variance in word reading times. The lexical level predictors, themselves, alone, in the current analyses account for a higher proportion of variance than those of Balota et al. In the regression analyses of Balota et al., lexical level variables account for an additional 15% (in the analysis with Cortese and Fugett's imageability) and 10% (in the analyses with Nelson's imageability) of the variance. However, in conjunction with surface level variables, a total of 49.5% (in the analysis with Cortese and Fugett's imageability) and 49.6% (in the analyses with Nelson's imageability) of variance was predicted in the regression analyses of Balota et al., which is very similar to the 53% of variance predicted by the surface and lexical level variables in the current analyses. Though the precise proportions of variance predicated by surface and lexical level factors are not identical between the analyses of this Chapter and those of Balota and colleagues, the difference in results between the two sets of analyses is small. As stated previously, any difference in results is likely due to the word sample used in the analysis; the final analysis of this Chapter used 244 words, whereas the analyses of Balota and colleagues used 2,342 (in the analysis with Cortese and Fugett's imageability) and 997 words (in the analyses with Nelson's imageability). Consistency between the published results of Balota et al. and those of the current Chapter, even with different word samples, demonstrates the reliability of the results presented in this Chapter.

There were two anomalous results in the lexical level results across the four regression analyses; they may be due to the specific set of words used within the various analyses. First, Coltheart's N was a significant predictor in each analysis with the exception of the final analysis that included age-of-acquisition. This could be due to the specific set of words used in the final analysis. Second, in the current analyses frequency is not a unique predictor of word reading times. In published regression analyses, frequency was a significant predictor of word readings times (Balota et al., 2004; Cortese & Khanna, 2007; Monaghan & Ellis, 2002; Strain & Herdman , 1999; Yap & Balota, 2009) [42]. Investigations from Balota et al. (2004) and Strain et al. (1995, 2002), though some results only approached significance, indicated that imageability effects might be present for low frequency words alone. A further regression analysis investigated the possibility of imageability by frequency interaction (Tabachnick & Fidell, 2007), but did not reveal an interaction of these factors.

The lack of frequency effect including the lack of any significant interaction, however, does not necessarily mean that frequency does not influence word reading times; it may indicate that the influence of frequency may be subsumed by other variables. Subjective familiarity was considered to be a likely candidate (Speiler &Balota, 2000), but further analyses did not confirm this[43].

The influence of surface and lexical level variables within the current analyses is clear, but it is the influence of semantic measures, semantic features and imageability, and age-of- acquisition, once these surface and lexical measures have been accounted for, that is important to the central aim of this thesis.

8.4.2 Semantic measures and age-of-acquisition

The main purpose of these regression analyses was to assess the relationship between semantic variables and age-of-acquisition, and word reading times. Of interest was

[42] Subsequent analyses were performed to ensure the current results were not due to inconsistencies in the frequency measure with previous analyses (Balota et al., 2004; Cortese & Khanna, 2007; Monaghan & Ellis, 2002). Regressions using the log of Celex values and mean centred Celex values did not differ from the current results in any meaningful way. Analyses using Kucera and Francis (1967) frequencies revealed the same non-significant results.

[43] A further regression analysis that included all predictor variables, with the exception of familiarity was performed. Frequency was not a significant predictor in this analysis.

whether semantic level variables (semantic features, imageability, and age-of-acquisition) significantly predicted unique proportions of variance in word reading times and whether specifically semantic features were a unique predictor of word reading times once other known predictors, including imageability and age-of-acquisition, had been accounted for.

The first two regression analyses identified semantic features and imageability in individual analyses as having a significant impact on word reading, when surface and lexical factors were entered in earlier blocks (Sections 8.3.2 and 8.3.3, respectively). Within the semantic features analysis, the model accounted for 51% of the variance in word reading times; semantic features contributed a small, yet significant, 1.3% to this model, which is consistent with the results of Pexman et al. (2002). Within the imageability analysis, the model accounted for 53% of the variance in word reading times with imageability accounting for 1.2% of unique variance within this model. Results from other analyses within the literature indicated that imageability might be a significant predictor in these types of analyses as well (Balota et al., 2004; Strain & Herdman, 1999; Strain et al., 2002).

Whether either semantic features or imageability are unique predictors of word reading times in the presence of one another and in presence of age-of-acquisition is of interest. When entered into the same analysis, imageability uniquely contributed to the variability of word reading times, whereas semantic features did not account for a unique proportion of the variance in reading times (Section 8.3.4). Previous regression analyses found that semantic features were a significant predictor of word reading times, but imageability was not included in those analyses (Pexman et al., 2002). Another semantic measure- that is, a connectivity measure based on Wordnet (Miller, 1990)- has previously accounted for unique portion of word reading time variance when imageability was also included in the analysis (Balota et al., 2004). This indicates that two semantic measures, as long as they are accounting for different types of information, can both be significant and unique predictors in a large analysis. However, the number of semantic features was not included in those analyses either (Balota et al., 2004).

In the current analyses, though semantic features were a significant predictor of word reading times when entered alone, they failed to contribute uniquely once imageability was also included. This result may indicate that these two semantic variables are not accessing distinct aspects of semantic information, but are instead measuring similar attributes. Imageability is a semantic measure (Plaut et al. 1996; Strain et al., 1995, 2002), and may partially, but not wholly, measure semantic features. If a word is high in imageability, then it may have a greater number of semantic features than words low in imageability; this is supported by the correlations. Imageability, however, may also account for other aspects of semantic knowledge as well.

Strain et al. (1995) describe ways in which highly imageable items are semantically rich and these are described here (see also Section 2.3.1). Highly imageable items may capture semantic features, but will also capture other aspects of semantic information as well. Imageability is highly, significantly, and positively correlated with concreteness (Stadthagen-Gonzalez & Davis, 2006). Highly imageability items may also have an inherent meaning in contrast to more context-dependent low imageable items. For example, the highly imageable “pear” brings to mind more information about itself, its meaning and its possible context than the lower imageable word “soften”. Moreover, highly imageable items may be connected to other areas of the brain that store information about how a “peach” feels; its weight and softness (Strain et al., 1995).

Within this definition, semantic features may be wholly subsumed within imageability, but imageability is not wholly described by semantic features. The results from the current behavioural analyses support this view.

When age-of-acquisition was entered into the analysis with imageability and semantic features, age-of-acquisition was a significant predictor of word reading times. Additionally, imageability remained a significant and unique predictor, while semantic features did not. The non-significant correlation of imageability and ELexicon reading times is discussed in Section 8.4.2.2. That imageability remains significant in the presence of age-of-acquisition, which contrasts with previously published regression investigations, within which age-of-acquisition and imageability did not account for unique proportions of the variance within a regression model (Cortese & Khanna, 2007; Monaghan & Ellis, 2002). The difference in the results of the current analyses and the previous analyses may be due to the set of words used. The words used in these analyses were one *and* two syllables in length, and though representatives of the whole spectrum of frequency were included, on average they were low in frequency, and, as they were selected to have McRae et al. (2005) semantic feature ratings, may also have been more concrete than words in previous analyses. It could be these, or unidentified, factors that resulted in both imageability and an age-of-acquisition effects being found within these analyses.

In the current regression, age-of acquisition explained a greater proportion of variance than imageability alone. Therefore, both age-of-acquisition and imageability were responsible for unique proportions of variance in word reading times. Though each measure's contribution was small, it was significant and unique, indicating that age-of-acquisition and imageability measure different aspects of words reading. Of importance is whether this result demonstrates a semantic contribution to word reading. Crucial to interpreting the results is whether age-of-acquisition might be semantic in nature and also imageability's lack of significant correlation with the criterion variable.

8.4.2.1 Age-of-acquisition locus

As mentioned previously, it is unclear as to exactly what information is captured by the measure age-of-acquisition. Earlier accounts suggest that age-of-acquisition does not measure semantic information (Brown & Watson, 1987; Ellis &Monaghan, 2002; Monaghan &Ellis, 2002; Morrison & Ellis, 1995; Zevin &Seidenberg, 2002), whereas more recent accounts allow for age-of-acquisition to be, at least partially, semantic in nature (Brysbaert et al., 2000; Brysbaert & Ghselinck, 2006; Hernandez & Li, 2007; Monaghan & Ellis, 2010; Lambon-Ralph & Ehsan, 2006; Steyvers &Tenenbaum, 2005; Zevin &Seidenberg, 2004). Similar to other known semantic measures, such as semantic features, if age-of-acquisition is semantic in nature, then this would mean that it measures semantic information in semantic memory, and which could effect the weights of semantic connections. Age-of-acquisition accounts are now discussed further.

Some accounts of age-of-acquisition indicate that it may be a non-semantic variable. Brown and Watson (1987) suggest that age-of-acquisition is a measure of the phonological form in the lexicon; this is known as the "phonological completeness hypothesis". The earlier a word is learned the more wholly represented the item is in the lexicon. Later learned words are not stored as whole forms, but are assembled when read, accounting for why later learned words often have slower reading times. However, data, using segmentation tasks or tasks other than reading, shows patterns of behaviour that make this theory the least viable (Hernandez & Li, 2007; Johnston & Barry, 2006).

Other hypotheses also suggest that age-of-acquisition does not measure semantic information, but instead captures qualities in the connections between orthography and phonology, interacting with spelling-to-sound regularity measures (Ellis &Monaghan, 2002; Monaghan &Ellis, 2002; Morrison & Ellis, 1995). Age-of-acquisition may account for many of the (non-semantic) word reading effects reported in the literature (Morrison & Ellis, 1995).

The "cumulative frequency account" of age-of-acquisition claims age-of-acquisition influences semantic and non-semantic tasks; age-of-acquisition, however, does not heavily influence adult word reading (Zevin &Seidenberg, 2002). In a later article, Zevin and Seidenberg (2004) claim that it may be impossible to disentangle age-of-acquisition from other measures. Within this account, effects attributed to age-of-acquisition are explained by "lifetime" frequency, not simply the age at which the word was first learnt[44]. In the latest account, age-of-acquisition could also be, in part, semantic in nature (Zevin &Seidenberg, 2004).

Furthermore, age-of-acquisition is also described as being semantic in nature in the "semantic locus hypothesis" (Brysbaert et al., 2000; Brysbaert & Ghselinck, 2006). This account claims age-of-acquisition effects are present in tasks that are not solely dependent on orthography-to-phonology, such as word reading, or in tasks that are solely phonological. Within this theory, age-of-acquisition affects connections between orthography and phonology, and also affects semantic organisation and connections of semantic memory to the other systems; the latter effects are emphasised by this account. A foundation, or framework, of semantic information is established by words when learnt early in life; later learnt semantic information is then connected to this early-

[44] For an alternative account of frequency in age-of-acquisition see Monaghan and Ellis, 2010.

acquired framework. Therefore, this account claims that age-of-acquisition is a measure of semantic memory representations.

Recent computational models attempt to account for age-of-acquisition effects in several research areas, not only those from word reading (Hernandez & Li, 2007; Lambon-Ralph & Ehsan, 2006; Steyvers &Tenenbaum, 2005). Connectionist simulations of Lambon-Ralph and Ehsan indicate that, as the semantic locus hypothesis theorises, age-of-acquisition may influence all connections, and that it especially affects semantic connections. As mentioned previously, it is proposed age-of-acquisition has a greater influence on routes with "arbitrary" mappings within this model (Lambon-Ralph & Ehsan, 2006, Monaghan & Ellis, 2010). For example, in English word reading, age-of-acquisition would affect phonology-to-semantic and orthography-to-semantic connections. It would have little effect on less arbitrary orthography-to-phonology connections, which are "quasi-regular" (Section 1.2.2) (Lambon-Ralph & Ehsan, 2006). Styvers and Tenenbaum create a semantic network that models the development of semantic memory. Within this model words acquired early in life have a greater number of semantic interconnections than later-learned words. Therefore, these recent models, in contrast to earlier research, suggest age-of-acquisition captures semantic information. Within this, a more recent, computational model of semantic memory, it is claimed that age-of-acquisition can be a measure of semantic information representations.

The current analyses indicated that age-of-acquisition significantly impacts on healthy adult word reading and that it may capture components of lexical and semantic measures. The data with this set of words from the current analyses indicated that age-of-acquisition can capture similar aspects of word reading to known lexical variables, such as Coltheart's N and frequency, as revealed in the significant correlations (Table 8.2).

Like the more recent theoretical and modelling accounts of age-of-acquisition, the current analyses also show that age-of-acquisition might not solely be a lexical measure; it may also be partially semantic in nature as well. Age-of-acquisition was significantly and negatively correlated with the semantic level variables imageability and semantic features. When age-of-acquisition was entered into the analyses (Section 8.3.5), imageability remained a significant predictor. When age-of-acquisition and semantic features alone were entered as block 3 predictors in an additional analysis [45], semantic features failed to be a significant predictor of word reading times once the usual surface and lexical level predictors had been accounted for. Age-of-acquisition, however, was a significant predictor. As a reminder, semantic features were significantly correlated with word reading reaction times in the original correlations and were a significant predictor of word reading times when entered on their own in block 3. In this additional analysis, however, age-of-acquisition subsumed the contribution of an explicitly semantic measure, that of semantic features.

Not only do recent accounts of age-of-acquisition, as reviewed previously in this section, indicate that age-of-acquisition may, in part, be semantic in nature (Lambon-Ralph & Ehsan, 2006; Monaghan & Ellis, 2010; Styvers & Tenenbaum, 2005), but results from the current regression analyses also indicate that age-of-acquisition may measure semantic information. Therefore, the age-of-acquisition effects found in the current analyses, as it was concluded that they are, in part, semantic, might also suggest a

[45] This analysis differed from previously reported analyses of this chapter in that imageability was not included in this analysis.

semantic contribution to word reading. However, age-of-acquisition is not the only significant block 3 predictor from these analyses, and there might be further evidence of a semantic contribution to word reading. The unique contribution of imageability to the model might also indicate that semantic information influences word-reading times.

8.4.2.2 Imageability effects

After other factors were accounted for, including the influence of age-of-acquisition, imageability significantly contributed to the regression model. In previous research, as demonstrated by the published debate of Strain and colleagues (1995; 2002), and Monaghan and Ellis (2002) and Ellis and Monaghan (2002), the individual effects of imageability and age-of-acquisition have been difficult to distinguish. However, in the current regression analyses, imageability contributed its own unique significant contribution to the regression model once age-of-acquisition had been accounted for. Though promising, this result is examined further here as the original bivariate correlation between imageability and ELexicon reading times was not significant. Imageability was a known predictor of word reading times in previous analyses (Balota et al., 2004; Strain & Herdman, 1999; Strain et al., 1995, 2002). However, in this and in other analyses (Balota et al, 2004; Strain et al., 2002), though imageability was a significant predictor in the final analyses, it was not significantly correlated with word reading reaction times in the original simple correlations (Balota et al., 2004; Strain & Herdman, 1999; Strain et al., 1995, 2002). Yet, when included with other measures, there was a significant relationship between imageability and reaction times. It is possible that this might be due to a relationship of imageability with another measure (Boniface, 1995).

In a suppression relationship, imageability, for example, would originally have a significant relationship with word reading reaction times, but with the addition of other predictor variables, this relationship would no longer be significant (Howell, 1992; Tabachnick & Fidell, 2007). Yet, the opposite occurred here. Imageability did not originally have a significant relationship with word reading reaction times; however, a significant relationship with word reading reaction times was present when other independent predictors, that are “interdependent” (p. 107) with imageability, are entered into the analysis (Boniface, 1995)[46]. This might suggest a synergistic relationship between imageability and another variable[47]. A relationship, between imageability and another “interdependent” measure could be responsible for the significant prediction of variability in word reading reaction times (Boniface, 1995).

A technique similar to one suggested for uncovering a suppressor variable was undertaken to investigate whether there was an identifiable synergistic variable working with imageability in order for it to be a significant predictor (Tabachnick & Fidell, 2007). When trying to understand a suppression relationship and trying to discover which variable is acting as a suppressor, it is suggested that independent variables individually are removed until the original significant relationship between the suppressed variable and dependent variable returns (Tabachnick & Fidell, 2007). Therefore, in the case of the suspected synergistic variable, the other independent variables were removed systematically from the analysis in order to return the relationship between imageability and reaction times to non-significant. However, no variable was uncovered. When all variables were removed, leaving only imageability as

[46] Synergistic relationships are often discussed in terms of Sum of squares values for F-tests, but as regression analyses are also reliant on sum of squares, the same principal could apply here.

[47] With thanks to Dr Peter Watson for help with this interpretation.

a predictor, then this regression model was not significant, and equalled the non-significant original correlation, as would be expected. No single variable was identified. Therefore, it could be due to an unidentified factor.

Previous sections in the present discussion identify imageability as being semantic in nature and as being partially composed of semantic features. However, imageability is also related to lexical variables. It was significantly correlated with them (Table 8.2), and also relied on them to predict some of the variance in word reading times. When these variables accounted for some of the variance, then imageability was a significant predictor as well. Imageability remained a unique predictor of word reading reaction times when age-of-acquisition was entered into the analysis, but it may be dependent on the relationships with other measures, including those not accounted for. However, other regression analyses interpreted significant imageability effect as indicative of a semantic contribution to word reading even when no original significant correlation was present and the same conclusion is drawn here. Imageability as a unique and significant predictor in these regression analyses provides evidence of a semantic contribution to word reading.

8.4.3 Conclusions

Of priority to the central aim of this thesis is whether within these analyses known semantic factors (semantic features and imageability) and age-of-acquisition significantly influenced word reading reactions times, and specifically whether McRae's semantic features was a unique predictor. Imageability and age-of-acquisition were unique, significant predictors of word reading times. Semantic features, however, only predicted word reading times when entered with lexical and surface level variables, and was not a significant predictor when imageability and age-of-acquisition were entered into the analysis. Therefore, the current regression, which used behavioural measures, suggests that semantic features are not measuring a unique aspect of semantic knowledge that is not measured by other predictors. Imageability and age-of-acquisition may also capture semantic feature information. Also, previous sections presented that these measures (imageability and age-of-acquisition), as suggested by recent published research, are semantic in nature.

These two semantic measures, i.e. imageability and age-of-acquisition, remain significant in the current regression analysis while length, familiarity and other factors are statistically controlled for. This may provide evidence of a semantic information contribution to orthography-to-phonology computation with the selection of words in the current analyses. The restrictions on the choice of words selected for these analyses may have limited the sample of stimuli. Stimulus words were from the frequency spectrum with 20% of words being high frequency, but the sample was low frequency. The restricted stimuli choice and frequency limitations may limit the external validity of the results. Therefore these results may support the view that a semantic contribution occurs with one and two syllable words, but this cannot perhaps be generalised to other words, though these conclusions may be most appropriate for low frequency words.

8.4.4 Subsequent ERP studies

The results of these regression analyses may provide evidence of a semantic contribution to word reading, specifically with the words used in the current analyses, which though included high frequency were on average low frequency, though no interaction of frequency was found. Furthermore, in the semantic priming experiments, reliable evidence of semantic priming was found for low frequency target word conditions (Chapter 7). The regression analyses also indicate that further investigations using imageability, age-of-acquisition, and semantic features may be of worth. The final investigations of this thesis, detailed in the following chapter, investigate the neurocorrelates of semantic features and imageability from the regressions' block 3 during low frequency (silent) word reading. These two experiments that use ERP measurements also have the potential to reveal differences between the two semantic measures that may have been obscured by behavioural measures. They may also reveal semantic effects of these measures early in the time course of word reading, meaning a semantic contribution prior to the completion of phonological processing is possible, as suggested in the previous investigations of this thesis.

Chapter 9

Two ERP Experiments

9.1 Introduction

The investigations of this thesis, thus far, have used behavioural measures, namely word reading reaction times of target words in priming paradigms and of single words, to investigate a semantic contribution to orthography-to-phonology computation in healthy adults. Results from the semantic priming experiments suggest significant priming in low frequency target word conditions (Experiments 1, 2, and 4, also see Chapter 7) indicating that low frequency target words may benefit from a related primes' remaining semantic activation. Priming of high frequency words was also found in an experiment with one intervening item (Experiment 2). Further investigations using regression analyses (Chapter 8) showed that semantic variables are significant predictors of word reading times for single words, even after accounting for other measures. This too suggests that semantic information might contribute to word reading with the set of words used in the regression analyses, which contained a spectrum of frequencies with a low frequency average. Therefore, the behavioural results from the investigations in Chapters 3 - 8 suggest that there may be a semantic contribution to orthography-to-phonology computation, and this was most reliable for low frequency words.

The two experiments presented in this chapter (Experiments 5 and 6) seek to uncover the time course of semantic measures (imageability and semantic features in Experiments 5 and 6, respectively) during silent word reading using low frequency words and the neurophysiological measure of even related potentials (ERPs), thus further investigating whether semantic information contributes to orthography-to-phonology computation. An advantage of using ERP measures is that it allows for time course observations of processes in the brain (Handy, 2005) (see Section 9.1.1.1). Within this chapter's experiments, evidence of early divergences in the ERP electrical signals for the semantic measure manipulations that occur prior to the completion of phonological processing would provide evidence that semantic information has the potential to contribute to the process of orthography-to-phonology computation.

Chapter 2 (Section 2.2.3) briefly introduced ERP methods and literature, but more details will be provided here as ERP data is used as the dependent variable in this Chapter's experiments. This introduction section: (a) highlights information about ERP as a measure, (b) reviews phonological effects and relevant early semantic effects from the ERP literature, which provide relevant time-windows and electrode sites of interest for this Chapter's analyses, (c) introduces the Experiments 5 and 6, and (d) concludes with predictions. Subsequent to this, methods, results, and discussion sections are presented separately for each ERP study, followed by a combined discussion.

9.1.1 Introducing Event-Related Potentials (ERPs)

9.1.1.1 ERPs as a measure of cortical processing

In the following sections, information about ERPs, including what is measured and how it is used as dependent measure, is presented. Electroencephalography (EEG) measures continuous electrical signals (voltage) of cortical neuronal cells (Coles & Rugg, 1996; Luck, 2005a, Rugg & Coles, 1996). Voltages are collected using electrode sensors, and this thesis focuses on research that uses electrodes at many, various locations (sites) on the scalp. Amplifiers and computers are then used to magnify and process the electrical signals. ERPs are created by precisely linking these continuous measured voltages (EEG data) to a specific external event, such as a stimulus in an experiment (Coles & Rugg, 1996; Handy, 2005; Luck, 2005a), resulting in an epoch of data for each stimulus of an experiment. How the measured electrical charge occurs in neurons is now described.

The measured electrical signal in ERPs is thought to originate from post-synaptic activity of pyramidal neuronal cells in the cortex of the brain (Coles & Rugg, 1996; Luck, 2005a). Signals are conveyed by neurons using chemical and electrical transmitters; transmitters in the brain are called neurotransmitters. The propagation of a signal occurs when a neurotransmitter from one neuron binds to the receptors of an adjacent neuron (Luck, 2005a). This causes ions[48] to inundate this adjacent neuron, creating an electrically charged signal within it, and simultaneously creating an oppositely charged extracellular space around the neuron and in the parts of the neuron that initially receive the signal (Coles & Rugg, 1996; Luck, 2005a). For example, when a neuron is signalled by a positively charged neurotransmitter, protons enter the neuron from both the

[48] Ions are positively or negatively charged molecules that have an unequal number of protons (+) and electrons (-).

adjacent neuron and the extracellular space, creating a positive charge in the neuron. This leaves the extracellular space and some other areas of the neuron with an excess of electrons and a negative charge. For this one neuron a dipole (adjacent positively and negatively charged areas separated by space) is created. Because the electrical charge of one neuron is very small, in order for an electrical charge to be detected at the scalp, a large number of neurons in the brain with this electrical charge must occur nearly simultaneously (and be positioned in the brain in such a way that the electrical signal is not cancelled out), and these positive and negative charges are the electrical signals measured in ERPs (Coles & Rugg, 1996; Barber & Kutas, 2007; Handy, 2005; Luck, 2005a). ERPs provide a measure of electrical voltages that occurs in the brain at a specific moment thereby providing an indicator of any cortical neural change in response to an experiment's stimulus, with precise timing, over the entire course of a stimulus (Barber & Kutas, 2007; Coles & Rugg, 1996; Rugg & Coles, 1996).

Though ERPs provide a good method for discovering the time-course of a process, such as word reading, they are not the best method for discovering the cortical location of processes (Handy, 2005; Luck, 2005a; Otten & Rugg, 2005). Using ERPs to measure the location of a particular process is difficult because a scalp distribution of the cortical electrical signals (topography) in a certain configuration could be the result of any number of patterns of electrical activations in the brain. Discovering the cortical sources of scalp-measured ERPs is difficult; this is known as the "inverse problem" (Luck, 2005a; Magnum & Hillyard, 1996; Slotnick, 2005).

9.1.1.2 ERPs as a dependent measure

Quantifiable aspects of the voltage data (measured in volts and depicted as continuous waveforms) will be used as dependent measures in the experiments of this chapter. Aspects of the data include the strength of the electrical voltage (average amplitude) and the time at which the differences in peak electrical voltage occur (peak latency) (Coles & Rugg, 1996; Handy, 2005). Similar to using reaction times as the dependent variable in a behavioural experiment, the electrical signals from an ERP experiment's trials are averaged across trials in a condition to create an average ERP measurement, depicted as a waveform, per experiment condition, per participant, and per electrode site. ERP data from specific time-windows (the time-span duration-of-interest in the data, such as 150ms-190ms) at the various scalp electrode sites is analysed (Rugg & Coles, 1996). If an ERP experiment's conditions are well controlled and electrical artefacts (extraneous electrical signals) have been eliminated, then any voltage difference found in the ERP data, such as average amplitude or peak latency, between the experiment's conditions is likely due to the experimental manipulation between conditions, reflecting differences in neural processes for the two experimental conditions (Luck, 2005a; Rugg & Coles, 1996).

Significant differences in average amplitude between experimental conditions may indicate that items in the experimental conditions engage a specific cortical process to different degrees (Handy, 2005; Rugg & Coles, 1996). Investigating average amplitude differences in various time-windows of the time-course can provide information as to when these differences occur. Peak latency differences may indicate differences in cortical processing efficiency. Analyses of average amplitude and peak latency can both reveal information about whether that items from the experimental conditions engage the same set of cortical processes at different points in time (Handy, 2005; Rugg & Coles, 1996). ERP methods offer the opportunity to investigate whether the electrical voltages of experimental conditions diverge from one another and if so, then when. For example whether or not the electrical voltages of experimental conditions differing in the number of semantic features diverge and whether this might occur early in the time course of word reading, before the settling of phonology, can be investigated (see Experiment 6). Of note here, are “components” of an ERP waveform, which can be difficult to define (Coles & Rugg, 1996; Handy 2005; Luck 2005a). Components are expected features, highs (peaks) or lows (troughs) in amplitude, of a waveform (based on previous extensive research) “thought to reflect a specific cognitive process” (Otten & Rugg, 2005, p. 5). These waveform features are present at a specific time-window and electrode site and are defined in relation to one another (Coles & Rugg, 1996) and in relation to the task being performed (Luck, 2005a, 2005b). Extensive inventories of components thought to reflect a particular neural process have been produced (see Coles & Rugg, 1996; Dien, 2009; Luck, 2005). For example, the N400 component is a negative deflection that peaks around 400 ms post-stimulus (and can occur in a time window of 300ms-500ms) at right hemisphere central-parietal electrode sites (Kutas & Hillyard, 1980; Luck, 2005). The N400 component is thought to accompany semantic processing and most likely reflects effects of context, predictability, and integration of total semantic information, and is not solely associated with stimuli of word modality (Barber & Kutas, 2007; Hinojosa et al., 2001; Kutas & Federmeier, 2000).

Components may not be ideal for every investigation using ERP methods. It is not always clear what particular neural activity a component reflects (Handy, 2005; Coles & Rugg, 1996); for example the P300 or P3 component is still not fully understood (Luck, 2005a). The hypothesis of a specific study may not correspond to a component (Luck, 2005a), and components may omit important early effects (Hauk & Pulvermuller, 2004; Pulvermuller

2001). When component-lead approaches are not appropriate, ERP methods can still be used to explore a hypothesis (Handy, 2005; Luck, 2005a). ERP methods can be applied by using *a priori* predictions established on appropriate literature, as the literature can provide time-windows and electrode sites that would be of interest to a new investigation of a similar nature. For example, the two ERP experiments of this chapter, which seeks to investigate the ERP correlates of semantic variables, will rely on the literature to identify appropriate time-windows and electrode sites to use when analysing the collected ERP data from this type of study. The following section reviews relevant ERP literature starting with phonological processing, and moving to early semantic effects.

9.1.2 Establishing a phonological processing marker for Experiments 5 and 6

Silent, covert word reading is the task used in Experiments 5 and 6 (Sections 9.2.1.3 and 9.3.1.3). Overt word reading, which was used in the previous investigations of this thesis (Chapters 3-6 and Chapter 8), may not be the most appropriate task for using with ERP measures, as moving the mouth to speak the word aloud would create electrical muscle artefacts that are larger in electrical voltage than the desired smaller cortical voltages, making the latter hard to detect and measure (Handy, 2005; Rugg & Coles, 1996). Silent reading offers the advantage of not producing a muscle voltage, as overt reading does. Additionally, as in overt word reading, silent reading might not bias orthography-to-phonology computation to include semantic memory or include decision processes as other button press tasks might (Section 2.2.2)[49]. Hauk & Pullvermuller (2004) in a review of ERP investigations demonstrate by listing ERP studies and the task used in each study that lexical decision and silent reading are used as tasks in

[49] Some ERP studies use button press tasks (Ziegler et al. 1997), including lexical decision task (Hauk, Davis, et al. 2006; Sereno et al. 1998) and semantic judgement tasks (Adbullev & Posner, 1998; Bentin et al. 1999; Zeigler et al. 1997). However neither of these was appropriate for the investigations of this thesis as semantic judgements may bias word reading to include semantic memory.

published ERP experiments. Other studies using ERP measures have also used silent reading as a task; for example, silent reading is used by Kutas & Hillyard (1980), Eulitz, Hauk, and Cohen (2000), Khateb et al. (1999) and Kutas and Hillyard (1984). Moreover, silent reading was used by Ellis, Burano, Izuri, Bromiley, and Venneri (2006), Michelli, Friston, and Price (2000); Pylkkanen and Marantz (2003), Yates, Friend, and Ploetz, (2007) to investigate orthography-to-phonology computation in neuroimaging (fMRI) studies. Because silent reading has been used in many published studies, including investigations that use ERP, it shows that it is a legitimate task to use, and is especially relevant to use in investigations where it is important for participants not to move, such as in ERP experiments.

Silent reading does have a disadvantage, that of determining when word reading has been completed, and therefore when phonological processing has finished. As a reminder, of importance is whether any evidence of semantic effects occurs prior to the completion of orthography-to-phonology processing in silent word reading, as this would be evidence of early semantic processing. Early semantic processing would mean it is therefore possible for semantic memory to contribute to word reading before the completion of orthography-to-phonology computation. Therefore a priority for the current studies is to identify in terms of ERP time course when phonology occurs in adult word reading.

In order to investigate whether semantic processing occurs early, prior to the completion of phonological processing, a marker of phonological processing is needed. To establish when phonology is processed in the English language[50], a brief review of the ERP literature that report when phonological processing occurs is presented here, thereby providing a literature-based marker of phonological processing to be used in Experiment 5 and 6. Markers of phonological processing include effects of phonological variables, such as phonological neighbourhood measures, and effects of phonological tasks, such as rhyming, or spelling-to-sound typicality effects.

A study of note that investigated the time course of frequency and regularity effects in word recognition can provide information about when phonological processing occurs (Sereno et al., 1998). Sereno et al., using a between-subjects design with healthy adult participants, recorded ERPs (with one group of participants) during a button press lexical decision task in which stimuli were either real words manipulated of frequency (high or low) and regularity (regular or exception) or non-words. In the eye-tracking study (with another group of participants) the same words were placed in at the end of a context sentence. In this study, fixation on a single word lasts for 300ms, after which, because eye movements move on the authors conclude that lexical word recognition is complete. In the ERP data a significant regularity effect, a difference in the peak amplitude for regular and exception words was seen in the P2 component- a positive peak in amplitude around 200ms post-stimulus at anterior and central electrode sites, after the N1 component- specifically at 168ms. Only participants with a regularity effect in their lexical decision response times showed this regularity effect in their ERP data. Because of this ERP effect at 168ms, Sereno et al. concluded that spelling-to-sound

[50] Though word reading processes and their interactions may be universal, the precise time when orthographic-to-phonological processing may vary as depending on whether the language has a deep orthography (Barber &Kutas, 2007, See Chapter 1). Therefore of specific interest to the current investigation is the time at which phonological processing occurs in the English language as this is the language used in the current experiments. Phonological processing, however, has also been studied in languages other than English. Proverbio et al. (2004) study these effects using the Italian language. Phonemic effects were found at 250 and 350 ms at occipito-temporal sites. Within the French language, effects of a phonetic rhyming task on ERP waves were found after 320 ms (Bentin, et al., 1999). Therefore within languages other than English phonological effects are seen after 250ms.

regularity effects in word recognition occur between 160ms and 180ms, thus providing a slightly earlier time-window of phonological processing, as compared to the remaining literature reviewed in brief here. The time-window of phonological processing might be earlier than all other reported phonological windows because the marker of phonology used in this word recognition (as opposed to word reading) study is regularity, and regularity might be capturing the process of orthography-to-phonology computation, including orthographic processes, not phonological processing alone. Therefore the remaining studies reveiwed in this section and the next section may provide a more appropriate phonological marker.

In a recent article, Dien (2009) reviews numerous ERP studies and outlines a framework, including timeline of word reading based on the average findings of the many other published findings. Within Dien's framework, the time-window in which phonological processing is likely to occur; this is 200ms-450ms. Because the literature he reviewed uses varying designs and methods, Dien suggests that identifying the exact moment of phonology processing within his model is difficult, hence a time-window that is 250ms in length. This review shows a slightly different time-window for phonological processing from Sereno et al. (1998) above, but the time-window of 200ms-450ms proposed by Dien is consistent with the study reviewed next and with the literature reviewed in the subsequent section.

Grainger et al. (2006) investigated the activation of orthographic and phonological factors using orthographic and phonological masked priming and ERP measures with healthy adult participants. They used a target word button press task of reporting when an animal name was seen. When target items where phonologically primed, Grainger et al. report priming effects in the ERP amplitude at the time-window of 250ms-350ms.

Further analyses of the ERP data reveal both orthography and phonology priming effects in the N250 component, with phonological priming effects occurring from 250ms-300ms, 50ms after orthography priming effects. Therefore, using phonological priming, Grainger et al. identified phonological processing as occurring *from* 250ms, slightly later that the start of phonological processing as suggested by Dien.

A phonological processing time-window of 200ms-500ms is inclusive of that proposed by Grainger et al. (2006) and Dien (2009), and endorsed by other studies (that are reviewed in the following section) (Ashby, Sanders, & Kingston 2009; Barber & Kutas, 2007; Landi & Perfetti, 2007). However, this time-window, slightly later than that proposed by Sereno et al. (1998), was chosen because, with the exception of Sereno et al., the literature identifies phonological processing as originating at 200ms. Therefore, 200ms will be used as a marker of the start of phonological processing in Experiments 5 and 6. Any difference in semantic factors prior to 200ms indicates early semantic activation which could potentially contribute to orthography-to-phonology computation. The possibility of early semantic effects is reviewed in the following section.

9.1.3 A review of early semantic effects within the ERP literature

Word reading and semantic processing has been investigated using ERP methods in the literature, as briefly referred to in Chapter 2, and this will now be reviewed in greater detail here. As mentioned in Section 9.1.1., semantic processing has been linked to the N400 component (Barber & Kutas, 2007; Hinojosa et al., 2001; Kutas & Hillyard, 1980; Kutas & Federmeier, 2000). However, any effect found at 400ms post-stimulus may occur after phonological processing has been completed because phonological processing has been suggested to originate at around 200ms (Section 9.1.2). Therefore

N400 effects might not suggest a possible semantic contribution to orthography-to-phonology computation. However, some studies have found early, i.e. prior to 200ms, semantic effects prior to the phonological processing marker identified above (Dien, 2009; Hauk, Davis, et al., 2006; Hinojosa et al., 2001; Hinojosa, Martin-Loeches, Munoz, Casado, & Pozo, 2004; Martin-Loeches, Hinojosa, Fernandez-Frias, & Rubia, 2001; Mari-Beffa, Valdes, Cullen, Catena, & Houghton, 2005; Pulvermuller, 2001), and therefore provide evidence that ERP methods can be used to detect early semantic effects, and these relevant studies are reviewed in this section. This literature review is also used to identify time-windows and electrode sites that will be relevant to the analyses of Experiments 5 and 6 (Section 9.1.4).

Hauk, Davis, et al. (2006) investigated whether various word measures, lexical and semantic, predicted the variance in ERP data. Healthy adults performed a button press word recognition task (i.e., lexical decision) and regression analyses were used to analyse the ERP data. First, a principal component analysis on word stimuli measures was performed to separate inter-correlated variables, so that factors that measure similar information were not entered in to the regression analysis and to identify factors relevant for the regression analysis. Four variables significantly predicted 80% of the variance in the ERP data, including the semantic variable of semantic coherence, first published by Ford, Marlsen-Wilson, and Davis (2003). Semantic coherence quantifies the similarity between the meanings for all forms of a word. An example offered in the article is “help”, which is high semantic coherence because “helper” and “helpful” are similar in meaning (Hauk, Davis, et al., 2006). By contrast, “depart” is low in semantic coherence, as other forms of the word, such as “department”, are not similar in meaning (Hauk, Davis, et al., 2006).

Because of the central aim of this thesis, of particular interest are the semantic effects found in this study. ERP amplitudes revealed found a significant effect of semantic coherence at 160ms at a left fronto-temporal electrode site (FT7), and right parietal electrode site (P2, not to be confused with the P2 component), with words higher in semantic coherence producing greater amplitudes than words lower in semantic coherence. This difference reoccurred at 314ms, and 500ms post-stimulus. Because of findings in the literature, Hauk, Davis et al. (2006) claim that word recognition is completed by 250ms. They conclude from the results of their study that because semantic coherence effects are present before this time, at 160ms, that semantic effects, as captured here by the factor semantic coherence, are involved in the word recognition process before the decision was made. Therefore, this study, which uses a lexical decision task, offers evidence of early semantic effects, and it will be of interest to investigate whether similar early semantic effects can be found with a task which emphasises reading.

Research by Abdullev and Posner (1998) used ERP measures during a delayed word reading task to investigate the differences between word reading and a semantic task (saying an item's use). Because this research used a word reading task, it may have more to say about finding an early semantic during a reading task (as opposed to the recognition tasks used in the previously reviewed literature). In the ERP data, Abdullev and Posner found semantic effects with the semantic task at early time-windows and electrode sites similar to the research of Hauk, Davis et al. (2006) - that is, the semantic task showed greater amplitudes than the word reading task at left inferior frontal electrodes at 170ms and at 220ms. Additionally around 200ms, greater negative amplitudes for the semantic task were shown in the electrodes on the left occipital-temporal area. Since there is a semantic effect in the ERP data prior to 200ms, this research demonstrates that early semantic effects can be found. However, Abdullev and Posner's effects were found in the semantic task, not in the word reading task. Therefore it is still of interest to investigate whether semantic effects can be found during a task that emphasizes reading.

In ERP investigations of comprehension skills in healthy adults, Landi and Perfetti (2007) also found indications of early semantic effects. In one condition, participants performed a phonological comparison task (judging if two sequentially presented words were pronounced in a similar way) and a semantic comparison task (judging if two sequentially presented words were similar in meaning). The P2 (150ms-250ms) and N400 (250ms-600ms) components of the ERP data were used as the dependant variable. Significant differences in the P2 component at frontal and central locations were found between related pairs (categorical and associated) and unrelated pairs; related pairs had more positive amplitudes than related pairs. The N400 component showed a similar effect as the P2 component, except the effect moved to posterior regions over time. Because an effect of relatedness was found at 150ms-250ms, and relatedness (judging if two sequentially presented words were similar in meaning) can capture semantic information, Landi and Perfetti concluded that semantic effects can be detected early using ERP measures. The results also confirm the previously highlighted phonological processing marker (Section 9.1.2.), with significant phonological effects in the P2 component. However, again, the task used in this study does not emphasise word reading.

Recently, Kounios et al. (2009) have investigated the ERP correlates of semantic features as measured by McRae et al. (2005), which is a semantic measure of interest to the current thesis (see Chapter 8). The healthy adult participants of Kounios et al.

judged, via button press, whether two words presented sequentially were semantically related (a semantic comparison task). The words were from one of three experimental conditions: concrete words high in number of semantic features, concrete words low in number of semantic features, and abstract words; the focus in this review is the two concrete word conditions that differed in number of features because this is most relevant to this thesis (see Chapter 8 and Section 9.3).

In ERP measures Kounios et al. (2009) revealed that average amplitude significantly differed (main effect of semantic features) between words high in number of semantic features and words low in number of semantic features in component P2 (time-window 200ms - 300ms) and a late time-window from 500-800ms with words high in number of features producing more positive amplitudes than words low in number of features. There was no significant interaction with hemisphere (though the number of features interaction with a dorsal/ventral division of electrode sites was significant). Because there was a difference in ERP signal between semantic feature conditions, the authors concluded that the amount of semantic information inherent in an item (as measured here by semantic features) can affect the amount of information activated in semantic memory, (i.e. it is possible that a word with a higher in semantic features will activate more information in semantic memory). There were no clear, significant peak latency effects in the ANOVA analyses, though t-tests revealed a significant difference between words high in semantic features and words low in semantic features in the P2 component, with words high in semantic features peaking later than words low in semantic features at a left parietal region only. Because there was no main effect of semantic features with latency measures and because of the timing of the small latency effect, Kounios et al. conclude that the number of semantic features does not affect the speed at which a concept is initially activated, but that it must affect slightly later mechanisms involved in response decisions.

To summarise the literature reviewed in this section thus far, it is possible to find semantic effects early, prior to 200ms, during word processing. It is, however, possible that the early significant semantic effects may be due to the task used. Abdullev and Posner (1998), Landi and Perfetti (2007), and Kounios et al. (2009) use semantic tasks which may bias word processing via semantic memory, and Hauk, Davis, et al. (2006) used lexical decision might also do this (Section 2.2.2). These button press tasks might also include effects of the decision process as well (Section 2.2.2). Button press decisions, especially explicitly semantic tasks, therefore, may not offer evidence of automatic semantic processing during the word reading (orthography-to-phonology) process, which is of importance to the central aim of this thesis. Further evidence of early semantic effects with ERP methods using a different, simpler task, such as silent word reading, and including semantic measures, such as imageability and semantic features, is relevant to the aim of this thesis, and this is explored in Experiments 5 and 6.

Before closing this literature review section, a review article must be mentioned- Barber and Kutas (2007). The article provides an extensive overview of ERP (and magnetoencephalography (MEG)) literature, resulting in a proposed time course of visual word *recognition*. Though very little of this article concerns early semantic effects, in the article, there is a very small comment made of this topic. Because the literature review highlights some semantic effects during lexical tasks, Barber and Kutas propose that in their framework for word recognition, early semantic effect could occur as lexical-semantic effects at the time-window of 150ms - 250ms. They also propose, based on their review, that any semantic effects may differ across hemisphere, especially time-windows closer to the traditional N400. After reviewing phonological effects, they conclude that the literature suggests that they start occurring in the time-window of 200ms- 350ms, which is consistent with the phonological marker chosen in Section 9.1.2. Barber and Kutas also conclude from the literature that an orthography-to-phonology-to-semantics pathway is used in some word recognition circumstances, and may be dependant on task as to whether this route is exposed. Because the literature shows evidence of orthography-to-phonology-to-semantics pathway, they claim that their proposed time course based on ERP experiments is consistent with claims of the triangle model of word reading: that is, in some circumstances (if semantic memory is emphasised, or if orthography-to-phonology is less efficient) semantic information could be accessed more quickly and efficiently using an orthographic-to-semantic-to phonology pathway than an orthography-to-phonology-to-semantic pathway (Section 1.3.3; Harm & Seidenberg, 2004). As a reminder this could apply to the use of semantic information for phonology computation, not only comprehension.

9.1.4 Time-windows and electrode sites of interest in Experiments 5 and 6

The literature presented in the previous reviews of this Chapter has provided time-windows and electrode sites that will be the focus of this chapter's analyses of ERP data. The specific time- windows will be specified shortly (Sections 9.2.2 and 9.3.2), but times-of-interest to be included in the time-windows are presented here. Phonological processing has been identified at 200ms (Section 9.1.2), and a first time-window will be prior to this. The earliest time-window investigating semantic effects in Experiments 5 and 6 will include the times of earliest semantic effects reported in the review: 150ms (Barber & Kutas, 2007; Landi & Perfetti, 2007), 160ms (Hauk, Davis, et al., 2006) and 170ms (Abdullev & Posner, 1998). A second time-window inclusive of 200ms will also be used, in keeping with the earliest findings of Kounios et al. (2009); this also encompasses semantic effects from the literature that were less early, i.e. Abdullev and Posner (1998), and the second half of the time-windows identified by Barber and Kutas (2007), and Landi and Perfetti (2007). Furthermore, a final third time-window centring on 400ms will be investigated in keeping with the later significant semantic effects presented by Hauk, Davis, et al., Landi and Perfetti, and Kounios et al.

Electrode sites of interest have also been highlighted in the literature review. As the data of Experiments 5 and 6 will be collected using high-density electrode nets, which is based on the 10-20 reference system, but includes many additional electrodes, a small cluster of electrodes around the identified sites will be used in the analyses. Data were recorded using the Cz reference (Section 9.2.1.4). During processing, data were re-referenced using an average reference (Section 9.2.1.5.2). Abdullev and Posner (1998) and Hauk, Davis, et al. (2006) found early semantic effects, at left frontal and temporal sites. Therefore one cluster, a left fronto-temporal cluster using electrodes 34, 35,39, 40, which are slightly anterior to site T3 and inclusive of electrode site F7, will be analysed. Early semantic effects were also found by Hauk, Davis, et al. at a right parietal site (P2); therefore a cluster of right occipital-parietal sites using electrodes 78, 85, 86, 92, which are between P4 and O2 electrodes and to immediate left of T8, will be used. Corresponding clusters in the opposite hemisphere will also be used (Barber & Kutas, 2007). Therefore a right fronto-temporal cluster using electrodes 109, 110, 115, and 116 will be analysed. Also a left occipital-parietal cluster using electrodes 52, 53, 61, and 60 will be analysed. Abdullev and Posner found semantic effects at left occipital-parietal sites at a time corresponding to the second time time-window of the current studies, and Kounios et al. (2009) found a hint of a latency effect at a right parietal site.

9.1.5 The two ERP experiments of this thesis

The two experiments of this chapter investigate the ERP correlates of semantic measures during silent word reading and whether any difference of the electrical signals for semantic conditions occurs before the settling of phonology. The studies concentrate on low frequency words. As reviewed previously in this chapter, the behavioural studies described in this thesis have provided some evidence that low frequency words, both regular and exception word types, receive a semantic memory contribution when computing phonology from orthography (Experiments 2 and 4, and Chapter 8). The behavioural literature has also suggested that a semantic contribution to word reading in healthy adults may occur with low frequency words (Strain et al., 1995, 2002; Harm & Seidenberg, 2004; Plaut et al., 1996; Woollams et al., 2007). Experiment 5 and 6 follow-up this evidence using neurophysiological measures, prioritising low frequency words. The two semantic variables imageability and semantic features are used with low frequency words in Experiments 5 and 6, respectively. Experiment 5 and Experiment 6 are now described in the following section.

The behavioural literature also indicates there may be differences in a semantic memory contribution for low frequency regular and low frequency exception words with the literature showing a semantic contribution for low frequency exception words as, since they have the least efficient orthography-to-phonology computation of the four word types, there is time for a semantic contribution with this word type (however, a semantic contribution is not eliminated for other word types) (Strain et al., 1995, 2002; Harm &Seidenberg, 2004; Plaut et al., 1996; Woollams et al., 2007). ERP methods, which allow for the examination of the whole time course of stimulus, including early time-windows, may detect semantic effects for low frequency regular and exception words prior to the completion of phonological processing, as suggested by some behavioural studies (Strain et al., 1995, 2002; Woollams, 2007). Therefore a study using the two low frequency word regularity types is warranted and forms the basis for Experiment 5. Imageability has been highlighted as a useful measure when investigating a semantic contribution to word reading by the regression investigations of this thesis (Chapter 8), in which imageability was a significant and unique predictor of word reading reaction times, and by the behavioural literature (Balota et al., 2004; Strain & Herdman, 1999; Strain et al., 1995, 2002; Woollams, 2005). Of special interest is semantic effects early in the word reading time-course and whether early semantic effects differ between regular and exception words. Experiment 5 uses silent word reading and four word conditions: low frequency regular and exception words manipulated on imageability, with ERP measures.

Experiment 6 investigates the ERP correlates of the semantic measure semantic features (McRae et al., 2005) when silently reading low frequency words. The regression chapter (Chapter 8) highlighted semantic features as an appropriate semantic measure for the investigations of this thesis, as they may capture qualities of the semantic elements of the priming experiments. Also the regression analyses of this thesis (Chapter 8) indicated that McRae's semantic features may be of worth in further investigations as semantic features were initially correlated with word reading reaction times and were also a significant predictor of word reading reaction times, as they were in behavioural regression literature (Pexman et al., 2002). However, semantic features failed to be a significant predictor in the investigations of Chapter 8 when imageability and age-of-acquisition were also included in the analyses. Therefore, whether an early effect during a word reading task is observed when semantic features are used as a semantic variable, while controlling age-of-acquisition and imageability, and using ERP measures is of interest.

9.1.6 Predictions

Experiments 5 and 6 investigate the ERP correlates of semantic measures imageability and semantic features, respectively, using silent reading of low frequency words. Significant semantic effects, with semantic conditions differing in average amplitude or peak latency, may provide converging evidence to the previous results of this thesis which suggest a semantic contribution to word reading, including low frequency words, and would provide information concerning the time course and processing of these semantic measures. Of special interest is whether any significant semantic effects occur prior to the identified phonological processing time-window beginning at 200ms. Both mean amplitude and peak latency will be analysed in Experiments 5 and 6.

In Experiment 5, an imageability effect (a difference between high and low imageability conditions) is predicted. Results of the priming experiments (Experiments 1, 2, and 4) of this thesis, and the behavioural literature (Shibahara et al., 2003; Strain et al., 1995; Woollams, 2005) suggest that semantic effects may occur in the low frequency exception word condition. Moreover, an imageability effect might also be expected in mean amplitude of high and low imageability low frequency regular word conditions, as in Experiments 2 and 4, and in the regression investigations of this thesis there were indications of semantic effects with word types other than low frequency regular words. In ERP Experiment 6, which uses low frequency words without a manipulation of word regularity type, a semantic features effect (difference in conditions high and low in number of semantic features) is predicted in mean amplitude, which would replicate Kounios et al. (2009).

Specific effects applicable to both experiments are now presented. Kounios et al. (2009) found that words higher in number of semantic features had a greater positive amplitude than words lower in number of semantic features. Hauk, Davis, et al. (2006) found that words higher in semantic coherence had greater amplitudes that words lower in semantic coherence. Therefore, for Experiments 5 and 6, a greater mean amplitude is predicted for conditions higher in the semantic measure than words lower in the semantic measure, i.e., imageability and semantic features, respectively. As the left fronto-temporal site, and both occipital-parietal sites have been identified in the literature as showing semantic effects at some combination of the three time windows (Abdullev & Posner, 1998, Hauk, Davis, et al., 2006) and because main effects of semantic features have also been found (Kounios, 2009), a main effect of semantic manipulation is predicted. Main effects of the semantic measures in mean amplitude indicate that semantic information is activated in the earliest time-window, prior to the completion of phonological processing , the start of which is marked as 200ms (Section 9.1.2). This will be used as evidence of early semantic effect prior to phonological completion. As there is specific interest in semantic effects prior to the 200ms marker of phonological processing, behavioural literature has found imageability effects (Shibahara et al., 2003, Strain et al., 1995, 2002, Woollams, 2005), and previous ERP investigations have found early semantic effects (Section 9.1.4.), in the absence of significant ANOVA effects with the semantic variables, planned comparisons will be performed using the two semantic conditions in the earliest time-window, and with each word regularity type as would be appropriate for Experiment 5. In the event that there is an interaction with time-window, any simple effects occurring analyses will be performed, any semantic effects within the second time-window, simultaneous with the 200ms phonological marker, but none prior to this in the earlier time-window, will need special consideration because as this would be occurring with the start of phonological processing could mean semantic activation before the completion of phonological processing.

As the aim of these two ERP experiments is to specifically investigate semantic effects using ERP methods, where there are significant effects of factors other than the semantic variables, such as effects of region or time-window or hemisphere that do not interact with the semantic variable, the statistic (F, MSe, and p values) will be reported, but not discussed further.

Also of interest are the possible semantic effects in peak latency in Experiments 5 and 6. Some evidence of latency effects was found by Kounios et al. (2009) at left occipital-parietal sites at time congruent with the second time-window of the current investigations, though this was found in the simple effects of a marginally significant interaction. As there is only weak evidence for latency effects, no specific peak latency effects are predicted. However, as this dependent measure was used by Kounios et al., which is a study directly relevant to the current experiments, and has the potential to provide useful information (Handy, 2005; Rugg & Coles, 1996) it is included as a dependent measure.

9.2 Experiment 5 – ERP

9.2.1 Methods

9.2.1.1 Participants

Ten volunteers (5 male, 5 female) aged 18 to 39 from the University of East London participated in both experiments and were given a gift voucher honorarium. All spoke English as their first language, were right-handed, and had normal or corrected-to-normal vision. The data of one participant was excluded due to a high number of artefacts in the ERP data. Results for Experiment 5 are therefore reported for 9 participants. It would have been useful to have a larger number of participants, however the sample size was constrained to the ten as this was largest number achievable given limits to time and resources.

9.2.1.2 Stimuli and design

Experiments 5 and 6 were administered in the same session, and participants were randomly assigned to one of two counterbalances. No stimuli were repeated within or across the experiments. The presentation of the words in each experiment was randomised.

One-hundred and eight low frequency English words made up the two (regular, exception spelling-to-sound) by two (high, low imageability) design of Experiment 5. There were four conditions, each with 27 words, to create a single-word, silent reading paradigm. The maximum number of stimuli for Experiments 5 was selected while matching the four conditions as far as was possible. A complete list of stimuli is available in the Appendix G. An example is provided here:

Regular, Low Imageability: dumb, sage, tuck

Regular, High Imageability: duck, spear, thigh

Exception, Low Imageability: dread, sew, ton

Exception, High Imageability: dough, shoe, tomb

The low frequency target words from the four priming experiments (Chapters 3-6) formed the foundation of the high imageability regular and exception stimulus lists as they were object names, and, as would be expected, were highly imageable. This provided 20 words for each high imageability condition, and no words for the low imageability condition. Additional stimuli were compiled from experiments in the literature that used low frequency words manipulated on imageability and regularity. The works consulted were Woollams (2005), Strain et al. (1995, 2002), Cortese and Simpson (2000), Ellis and Mongahan (2002), Cortese et al. (1997), and Strain and Herdman (1999).

Given the design, the number of trials was determined by the number of word stimuli available. A larger number of stimuli reduces the signal-to-noise ratio in the ERP data, meaning that the data are less noisy, i.e., the ERP signal in the data, in response to the stimuli, is discernable from the extraneous electrical noise in the data. The number of stimuli for Experiments 5 and 6, 27 and 30 stimuli per condition, respectively is greater than the minimum number of stimuli specified by Rugg and Allan (2000, p521) and Fabiani, Gratton, and Coles (2000, p54) , which is 20. However, Luck (2005b, p30) specifies a minimum of 30 stimuli per condition for examining large known components, with smaller effects more stimuli are recommended to reduce the noise in the data. The number of stimuli in Experiment 5 (27 per condition) was slightly less than that specified by Luck, and the number of stimuli in Experiment 6 (30 per condition) was equal to the minimum suggested number by Luck (2005b). Measures were implemented to increase the signal-to-noise ratio, therefore reducing noise, using methods specified by Luck (2005b, p30-31); this included, but was not limited to using a healthy adult population who were instructed to limit muscular movements, including limiting blinking to between stimuli trials, recording in an acoustically shielded both with electrical equipment, as far as was possible, outside of the shielded booth and shielding cables in the booth, and designing the experiments with short blocks so that participants stayed alert.

9.2.1.2.1 Frequency, regularity, and imageability thresholds

Words had a Kucera and Francis (1967) written frequency count equal to or below 16 per million, as provided by the English Lexicon Project Database (Balota et al., 2007) and a Celex written frequency count equal to or below 14 per million, as provided by MCWord: An Orthographic Wordform Database (Medler & Binder, 2005). These are the same low frequency thresholds that were used in the priming experiments (Chapters 3-6). Moreover, the four conditions were matched for Kucera and Francis (1967) frequency and Celex lexical database frequency (Baayen et al., 1993). Spelling-to-sound regularity and exception categorisation was checked using the N-Watch program (Davis, 2005). Regular word conditions were listed as ‘regular’ and exception words were listed as ‘exception’ by the program. The MRC Psycholinguistic database provided imageability ratings for the words. High imageability words were above 538 in the MRC database, which uses 100 times multiplication of a one to seven rating. Low imageability words were below 531.

9.2.1.2.2 Stimuli matching

Averages of each measure as a function of condition and significance values (p-values) of the matching analyses are presented in Table 9.1. The four lists did not differ in frequency, number of phonemes, length in letters and syllables, number of orthographic neighbours (Coltheart's N), or number of phonological neighbours; nor did the high and low imageability lists differ on these measures within word type or when collapsed over word type. Regular and exception word lists when collapsed over imageability rating also did not differ on these measures. Due to constraints of the experimental manipulations and the use of the target words from the priming experiments, high and low imageability conditions were not matched on age-of-acquisition (Cortese & Khanna, 2008) or familiarity (Wilson, 1988). As would be expected with a manipulation of imageability, high and low imageability conditions significantly differed on imageability.

Variable	Exception			Regular			Four conditions	Colapsed over Regularity			Colapse over Imageability		
	Low Imageability	High Imageability	p-value	Low Imageability	High Imageability	p-value	p- value	Low Imageability	High Imageability	p-value	Exception	Regular	p-value
KF Frequency	5.81	6.48	0.61	5.04	5.11	0.95	0.58	5.43	5.80	0.66	6.15	5.07	0.20
Celex Total Frequency	5.74	6.11	0.73	5.31	6.79	0.23	0.61	5.53	6.45	0.25	5.92	6.05	0.88
Celex Written Frequency	6.00	6.41	0.70	5.49	7.14	0.20	0.55	5.74	6.78	0.21	6.21	6.32	0.90
Number of Phonemes	3.56	3.78	0.43	4.07	3.59	0.18	0.35	3.81	3.69	0.57	3.67	3.83	0.46
Syllable	1.15	1.22	0.49	1.22	1.19	0.74	0.89	1.19	1.20	0.81	1.19	1.20	0.81
Length (in number of letters)	4.70	4.93	0.46	5.04	4.67	0.27	0.59	4.87	4.80	0.74	4.81	4.85	0.87
Familiarity	452.48	479.91	0.19	439.74	506.43	0.001	0.007	446.11	493.17	0.001	465.10	470.42	0.72
Coltheart's N	5.89	3.70	0.11	5.37	6.37	0.45	0.22	5.63	5.04	0.54	4.80	5.87	0.26
Phonological Neighbourhood	13.63	11.33	0.33	14.41	15.08	0.79	0.43	14.02	13.17	0.62	12.48	14.74	0.18
Age-of-Acquisition (Cortese &Khanna, 2008)	4.91	3.94	0.001	4.99	3.35	<.001	<.001	4.95	3.64	<.001	4.46	4.15	0.20
Imageability (MRC Psycholinguist Database)	438.07	620.48	<.001	401.89	626.30	<.001	<.001	419.98	623.39	<.001	529.28	514.09	0.49

Table 9.1. The average values of various measures for stimulus words of the four conditions of Experiment 5, with significance p-values from matching analyses. See text for details. KF = Kucera and Francis.

9.2.1.3 Procedure

Once the EEG net of electrodes had been placed on their head, participants sat in the recording booth (Section 9.2.3) one metre from the computer screen and were given a button-box with which to respond. E-prime version 1.2 (Psychological Software Tools, Pittsburgh, PA) was used to present the stimuli and collect the response times of the button presses[51]. Each word was presented individually in lower case letters in white 32 point Arial font on a grey background in the centre of the computer screen for 500ms. Participants were instructed to name each word silently to themselves and to press a button with their right index finger when covert silent reading was completed. Emphasis was placed on the importance of reading the word and preparing a response as if it was to be read aloud, yet not speaking aloud, as this would interfere with the recording. Following each word, a fixation point that varied in length between 1800ms and 2800ms was displayed until the next word was presented. Participants were also instructed to limit their movements especially when words were on the screen and that they could blink during the small break in which the fixation cross appeared. Participants received a set of practice trials. Practice trials consisted of 16 words unique to the practice were presented in keeping with the procedure. The experiment was divided into two mini-blocks of an equal number of stimuli (54) in each half; the short break between the two mini-blocks was self-paced. The experiment lasted a maximum of 10 minutes.

[51] Due to a programming error, button press responses were not reliably collected by E-Prime.

9.2.1.4 ERP recording

Participants sat within an acoustically shielded, dimly lit booth for the recording. They were fitted with a high-density 128-channel HydroCel Geodesic sensor net, which is based on the 10-20 reference system, but includes many additional electrodes between these sites. The electrode impedances did not exceed 50kΩ. The sample rate was 250Hz. EEG data were recorded using Net Station software, version 4.2.4 (Electrical Geodesics Inc, EGI), using Cz as a reference and the Geodesic sensor net layout 128 2.1. The EEG signals were amplified using the Net Amps 200 amplifier for 128-channel EEG, provided by Electrical Geodesics Inc (EGI). EEG signals were continuously recorded for the experiment.

9.2.1.5 ERP data analysis

9.2.1.5.1 Pre-Processing

EEG data were processed using Net Station tools available in version 4.2.4. EEG recordings were filtered off-line using a high-pass filter of 1.0 Hz and a low-pass filter of 30Hz. Recordings were divided into epochs that encompassed stimulus onset and presentation duration per stimulus per participant. Data epochs were 700ms in length, beginning 200ms pre-stimulus onset and finishing at stimulus off-set, 500ms after the stimulus onset. “Bad” channels, eye blinks, and eye movements were marked in the data files using the Net Station artefact detection tool and default settings. An epoch from an electrode was excluded from further analysis if the maximum amplitude minus the minimum amplitude exceeded 200.00µv. An epoch from an electrode was eliminated due to an eye-blink if the maximum amplitude minus the minimum amplitude exceeded 140.00μv. An epoch from an electrode was excluded due to eye-movements if the maximum amplitude minus the minimum amplitude exceeded 55.00μv. The tool used a moving average of 80ms, the default. An entire electrode was excluded if more than 20% of segments were marked for elimination by the previous settings. An epoch for a participant was eliminated if 10 of the 128 channels were marked. In addition to this visual inspection of the data ensured the software correctly marked the data.

9.2.1.5.2 Processing

After pre-processing, the data were also processed using Net Station tools to replace "bad channels", correct the data with a baseline, and re-reference. Firstly, bad channel replacement was carried out using the bad channel replacement tool. Electrodes above a specified impedance threshold or containing a majority of "bad" epochs were replaced. Bad channels and bad epochs were replaced using a default calculation within Net Station, which used data from "good" epochs from the recording. Data was baseline corrected using the 200ms before stimulus onset, and was re-referenced using an average reference. ERP averages were then created for each participant in each of the four conditions at each electrode site. It is these data that were statistically analysed (Section 9.2.2). Based on the number of stimuli and the number of participants, the maximum number of segments that could have contributed to each conditions average was 270, and the maximum per subject per condition was 27. The number of segments that actually contributed to the average ERP data are as follows: for the low imageability exception word condition 206 segments were used, with each participant contributing 22.89 segments on average; for the high imageability frequency exception word condition 207 segments were used, with each participant contributing 23 segments on average; for the low imageability regular word condition 219 segments were used, with each participant contributing 24.44 segments on average; for the high imageability regular word condition 206 segments were used, with each participant contributing 23.78 segments on average. Finally, grand averages, providing an average across all participants for each of the four conditions at each electrode site was created for each 700ms epoch; grand averages and their topography are presented in Figure 9.1. Within this figure the four electrode site regions of interest, introduced in Section 9.1.4 and presented further in Section 9.2.2, are magnified.

9.2.2 Results

The waveform averages of each participant and each condition were inspected to identify the precise boundaries of the three time-windows that encompass the *a priori* times of interest listed in Section 9.1.5. As suggested by Luck (2005), care was taken to ensure time windows were at least 40ms in length.

The three time-windows are: 150ms-190ms, 190ms-250ms, and 350ms-450ms. The earliest time-window analysed (150ms-190ms) occurred prior to the phonological processing marker. The second time-window (190ms-250ms) occurred simultaneous with the phonological processing marker. The third time-window (350ms-450ms) selected for analysis was after the phonology marker, though it is possible that phonological processing could occur in this time-window.

Electrode sites of interest have also been identified (Section 9.1.4). In the analyses, as the data was recorded using a high-density electrode net (Section 9.2.3) an average of four electrode sites (a cluster) in each *a priori* region and hemisphere were used in the analyses; these are presented in Table 9.2.

Fronto-temporal		Occipital-Parietal	
Right	Left	Right	Left
109,110,115,116	34, 35, 39, 40	78, 85, 86, 92	52, 53, 61, 60

Table 9.2. Electrode site clusters used for analyses. Individual electrode sites averaged as a cluster, listed by region (Fronto-temporal or Occipital-Parietal) and hemisphere (Right or Left).

Grand averages for each of the four stimulus word conditions at each electrode site are presented topographically in Figure 9.1.; the four electrode clusters used in the analyses are magnified in Figure 9.1 and also provided with more detail in Figure 9.2.

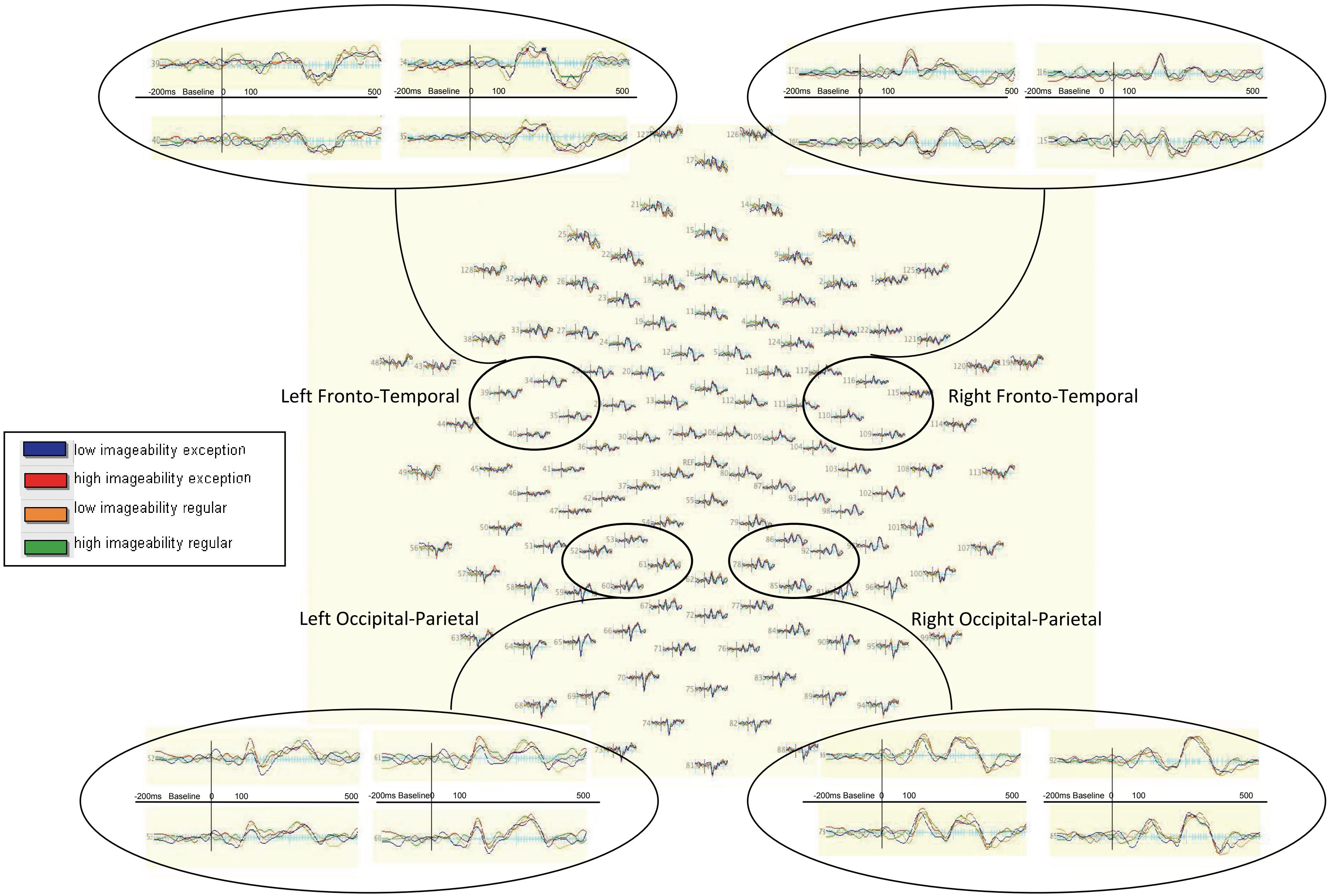

Figure 9.1. Grand average waveforms of Experiment 5 presented topographically. Waveforms are the average voltage across participant for the four conditions at each topographic electrode site for the 700ms epochs. The four electrode clusters of interest are presented in magnification with the baseline section marked. The four electrode clusters are presented in further magnification on the following page in Figure 9.2.

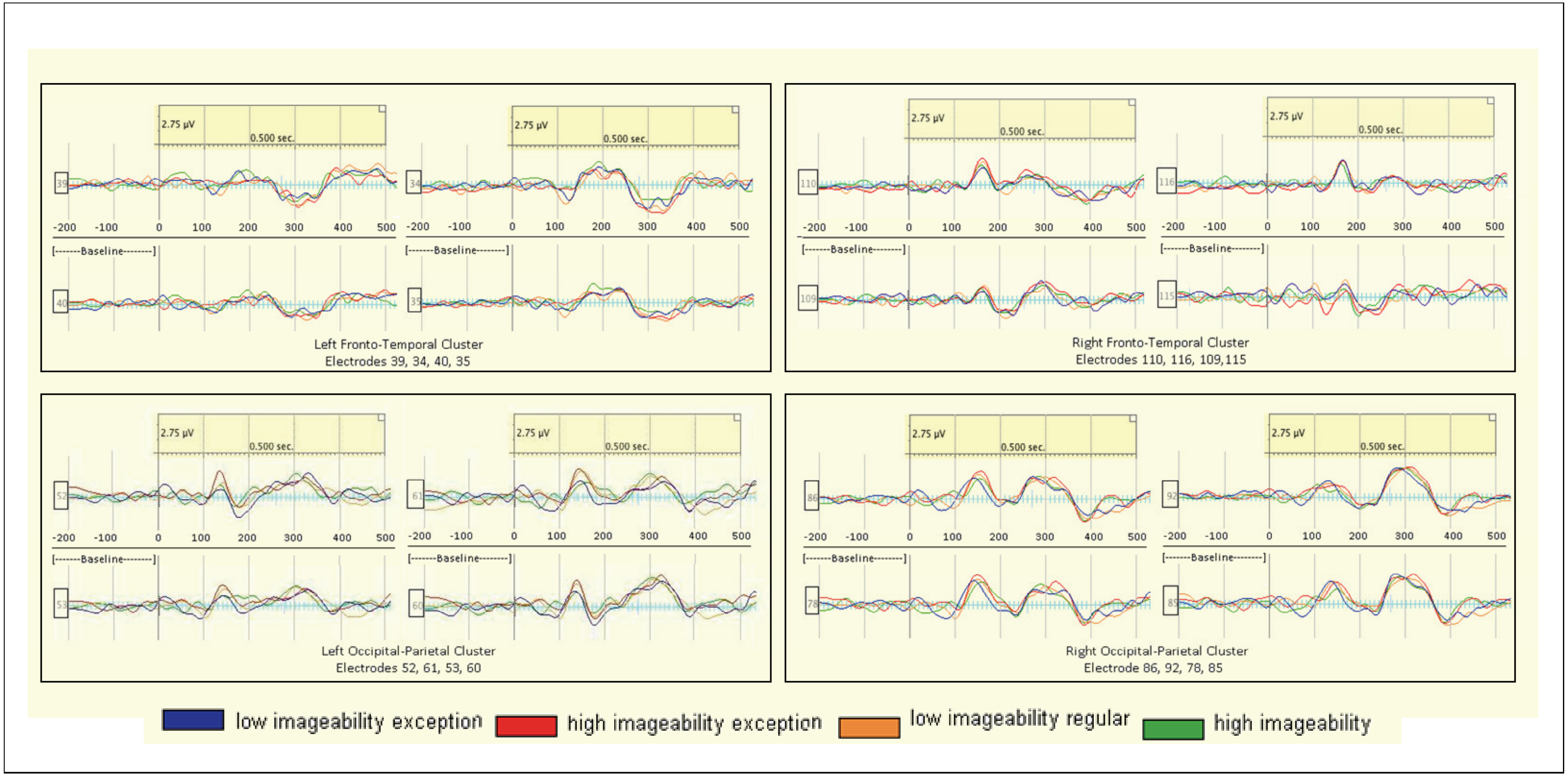

Figure 9.2. Four electrode clusters of Experiment 5 expanded. Waveforms are the average voltage across participant for each of the four conditions at each electrode site within each cluster of interest. Individual electrodes, clusters of electrodes, baselines, and timescales are labelled.

9.2.2.1 Mean amplitude analyses

Average amplitude ERP data were analysed using a 2 x 2 x 3 x 2 x 2 ANOVA, i.e. Regularity (Regularity, Exception) x Imageability (High, Low) x Time-window (150ms-190ms, 190ms-250ms, 350ms-450ms) x Region (Fronto-temporal, Occipital-parietal) x Hemisphere (Right, Left). There was a significant main effect of imageability, $F(1,8) = 14.72$, MSe = .60, $p = 0.005$. On average, high imageability words had a higher mean amplitude than low imageability words. No other main effects were significant. One interaction approached significance, regularity x imageability x region x hemisphere, $F(1,8) = 4.55$, MSe = 0.26, $p = .065$, and there were four significant interactions; two involving location, but not imageability, region x hemisphere, $F(1,8) = 7.95$, MSe = .73, $p = 0.023$, time-window x region x hemisphere, $F(2,16)$, MSe = .75, $p = 0.026$, and two involving imageability, imageability x region, $F(2,16) = 5.61$, MSe = 0.65, $p = 0.045$, imageability x time-window x hemisphere, $F(2,16) = 3.68$, MSe = 0.15, $p = 0.048$. There were no other significant interactions.

Topographic maps of the amplitudes of the four word conditions at each of the three time-windows are displayed in Figure 9.3. These topographies show the directional charge and size of the voltage as measured at the scalp. As there was not a significant interaction with regularity, the figure is organised to highlight high and low imageability conditions. Of the three time-windows, the voltage distributions for high and low imageability words in the earliest time-window (150ms-190ms) appears be the most different. Within this time-window, there appears to be differences between high and low imageability conditions in voltage strength for left and right hemisphere, and fronto-temporal and occipital-parietal. The strongest voltages (both strong positive and strong negative voltages) appearing in the right occipital-parietal region. The second time-window (190ms-250ms) has the weakest average voltages, with small differences appearing in the pattern of voltages for fronto-temporal regions versus occipital-parietal regions between high and low imageability conditions. The latest time-window (350ms-450ms) appears most similar topographic voltage patterns for high and low imageability conditions.

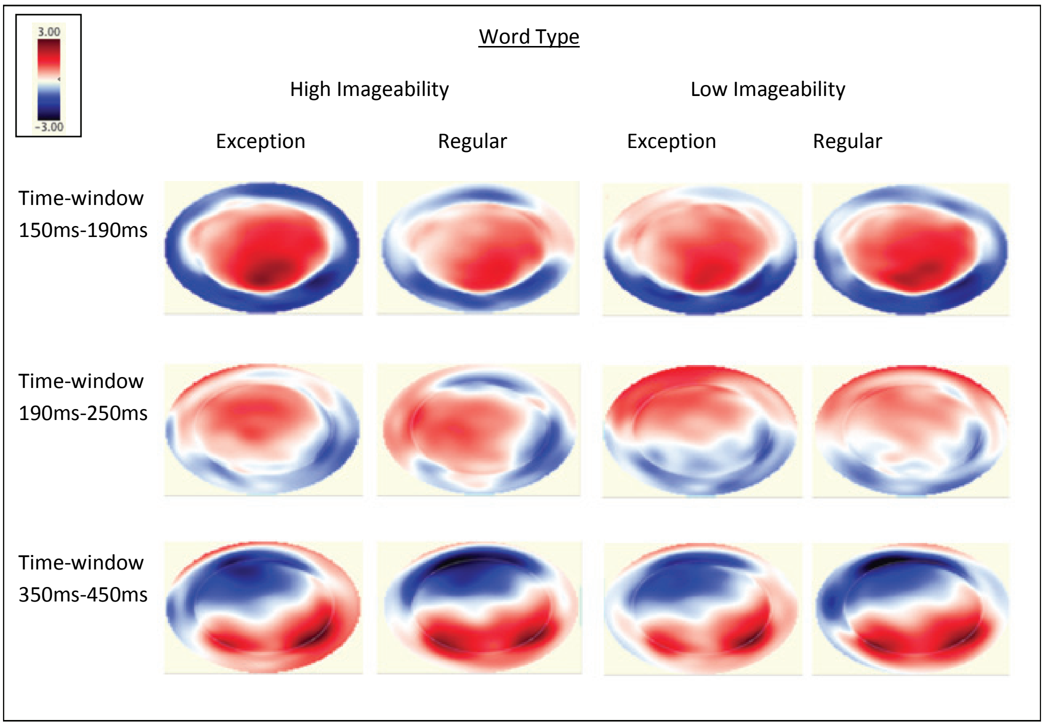

Figure 9.3. ERP topography voltage distribution plots for Experiment 5. The data are displayed for each word regularity type within each imageability type for the three time-windows. Within each diagram, dorsal is top and ventral is bottom, with positive voltages depicted in red and negative voltages depicted in blue. Simple effects are presented in Sections 9.2.2.1.

Simple effects analyses of the two significant interactions involving imageability were performed (Howell, 1992). To reduce the possibility of Type I error; critical p-values were set using a Bonferroni correction, effects were only considered significant if *p* was less than the critical p-value reported with the analysis (Howell, 1992).

Simple effects analyses of the imageability by region interaction compared high and low imageability conditions at each region, critical $p = .025$. The average amplitude of high and low imageability conditions significantly differed in the occipital-parietal regions, $t(8) = 3.88$, SEM = 0.12, $p = .005$, but not in the fronto-temporal regions, $t(8) = 1.11$, SEM = 0.09, $p = .299$. In the occipital-parietal regions high imageability words have a greater mean voltage than low imageability words. Figure 9.4. presents the interaction data.

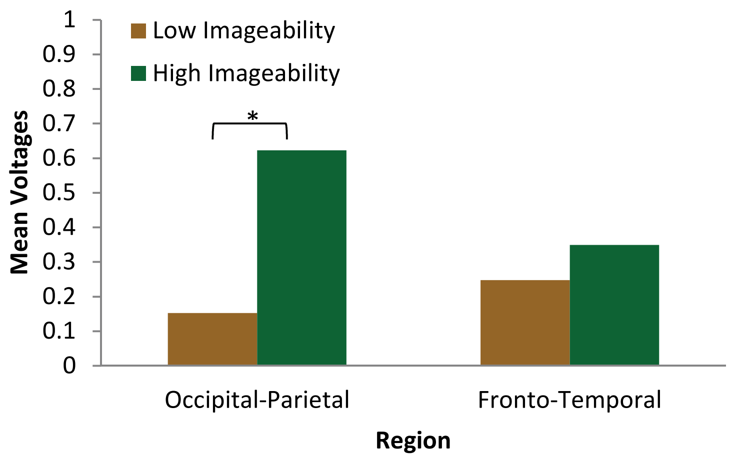

Figure 9.4. Average voltage amplitude for stimulus words in high imageability and low imageability conditions in Experiment 5 as a function of region, but collapsed across the three time-windows. There is a significant imageability effect at occipital-parietal electrode sites, as is specified in the figure (*). See text for details.

Simple effects analyses of the imageability by time-window by hemisphere interaction compared high and low imageability conditions at each time-window in each hemisphere. There was a significant difference between high and low imageability conditions in the 150ms-190ms time-window in the left hemisphere, $t(8) = 4.35$, SEM = 0.12, $p = .002$, with high imageability words having a greater mean amplitude than low imageability words. No other comparisons reached significance, critical $p = .0083$. Figure 9.5 depicts the interaction data.

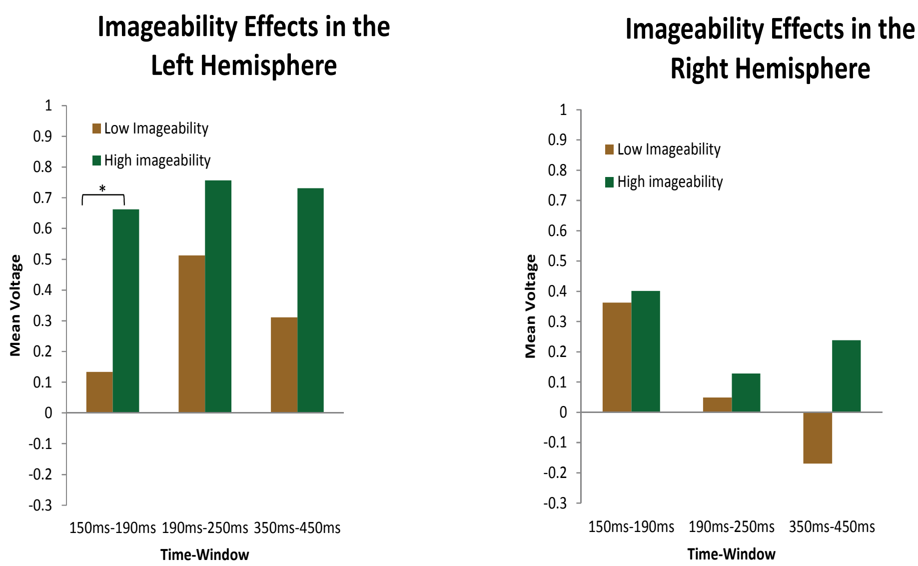

Figure 9.5. Average amplitude as a function of stimulus word imageability for left hemisphere as a function of time-window and for right hemisphere as a function of time-window Experiment 5. There was a significant difference between high and low imageability conditions in the 150ms-190ms time-window in the left hemisphere, as is specified in the figure (*). See text for details.

9.2.2.2 Peak latency analyses

Peak latency ERP data were analysed using a similar ANOVA analysis. The time-window factor was not included in the peak latency analysis as the time-windows by definition have peaks at different time points. The inclusion of this factor would yield significant latency effects between time-windows, when the true interest is in significant latency effects within a time-window. Therefore latency was analysed using a 2 x 2 x 2 x 2 ANOVA at each of the three time-windows, i.e., Regularity (Regularity, Exception) x Imageability (High, Low) x Region (Fronto-temporal, Occipital-Parietal) x Hemisphere (Right, Left).

The main effect of imageability was significant in the latest time-window of 350ms-450ms, with high imageability words peaking significantly earlier than low imageability words, $F(1,8) = 13.73$, MSe = 234.144, $p = 0.006$. The interaction of imageability and region approached significance in this time-window (350ms-450ms), $F(1,8) = 5.13$, MSe = 369.53, p=0.053. There was a significant main effect of hemisphere in the 150ms-190ms time-window, $F(1,8) = 15.91$, MSe = 312.32, p = 0.004; the left hemisphere had a later peak amplitude than the right hemisphere This main effect approached significance in the 190ms-250ms time-window, $F(1,8) = 4.67$, MSe = 289.74, $p = 0.06$, and in this time-window (190ms-250ms), the interaction of hemisphere and region approached significance, $F(1,8) = 4.69$, MSe = 521.77, p = 0.06. In this time-window (190-250ms) the right hemisphere had a later peak latency than the left hemisphere. There were no other significant main effects or significant interactions.

9.2.3 Discussion

Experiment 5 investigated the ERP correlates of the semantic variable imageability in low frequency silent word reading with a manipulation of spelling-to-sound regularity. Previous investigations both of this thesis (Chapters 3-8) and in the literature (Shibahara et al., 2003, Strain et al., 1995, 2002, Woollams, 2005) have suggested that semantic information may affect reading of low frequency words. There have also been reports of early semantic effects in the ERP literature using tasks other than word reading. Effects have been reported with lexical decision at 160ms (Hauk, Davis, et al., 2006), and with semantic judgements at 170ms (Abdullev & Posner, 1998) and 150ms-250ms (Landi & Perfetti, 2007). Of interest is whether there are significant semantic effects prior to the phonological marker (Section 9.1.2) in this experiment that used a silent word reading task and stimuli manipulated on imageability. Also of interest in whether there is a difference in semantic effects early in the word reading time-course between low frequency regularity word types.

Average amplitude analyses showed significant imageability effects in main effects, interactions, and simple effects. Though high and low imageability words were well matched, the exceptions were age-of-acquisition ratings and familiarity (Table 9.1); therefore there could be an influence of familiarity and age-of-acquisition on the ERP data. However, the results are likely due to an effect of the imageability manipulation since the regression analyses of Chapter 8 show that imageability effects are present even when familiarity and age-of-acquisition have been controlled. This argument is presented further following Experiment 6 in the general discussion of this Chapter (Section 9.4). The semantic measure, here, imageability, likely affects ERP voltages, thus reflecting differences in neural processes for the two imageability conditions during silent word reading (Luck, 2005; Rugg & Coles, 1996). Imageability also significantly interacted with region in the ANOVA of mean amplitudes; the differences in processing of high and low imageability words are reflected at occipital-parietal scalp locations. Differences in scalp voltage distributions between experimental conditions may be the result of any number of different patterns of cortical activation (Handy, 2005; Luck, 2005; Otten & Rugg, 2005), therefore what precise cortical activation differences are reflected between high and low imageability words by different scalp voltage distributions in different scalp regions cannot be known; (this also applies to differences in scalp hemisphere distributions as seen in the other significant interaction with imageability, discussed next). This interaction of imageability and region in mean amplitude, however, confirms the differences between processing for high and low imageability words. Therefore, the ERP data of

Experiment 5 offers evidence that processing differs for high and low imageability words during silent word reading.

Simple effects, which followed-up the interaction of imageability with time-window and hemisphere within the ANOVA of mean amplitudes, show imageability effects in the earliest time-window (150ms-190ms) at scalp electrode sites over the left hemisphere. Therefore the ERP data of Experiment 5 specifically provides evidence that semantic effects occur prior to the phonological processing marker at 200ms (Section 9.1.2). It is possible for this early semantic processing to contribute to orthography-to-phonology computation. There was no clear evidence of any effects of the regularity manipulation; therefore, there is some indication that semantic effects are present in low frequency regular words and low frequency exception words. There is certainly no evidence for differences between word regularity types in the early time-window.

There was only one significant semantic peak latency effect; this occurred in the latest time-window, 350ms-450ms, with high imageability words peaking later than low imageability words. The later semantic effect (in peak latency) in combination with the early significant semantic effect (in mean amplitude) suggest that semantic information is used repeatedly during word reading, perhaps for different purposes. The early semantic effects might indicate a semantic information contribution to orthography-to-phonology computation, as they occur prior to the phonological processing marker, whereas the later semantic effects in peak latency later in the time course could indicate cascaded processes (Hauk, Davis, et al., 2006), with semantic information again being accessed. Accessing semantic information again may suggest it's use to "clean-up" phonological processing (Plaut et al., 1996; Harm & Seidenberg, 1999) or its use for comprehension of the word (Hauk, Davis et al. 2006; Barber &Kutas, 2007). The ERP data of Experiment 5 reveals that semantic information is active during silent word reading, both before and after the phonological marker of 200ms.

The significant early semantic effects in the 150-190ms time-window (as shown by the simple effects in the imageability, time-window, and hemisphere interaction) replicated the early semantic effects reported in the literature (Abdullev & Posner, 1998; Hauk, Davis, et al., 2006; Landi & Perfetti, 2007). In Experiment 5, the simple effects revealed early semantic effects in the time-window of 150ms-190ms, at left hemisphere sites, though this did not interact with region. Hauk, Davis, et al. (2006) found an early semantic effect (160ms) using semantic coherence and a lexical decision task at left frontal-temporal sites and right occipital-parietal sites. Early semantic effects using semantic tasks reported at 150ms (Barber & Kutas, 2007; Landi & Perfetti, 2007) and at 170ms at left-fronto-temporal sites and right occipital-parietal site (Abdullev & Posner, 1998). The results of Experiment 5 confirm these effects with a task of word reading and the semantic variable of imageability. The early semantic effects found in Experiment 5-ERP, however, were earlier than those reported by Kounios et al. (2009), who reported effects of semantic features at the slightly later time window of 200ms-300ms. Finally, in Experiment 5 the condition higher in the semantic measure (high imageability) produced higher amplitudes, replicating the effects reported by Hauk, Davis, et al. (2006) and Kounios et al. (2009). Experiment 5 extended the previous work by using a different task (silent word reading) and different semantic measure (imageability) than the reviewed studies.

9.3 Experiment 6- ERP

Experiment 6- ERP investigated the ERP correlates of semantic features using mean amplitude and peak latency of ERP data. In Experiment 6, as a reminder, low frequency words were manipulated on the number of semantic features (high or low in the number of semantic features; McRae et al., 2005).

9.3.1 Methods

9.3.1.1 Participants

Experiment 6 data was collected in the same recording session as Experiment 5; therefore participant details are the same as those described in Section 9.2.1.1. For Experiment 6, the data is reported from 10 participants.

9.3.1.2 Stimuli and design

To create the initial stimulus lists, McRae et al.'s (2005) complete list of words with a count of semantic features inclusive of taxonomic features was ordered by number of semantic features; this is the same semantic feature measure that was used in the regression analyses of Chapter 8. To achieve a good separation of between the two conditions: words high in the number of semantic features and words low in the number of semantic features words, the words in the top quarter of this list with the highest number of semantic features and the words in the bottom quarter of this list with the lowest number of semantic features were selected. These words (106) were presented in the recording session.

For the analyses of this chapter, a subset of 30 well matched (Table 9.3) one- and two-syllable words from each condition, i.e. high in number of semantic features or low in number of semantic features, was selected. Words were in keeping with the target stimuli in the priming studies. A stimulus list is available in Appendix H; an example is provided here:

High in semantic features: hare, sofa, turkey

Low in semantic features: harp, shell, turnip

The maximum number of words was selected while matching the two conditions on a number of factors (Section 9.3.1.2.2). Though this lowered the number of trials per condition to 30 and therefore potentially increased the noise in the data, there were still a greater number of trials per condition than the minimum for adequate signal-to-noise ratios in an experiment, which is 20 (Fabiani et al., 2000, p54; Rugg & Allan, 2000, p521), and an equal number of trial per condition as the minimum suggested by Luck (2005b, p30) for large components. As noted in Section 9.2.1.2., measures were taken to increase the signal-to-noise ratio in the data.

9.3.1.2.1 Frequency and number of semantic features

Stimulus words were low in frequency. Twenty-six of the 30 low in features words and 28 of the 30 high in feature words were below 15 instances per million using Kucera and Francis frequency counts (1967), which was the low frequency threshold used in the priming experiments of this thesis (see Section 3.2.3.1) . All words had a frequency count of less than 37 instances per million using the Kucera and Francis frequency counts (1967). Words with a high number of semantic features ranged in number from 18 to 26 features; words with a low number of semantic features ranged in number from six to nine features.

9.3.1.2.2 Stimuli matching

High-in-feature and low-in-feature conditions were matched on Kucera and Francis (1967) frequency, Celex lexical database total and written frequencies (Baayen, Piepenbrock, & Van Rijn, 1993), number of phonemes, length in syllables and letters, familiarity, Colthearts's N, phonemic neighbourhood, age-of-acquisition, and imageability. The lists were also matched in terms of initial phoneme as far as was possible. The two lists significantly differed on number of semantic features, as would be expected for this type of manipulation. Averages of each measure as a function of number of features condition and significance values are provided in Table 9.3.

Measure	Low in Features	High in Features	p-value
KF Frequency	6.27	6.37	0.96
Celex total Frequency	7.58	7.85	0.91
Celex written Frequency	7.91	8.29	0.88
Number of phonemes	4.47	4.43	0.93
Syllable	1.6	1.53	0.61
Length (in number of letters)	5.67	5.67	> 0.99
Familiarity (McRae et al., 2005)	5.55	6.06	0.32
Coltheart's N	4	3	0.4
Phonological neighbourhood	9.04	8.32	0.68
Age-of-acquisition (Cortese & Khanna, 2008)	3.8	3.5	0.23
Imageability (MRC Psycholiguistic Database)	579.14	591.18	0.18
Number of Semantic Features (McRae et al., 2005)	8.2	19.57	< 0.001

Table 9.3. The average values of various measures for stimuli from the two conditions of Experiment 6, with significance p-values from matching analyses. See text for details. KF = Kucera and Francis (1967).

9.3.1.3 Procedure, recording, and ERP data analysis

Procedural details were the same as in Experiment 5 (Section 9.2.1.3), exceptions are detailed here. Experiment 6 was divided into two mini-blocks of equal number of stimuli (53 words in each block). Participants received a short self-paced break between the two mini-blocks. The experiment lasted a maximum of ten minutes.

ERP data recording and ERP data analysis (pre-processing and processing) were the same as detailed in Experiment 5 (Sections 9.2.1.5), with the exception of baseline correction processing, which is detailed here. Based on the number of stimuli and the number of participants, the maximum number of segments that could have contributed to each conditions average was 300, and the maximum per subject per condition was 30. The number of segments that actually contributed to the average ERP data were as follows: for the low in number of semantic features condition 278 segments were used (maximum possible 300), with each participant contributing 27.80 segments on average (maximum possible 30); for the high in number of semantic features condition 271 segments were used, with each participant contributing 27.1 segments on average.

As in Experiment 5, epochs were 700ms long, with 200ms pre-stimulus and 500ms post-stimulus onset. Visual inspection of the waveforms revealed the data from the two semantic feature conditions varied from -200ms to -100ms. The waveforms for the two semantic features conditions were more similar for -100ms to 0ms. These two times are noted in the magnified electrode views of Figures 9.6 and 9.7. As a 100ms length of baseline is sufficient according to Luck (2005) and has been used by others

(Hauk, Davis, et al., 2006; Landi & Perfetti, 2007; Sereno et al., 1998), data for Experiment 6 were baseline corrected using the 100ms before stimulus onset.

9.3.2 Results

Data analysis details are the same as Experiment 5 (Section 9.2.2), unless detailed here. The three previously identified time-windows of Experiment 5 were also appropriate for the data of Experiment 6. The three time-windows used were: 150ms-190ms, 190ms-250ms, and 350ms-450ms. The four electrode clusters (left and right fronto-temporal sites and left and right occipital-parietal sites) identified in Section 9.2.2 and Table 9.2 were also used in Experiment 6. Grand averages for each condition at each electrode site are presented topographically in Figure 9.6; the four electrode clusters used in the analyses are magnified and also provided in Figure 9.7.

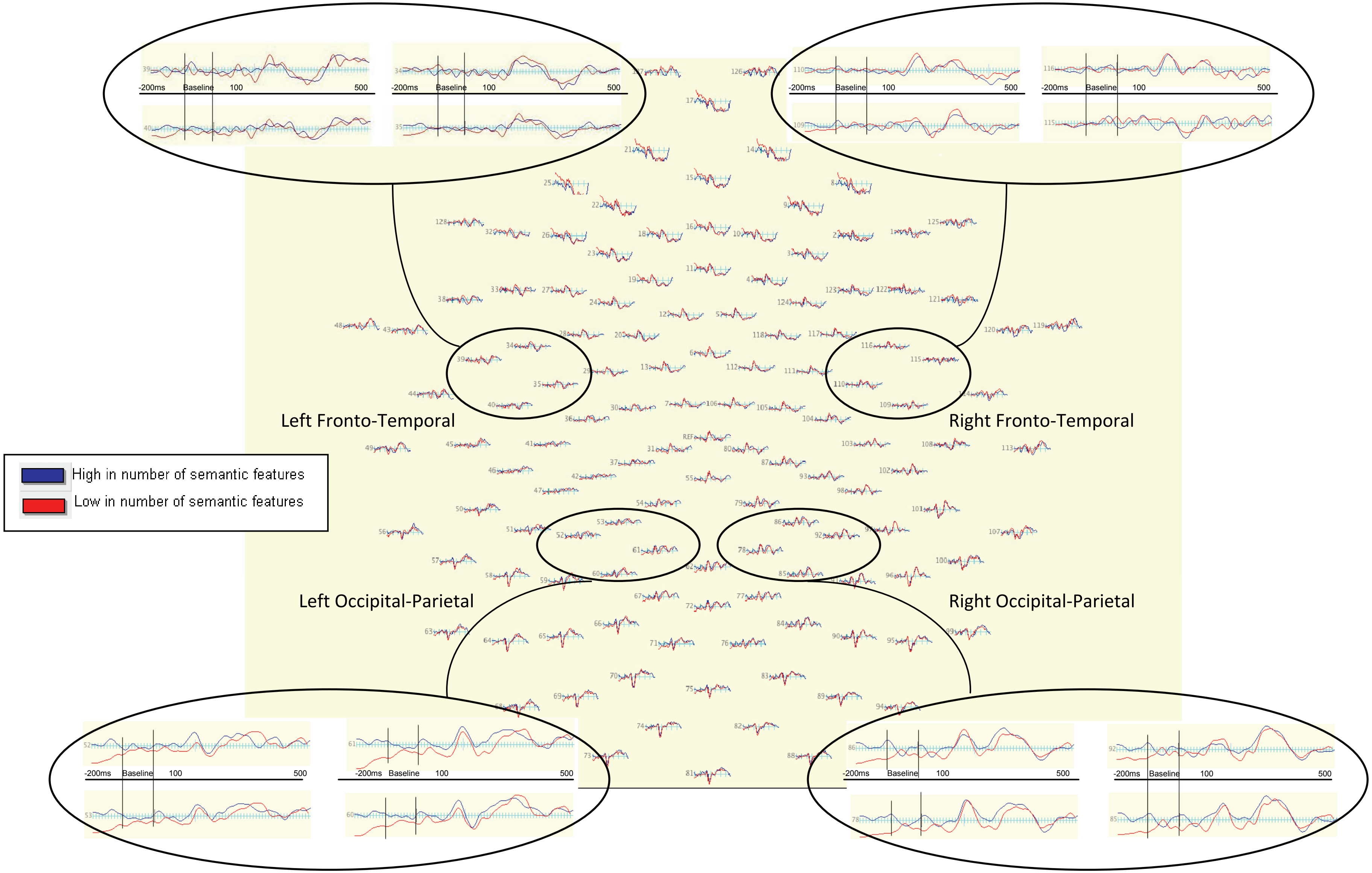

Figure 9.6. Grand average waveforms of Experiment 6 presented topographically. Waveforms are the average voltage across participant for the two semantic feature conditions at each topographic electrode site for the 700ms epochs, baseline corrected using the -100 to 0 time-window. The four electrode clusters of interest are presented in magnification with the labelled baseline section. The four electrode clusters are presented in further magnification on the following page in Figure 9.7.

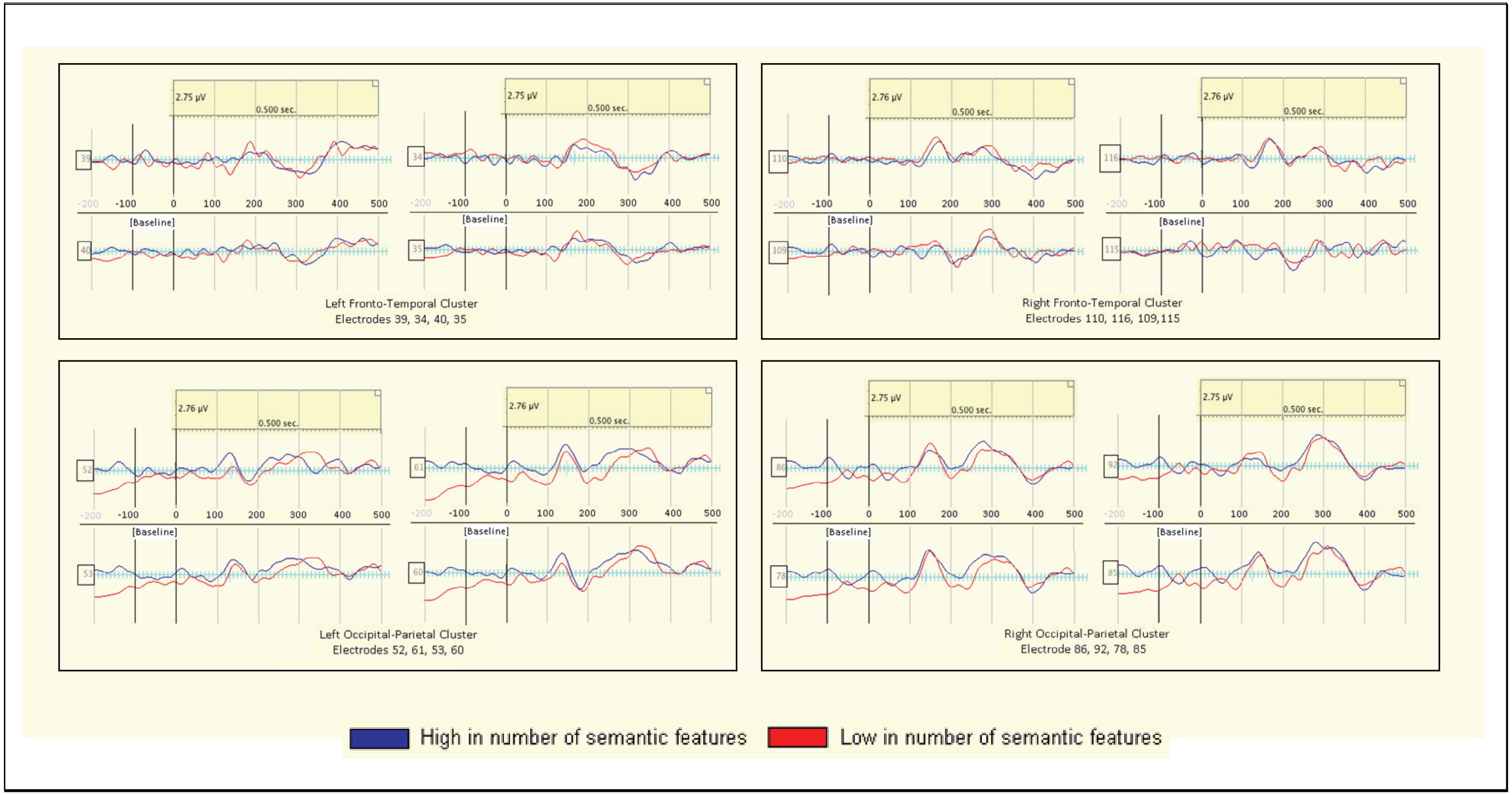

Figure 9.7. Four electrode clusters of Experiment 6 expanded. Waveforms are the average voltage across participant for each of the four conditions at each electrode site within each cluster of interest. Individual electrodes, clusters of electrodes, baselines, and timescales are labelled.

9.3.2.1 Mean amplitude analyses

Mean amplitudes were analysed using a 2 x 3 x 2 x 2, i.e. Semantic features (high, low) x time-window (150ms-190ms, 190ms-250ms, 350ms-450ms) x Region (Fronto-temporal, Occipital-parietal) x Hemisphere (Right, Left), ANOVA. There was a main effect of semantic features was significance, $F(1,9) = 16.56$, MSe = .24, $p = .003$. On average, low number of semantic feature words had a lower mean amplitude than high number of semantic feature words. There was a significant interaction of semantic features and time-window, $F(2,18) = 3.824$, MSe = 0.16, $p = .04$. There was also a significant interaction of number of semantic features and region, $F(1,9) = 14.06$, MSe = 1.01, $p = .005$. There was a significant interaction of time-window and region, $F(2,18) = 6.58$, MSe = 1.65, $p = .007$. No other main effects or interactions were significant.

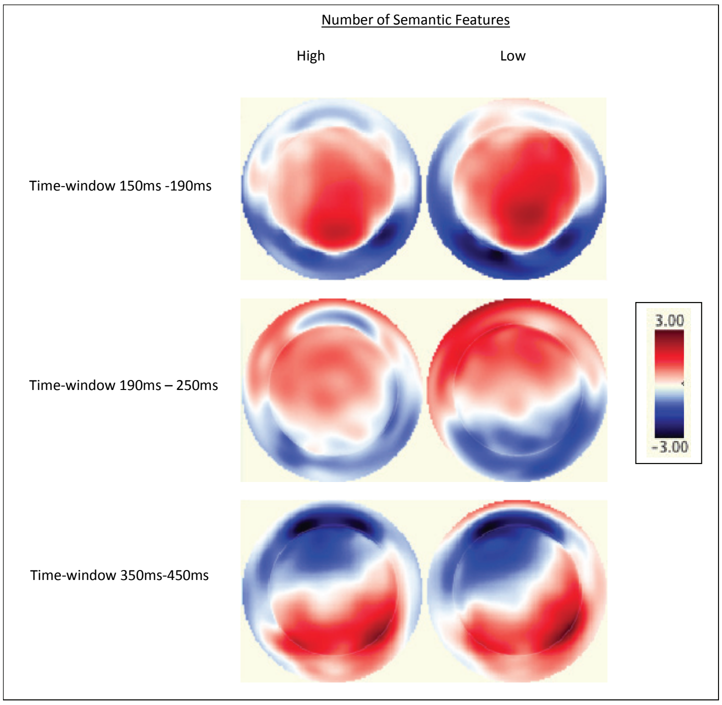

Figure 9.8. ERP topography voltage distribution plots for Experiment 6. The data are displayed for each semantic feature condition (high, low) for the three time-windows. Within each diagram, dorsal is top and ventral is bottom, with positive voltages depicted in red and negative voltages depicted in blue. Simple effects are presented in Sections 9.2.3.1.

Of particular interest to the aim of this thesis are semantic feature effects, specifically whether there are early semantic feature effects in the 150ms-190ms time-window prior to the phonological processing marker. Therefore simple effect analyses were performed to investigate the two significant interactions with semantic features (Howell, 1992). To reduce Type I error critical p-values were set using a Bonferroni correction (Howell, 1992); these critical values are reported with the statistics.

Simple effect analyses investigated semantic features and region interaction; the conditions of high and low in number of semantic features were compared at each region, critical $p = .025$. The average amplitude of high and low semantic feature conditions significantly differed in the occipital-parietal regions, $t(9) = 7.09$, SEM = 0.11, $p < .001$, but not in the fronto-temporal regions, $t(9) = 1.30$, SEM = 0.18, $p = .226$. High in semantic feature words have a greater mean amplitude than words low in semantic features. Figure 9.9 presents the interaction data.

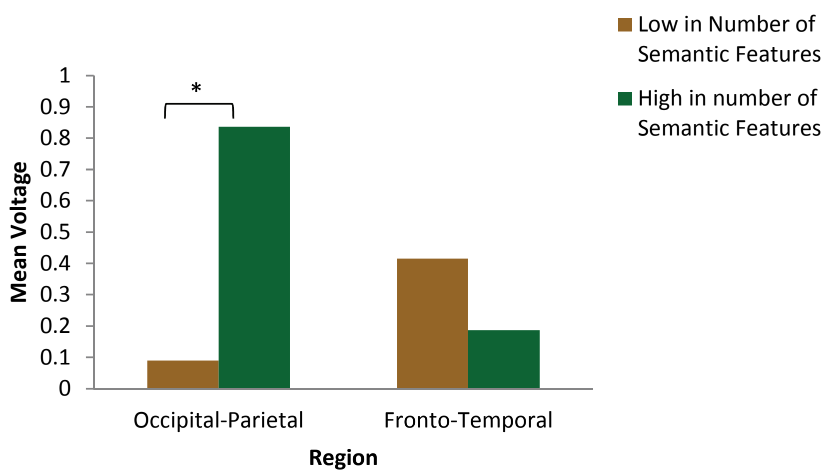

Figure 9.9. Average amplitude as a function of stimulus word semantic features and region in Experiment 6. The average amplitude of high and low semantic feature conditions significantly differed in the occipital-parietal regions, but not in the fronto-temporal regions, as is specified in the figure (*). See text for details.

Simple effect analyses were also performed to investigate the interaction of semantic features and time-window; the conditions of high and low in number of semantic features were compared at each time-window, critical $p = .017$. There was a significant difference between high and low imageability conditions in the 190ms-250ms time-window, $t(9) = 4.10$, SEM = 0.11, $p = .003$, but the difference was not significant at the other time-windows, 150ms-190ms, $t(9) = 1.53$, SEM = 0.10, $p = .16$; 350ms- 450ms $t(9) = 2.2$, SEM = 0.07, $p = .054$. Figure 9.10 presented the interactions data.

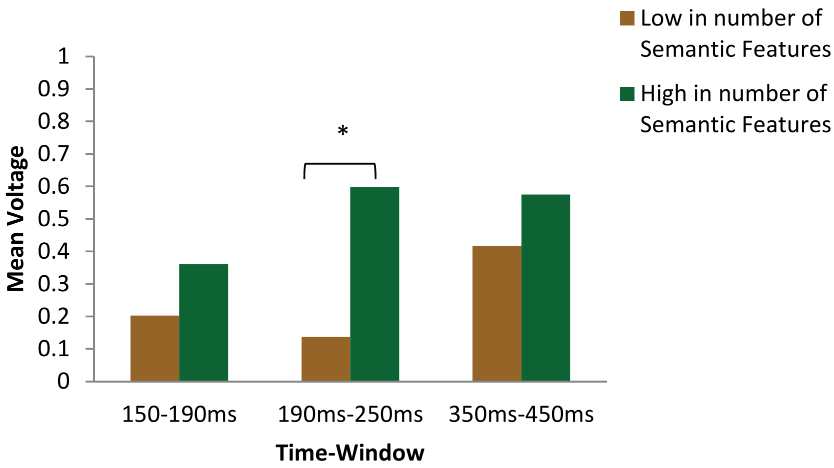

Figure 9.10. Semantic feature effects in average amplitude as a function of time-window in Experiment 6. There was a significant difference between high and low imageability conditions in the 190ms-250ms time-window, but the difference was not significant at the other time-windows, as is specified in the figure (*). See text for details.

9.3.2.2 Peak latency analyses

Peak latency ERP data was analysed as in Experiment 5 (9.2.2.2). Peak latency was analysed using a 2 x 2 x 2, i.e., Semantic Features (High, Low) x Region (Fronto-temporal, Occipital-Parietal) x Hemisphere (Right, Left) ANOVA at each of the three time-windows. There were no significant main effects or interactions involving semantic features. There were also no significant effects in the planned comparisons of the two semantic feature conditions in the earliest time-window (150ms-190ms). There were significant effects of topographic location of voltages. In the 150ms-190ms time-window there was a significant main effect of hemisphere, $F(1,9) = 8.61$, MSe = 206.28,

$p = 0.02$; the left hemisphere had a significantly later peak amplitude than the right hemisphere. In this time-window there was a significant interaction of region and hemisphere, $F(1,9) = 8.39$, MSe = 63.23, $p = 0.02$. In the 190ms-250ms time-window, the main effect of hemisphere approached significance, $F(1,9) = 4.41$, MSe = 177.14, $p = 0.07$; the means suggest that the right hemisphere had a later peak latency than the left hemisphere. In this time-window, there was also significant main effect of region, $F(1,9) = 6.93$, MSe = 309.12, $p = 0.03$; peak latency was later in the front-temporal region than in the occipital-parietal region. There were no other significant main effects or significant interactions, including the results of the analysis of time-window 350ms-450ms.

9.3.3 Discussion

Experiment 6 investigated the ERP correlates of the variable semantic features in low frequency silent word reading. Of interest is whether ERP correlates differ for words that have a high number of semantic features and words that have a low number of semantic features. Specifically of interest in the current investigations is whether semantic feature effects are found using a silent reading task and whether there are significant semantic effects in the "early" time-window (150ms-190ms), prior to phonological processing (Section 9.1.2). As a reminder, previous investigations using ERP measures and semantic features as a semantic measure have suggested that there may be early semantic effects. However as presented in section 9.1.3, these effects were not investigated in time-windows prior to 200ms, and the task used (judging semantic relatedness of sequential items using a button pres) was not word reading and may have biased word processing via semantic memory (Kounios et al., 2009).

Significant semantic features effects were found in analyses of mean amplitude. There was a significant difference between words high in number of features and words low in number of features with the former having a higher mean amplitude than the later. Differences in mean amplitude suggest that these two semantic conditions are processed differently within the brain, capturing semantic information activation during a word reading task. In Experiment 6, there were no significant semantic effects in peak latency analyses.

Significant interactions and simple effect analyses of mean amplitude revealed additional information about semantic feature effects. The simple effect analyses of the interaction of semantic features and region showed that the semantic effect was significant in occipital-parietal topography regions, but not in fronto-temporal topography regions. As noted in the discussion of imageability effects (Section 9.2.3), statistical interactions of semantic measures with region highlight processing differences between the two semantic feature conditions, but do not offer additional evidence as to the cortical sources of the topographical differences.

Semantic effects in Experiment 6 significantly interacted with time-window. Simple effects showed that semantic effects were limited to the second time-window (190ms-250ms). This time-window occurs simultaneous with the phonological processing marker of 200ms, and not prior to it. This time-window effect replicates the results of Kounios et al. (2009), who found ERP amplitude differences between semantic feature conditions in a 200ms-300ms time-window; though they did not investigate an earlier time-window and they did not use word reading.

Semantic effects found simultaneous with phonological processing could indicate a semantic contribution to word reading. The triangle model of word reading has clean-up units so that phonology and semantic memory reverberate to settle phonology (Section 1.3.3; Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989). Therefore, within this model semantic effects simultaneous to the phonological processing marker could indicate this “clean-up” process. The semantic effect in the second time- window (190ms-250m) might also indicate that it is possible for semantic information to be activated during phonology computation. These results are discussed further and in conjunction with the results Experiment 5 in the next section.

9.4 General discussion of ERP Experiments 5 and 6

Experiments 5 and 6 used ERP methods to investigate the time course of semantic measures during silent reading of low frequency words, as early semantic effects found prior to phonological processing (Section 9.1.2) may provide converging evidence of a semantic contribution to orthography-to-phonology computation, as suggested by the previous investigations of this thesis. Each ERP experiment used a different semantic measure, imageability and semantic features, respectively, to further investigate the effects of these measures on word reading following the conclusions of Chapter 8. Experiment 5 used low frequency words in a two-by-two manipulation of regularity and imageability. Experiment 6 used low frequency words that were manipulated on number of semantic features. Both experiments used a silent word reading task. As suggested in Chapter 8, semantic features and imageability may measure somewhat similar, but not entirely identical, semantic information.

There were significant statistical effects at $p = .05$ with the semantic measures in both experiments (imageability and semantic features, respectively) in mean amplitude analyses. In both experiments there were significant main effects of the semantic measures with words high in the semantic measure having greater mean amplitudes than words low in the semantic measure. Experiment 5 did not reveal any clear interaction of word regularity type, and there were no interactions of word regularity type and time-window. The significant interaction of the semantic measures and region was consistent across Experiments 5 and 6, with semantic effects in occipital-parietal regions, but not fronto-temporal regions, present in both experiments. These results show that words high or low in a semantic measures are processed differently within the brain. In both experiments, semantic effects occurred during silent low frequency word reading. However, as there were also significant interactions between the semantic measures and time-window in both experiments, of particular interest is whether either or both experiments exhibit semantic effects prior to the phonological marker of 200ms (Section 9.1.2).

The significant simple effects in earlier time-windows found in Experiments 5 and 6 indicate that semantic effects may occur prior to the completion of phonological processing. Moreover, the effects found in Experiments 5 and 6 are earlier than the "semantic" N400 component (Kutas & Hillyard, 1980; Sections 9.1.1.1 and 9.1.3). In Experiment 5, simple effects in the interaction of imageability, time-window and hemisphere showed a significant semantic effect in the time-window of 150ms-190ms in the topographical left hemisphere. This was not replicated in Experiment 6; the semantic measure significantly interacted with time-window, but not with hemisphere, and the effect was in the slightly later time-window of 190ms-250ms, but not the earlier time-window of 150ms-190ms. Therefore semantic effects as measured by imageability occurred in an earlier time-window than the semantic effects as measured by semantic features, and this is discussed further in the Final Discussion Chapter (Chapter 10). Considering the results of Experiments 5 and 6 and the phonological processing literature presented in Section 9.1.2, it seems unlikely that the semantic effects in the two earlier time-windows (150ms-190ms and 190ms-250ms) occur *after* the completion of phonological processing.

The simple effects with the semantic measures in the earlier time-windows (150ms-190ms; 190ms-250ms) in Experiments 5 and 6 seem to occur either prior to phonological processing (Experiment 5) or simultaneous with phonological processing (Experiment 6), though in both cases the semantic effects would seem to occur prior to the completion of phonological processing, as indicated by the phonological marker used here (Section 9.1.2). Dien (2009) claims phonological processing occurs from 200ms-500ms, and 200ms is used in these experiments as maker of phonological processing, as it likely begins at this time point. Therefore phonological processing is likely not completed when the significant mean amplitude effects of Experiment 6 occur in the time-window of 190ms-250ms. Concurrent semantic and phonological effects are possible within the triangle model of word reading (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989; Strain et al., 1995; Sections 1.3.3 and 9.3.3) as the two systems may work together to settle phonology with the use of clean-up units. The semantic effects observed in Experiments 5 and 6 provide evidence of semantic information activation early in the time-course of silent word reading as they likely occur prior to the *completion* of phonological processing. Moreover, semantic activation early in the time-course of word reading has the potential to contribute to orthography-to-phonology computation.

Of note are confounds of imageability with familiarity and age-of-acquisition in Experiment 5, though word regularity types are matched on these factors (Table 9.1). It is possible that the semantic imageability effects in Experiment 5 are due to age-of-acquisition or familiarity differences in the two conditions, not differences in imageability. However the results of the regression analyses in Chapter 8 reveal that imageability effects are still present even when significant age-of-acquisition effects are statistically controlled for (See Section 8.4.2). Likewise in the analyses of Chapter 8, imageability effects are also still present when familiarity is statistically controlled for. Therefore, using the results of the regression analyses in Chapter 8 as a guide, the imageability effects in Experiment 5 are not likely, wholly due to age-of-acquisition effects or familiarity effects as each factor likely has its own unique influence. These same factors of familiarity and age-of-acquisition are not confounded in Experiment 6 (high and low semantic feature conditions were matched on age-of-acquisition and familiarity, in addition to other factors) and a semantic effect is found prior to the completion of phonological processing, as defined by the literature. Additionally, age-of-acquisition is interpreted as being, at least in part, semantic in nature, (Section 8.4.2.1). Therefore the imageability effects found in Experiment 5 are likely a genuine semantic effect early in the time-course of word reading and not solely due, instead, to un-matched factors. The measures of imageability and age-of-acquisition and their relationship with one another are discussed further in Chapter 10.

In addition to the semantic effects in early time-windows of Experiments 5 and 6, in Experiment 5, there was also a semantic effect in peak latency analyses in the late time-window (350ms-450ms), with high imageability words peaking later than low imageability words. This was the only significant semantic effect in the peak latency analyses in both Experiments, and this was the only semantic effect specifically in this late time-window (350ms-450ms) in both Experiments. As these effects are after the phonological processing marker and likely after the completion of phonological processing, at least by reference to other literature (Section 9.1.2), it is unlikely that these semantic effects indicate semantic contribution prior to the completion of orthography-to-phonology computation. This later semantic effect in the time-window 350ms-450ms, however, may be evidence of semantic information being used again, possibly using phonology-to-semantic connections as part of the comprehension process (Hauk, Davis, et al., 2006), which would be consistent with recent interpretations of N400 effects (Barber & Kutas, 2007; Hinojosa et al., 2001). As these later effects were not found in Experiment 6 with semantic features, it may provide further evidence that the information captured by semantic features is not wholly subsumed by imageability, and that semantic feature information might be used for orthography-to-phonology computation, but does not benefit later comprehension processes.

Finally, within these experiments a marker of phonological processing has been defined by the literature. Therefore future experiments aim to replicate these semantic effects early in the time-course of word reading while employing a phonological marker within the experiment themselves, by including manipulations of semantic measures and phonological measures and matched low frequency words.

9.5 Conclusions

The experiments described above aimed to investigate the neurocorrelates of semantic measures imageability and semantic features using a word reading task prior to phonological processing completion, using ERP measures to obtain converging evidence for a possible semantic contribution to word reading. The semantic effects in both experiments presented here indicate that there are early semantic effects during silent word reading of low frequency words. Experiments 5 and 6 replicated the earlier semantic effects published in the literature (Abdullev & Posner, 1998; Barber & Kutas, 2007; Hauk, Davis, et al., 2006; Kounios, et al., 2009; Landi & Perfetti, 2007) using a different task (silent word reading) and different semantic variable (imageability and semantic features). Experiments 5 showed semantic effects in the time-window of 150ms-190ms, and Experiment 6 showed semantic effects in the time-window of 190ms-250ms. As semantic information is processed prior to or simultaneous with the phonological marker of 200ms, it is therefore possible that semantic information is used in orthography-to-phonology computation. Together with the behavioural investigations described in Chapters 3-8, the semantic effects in the ERP experiments presented here seem to provide converging evidence that semantic information may contribute to orthography-to-phonology computation. All investigations of this thesis are discussed further in the General Discussion Chapter that follows.

Chapter 10

General Discussion

10.1 Introduction to general discussion

The central aim of this thesis is to explore whether semantic information contributes to orthography-to-phonology computations in healthy adult word reading. This was investigated using behavioural and ERP measures, and factorial and regression designs. The three lines of investigation have individually provided evidence that suggests a semantic contribution to orthography-to-phonology computation. Together these investigations can improve the understanding of semantic information's role in healthy adult word reading and the simulation of this in current computational models. The research of this thesis can also aid in the understanding of semantic measures, imageability and semantic features. Furthermore, additional aims of this thesis were listed at the end of Chapter 2 and how these were met is highlighted in this chapter.

10.2 An overview of results

The semantic priming experiments (Chapters 3 to 7) addressed the specific additional aim (1) of investigating whether a semantic contribution is evident in word reading using long lag cross modal semantic priming paradigms and various word types as targets. A picture prime was followed either two trials later (Experiments 1, 3, and 4) or one trial later (Experiment 2) by a word target. The rationale for this paradigm was presented in Chapters 3 and 4, and detailed discussions of these experiments' results and additional across-experiment analyses were provided in Chapter 7. One advantage of the priming paradigm used in these studies is that primes and targets appeared in both related and unrelated conditions, and target reaction times are compared against themselves in related unrelated conditions, hence evidence for semantic effects within word type cannot readily be interpreted as due to uncontrolled factors. The conclusions drawn in Chapter 7 will be summarised here, with further discussion in relation to the subsequent regression and ERP investigations.

There were some inconsistencies in the data across priming experiments and the significant results in the main came from combining data across the four priming experiments or from limiting the analyses to include only certain stimuli; so, results from the priming experiments must be considered with some caution. The priming studies indicate that a semantic contribution to orthography-to-phonology computation is present in low frequency exception word reading and low frequency regular word reading. Low frequency exception target word reading was quicker when preceded by a related picture target in Experiments 1, 2, and 4, and low frequency regular target words were primed in Experiments 2, and 4. A semantic contribution to word reading cannot be eliminated for high frequency words, when semantic priming is used as a measure of semantic contribution (Experiment 2, Chapters 4 and 7).

It was concluded that priming might be influenced by strength of prime-target pair semantic relationship and block order. In particular, the data from the semantic priming experiments indicated that qualities of intervening filler items, such as number, category membership, modality -including modality pattern and modality of individual filler items- and the filler item itself, affect whether the prime's semantic activation survives for the target. Chapter 7 presented these conclusions in detail, and it was reasoned that the effects of filler qualities should be studied further in their own right.

The regression investigations investigated whether additional behavioural evidence of a semantic contribution to word reading could be found using single word reading (Chapter 8). These investigations addressed additional aim (2) by exploring whether semantic features and imageability were unique significant predictors of ELexicon single word reading reaction times for words that varied along a spectrum of frequency, but were on average low frequency, while statistically controlling for age-of-acquisition and other correlated measures. Semantic features were not a significant predictor of word reading times when entered into the analyses with imageability and/or age-of-acquisition. Imageability and age-of-acquisition were significant and unique predictors of word reading times in these analyses, suggesting that semantic information contributes to word reading with the set of words used in the regression analyses. Thus, both the priming studies and the regression analyses suggest that there is a semantic contribution to word reading, especially low frequency word reading, using different methods (factorial and regression), and different measures of semantic information (semantic priming, semantic measures imageability, and semantic features).

The two ERP experiments of this thesis investigated the neurocorrelates of the semantic variables imageability and semantic features, and their potential for an early semantic contribution to low frequency (silent) word reading, prior to phonological processing (as established in the literature), thus addressing additional aim (3). The results of ERP Experiments 5 and 6 suggested an indication of semantic processing in relation to imageability and semantic features early (150-190ms in Experiment 5 and 190-250ms in Experiment 6) in the time course of low frequency word reading, prior to the completion of phonological processing as defined in the literature. These early semantic effects are consistent with a possibility of a semantic contribution to low frequency word reading in healthy adults, providing converging evidence to the behavioural studies of this thesis. However phonological processing in this thesis' ERP experiments was provided by establishing a marker from the published literature of when phonological processing might occur. Future experiments should seek to replicate the semantic effects of Experiments 5 and 6 and also include a manipulation of a phonological variable, such as phonological neighbourhood, in order to mark the onset of phonological processing within the experiment itself.

The results from the three lines of investigation, together, lead to a conclusion that semantic information is involved in healthy word reading. The next section considers the computational model of word reading and to what extent they account for these results.

10.3 Computational models of word reading

Chapter 2 presented two groups of computational models of word reading; one group has not implemented semantic memory in the process of orthography-to-phonology computation and the other group has. The DRC model and CDP +and ++ models of word reading theoretically include semantic memory on the lexical route, but it is not implemented within the computational models (Section 1.3.2; Coltheart et al., 1993; Coltheart et al. 2001; Coltheart 2006a, 2006b; Perry et al., 2007; 2010b). Furthermore DRC modellers claim that semantic memory is not needed for orthography-to-phonology computation (Coltheart, 2006a, 2006b) (Section 1.3.2). Therefore to account for the evidence from this thesis' investigations of a semantic contribution to word reading, the DRC and CDP modellers, since the semantic memory component is not implemented in the model itself, would at least need claim that the theoretical semantic memory component of their respective models contributes to orthography-to-phonological processing, but at present they do not argue this (see Section 1.3.2).

The computational model of word reading that allows for a semantic memory contribution to orthography-to-phonology computation and that has implemented a semantic pathway, i.e., the connectionist triangle model (Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989; Plaut et al., 1996), can account for the results of this thesis. The triangle model proposes that orthography, phonology, *and* semantic information are used to compute phonology; moreover semantic memory has been implemented in this model (see Section 1.3.3; Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989). Within this model, the orthography-to-semantics-to-phonology pathway and the orthography-to-phonology pathway can work together to settle phonology (Figure 1.2) (Harm &Seidenberg, 2004; Plaut et al., 1996). Semantic information can potentially contribute to the reading of all words as semantic memory is connected to orthography and phonology, but semantic memory is most likely to contribute to phonological computation when the orthography-to-phonology pathway is slow and less efficient. When this phonological pathway is slow, the orthography-to-semantics-to-phonology pathway and/or the "clean-up" units of the orthography-to-phonology-to-semantics-to-phonology would have time to contribute to the computation of phonology. The data from the three lines of investigation will now be considered further in order to comment on a semantic contribution to the various word types, and how this might be accounted for within the triangle model of word reading.

10.4 Semantic effects in relation to word types

10.4.1 Semantic contribution to various word types and the triangle model of word reading

In Chapter 2, it was detailed that the investigations of this thesis would provide information on semantic effects in relation to “word types” commonly investigated in the literature (Section 1.2.3), and implicated in word reading models accounts (Section 1.3). By considering the four target word types in the priming experiments, a spectrum of words (average low frequency) in the regression analyses, and the low frequency words in the ERP experiments, comment can be made as to whether this thesis’ investigations show a semantic contribution for only low frequency exception words, as might expected when considering the results of some behavioural literature (see Section 2.3.2; Shibahara et al., 2003; Strain et al., 1995, 2002; Woollams, 2005) or whether there is any evidence for semantic activation during the reading of other word types. In short, as indicated above, the triangle model can account for a semantic contribution whenever orthography-to-phonology computation might be slow.

Previous research has found evidence of a semantic contribution to low frequency exception words reading (Strain & Herman, 1999; Strain et al., 1995, 2002; Shibahara et al., 2003; Woollams, 2005), and the results of thesis are consistent with this. Low frequency exception words showed some evidence of priming in Experiments 1, 2, and 4, and showed semantic effects early in the time course of word reading in ERP

Experiment 5. Additionally, though the majority of words in the regression analyses and ERP Experiment 6 were low frequency regular, these investigations may also support a semantic contribution to low frequency exception word reading as there were a few stimuli of this type in these investigations. As orthography-to-phonology computation is less efficient for low frequency exception words, these findings are consistent with the triangle model.

The investigations of this thesis have provided an additional indication of a semantic contribution to low frequency regular word reading, as priming was also suggested in Experiments 2 and 4 and also showed evidence of a semantic effect early in the time course of word reading in ERP Experiment 5, i.e. there was no interaction of the semantic manipulation with word type. Likewise, the suggestion that semantic information contributes to low frequency regular words is evidenced by the results of the regression analyses, which used words that varied along a spectrum of frequency, but that were low frequency on average and in the main were regular in spelling-to-sound correspondence. It is also supported by the results of ERP experiment 6, which used matched low frequency words not manipulated on regularity type.

A contribution to low frequency regular word reading was initially considered in Chapter 3, in relation to the results of Experiment 2 of Vitkovitch et al. (2006), in which the majority of the primed target words were low frequency regular. However, as the target words were not manipulated on word type and were not all one word type, it could not be firmly concluded that the significant priming effect in Experiment 2 of Vitkovitch et al. indicated a semantic contribution to low frequency regular word reading. Therefore the priming experiments of this thesis explored this further. A semantic contribution to low frequency regular word reading is consistent with regression analyses of Balota et al. (2004), Pexman et al. (2002) and Cortese and Khanna (2007) that concluded that semantic measures within their analyses were significant predictors of single word reading times for a spectrum of words, including low frequency regular words. However, these three regression analyses had their limitations of not including imageability, semantic features, and age-of-acquisition together in the analyses. This was a contribution of the investigations in the regression analysis of Chapter 8. Thus, the current results point to evidence of a semantic contribution to low frequency regular word reading, not only low frequency exception word reading.

A possible explanation as to why semantic effects with low frequency regular words have been found here may be due to the very low frequency of the this word type. The low frequency words used in this thesis, both regular and exception words, were very low frequency, under 16 per million in Kucera and Francis frequency (1967) and under14 per million in Celex frequency; this is lower than the threshold in other published investigations (Andrew, 1982; Andrews et al., 2005; Monaghan & Ellis, 2002; Strain et al., 2002; Section 1.4.2). Hence within the account of the triangle model, these words, though regular, would have less efficient orthography-to-phonology computations because of their very low frequency.

In the experiments of this thesis semantic effects with high frequency word types were also found in Experiment 2. Priming of high frequency target words was only found over one intervening item (Experiment 2) , but not two (Experiments 3 and 4), and it was argued that this may have been due to the strength of semantic relationship between prime and target and filler qualities (see Chapter 7). Possible effects of prime-target relationship and filler effects warrant further investigations, including systematic manipulations of relationship and semantic strength as this could be an important factor effecting whether semantic priming is found (see Chapter 7). Nevertheless, evidence for high frequency priming was evident, and this needs to be considered in relation to other empirical and theoretical accounts, as follows.

The ERP studies, which are sensitive to small time-course effects that may not be seen in behavioural measures, only used low frequency words. The results of the regression analyses do not eliminate a semantic contribution to high frequency word reading as 20% of the stimulus words in these analyses met the high frequency threshold used in the priming Experiments. It also included 25% of the high frequency exception words and 55% of the high frequency regular words from the priming experiments. However, the average frequency for all of the stimulus words in the regression was low. Additionally the interaction of imageability and frequency in the regression analyses was not significant; therefore there was no indication that the significant imageability effects were limited to the low frequency words of this analysis. However, as word stimuli choice was constrained, e.g., word stimuli for the regression analyses were chosen based on availability of semantic feature measures, and word targets for the priming study were required to have a category coordinate in picture modality, this may limit the extent to which the results can be generalised to other word samples, including other high frequency words that were not used in these investigations. Though semantic effects may not be expected in high frequency healthy word reading and may not readily be generalised, there has been some indication, in the results of Experiment 2 of this thesis and also possibly in the regression analyses, that it is possible to find a semantic contribution with high frequency word reading. An indication of a semantic contribution to high frequency word reading has also been found in the research with impaired reading in people with semantic dementia (Woollams et al., 2007). These results suggest that a semantic contribution to high frequency word reading is not impossible, even though, according to the triangle model, it would be highly unlikely (Harm & Seidenberg, 2004).

So, how should the evidence for semantic effects with high frequency word reading be accounted for within the triangle model? One possibility is individual difference of the readers. Though high frequency words ordinarily have efficient orthography-to-phonology computation, it is possible that if the participants are lower skilled readers, then their orthography-to-phonology computation may be slow, creating a greater reliance on the pathways that provide a semantic contribution, e.g. orthography-to-semantics-to-phonology. Strain & Herdman (1999) investigated this, and data from lower skilled readers showed a main effect of imageability, but no interaction with regularity when reading low frequency words (see Section 2.1.2.2). The possibility of individual differences was considered with the current semantic effects in high frequency, but the investigations that showed a semantic contribution (Experiment 2 and the regression analyses) used healthy adult participants from different samples[52]. It is unlikely that all of these samples have less efficient orthography-to-phonology due to lower reading skill.

The results of some investigations of this thesis suggest a semantic contribution to high frequency word reading with healthy adults, and this is not expected by the triangle model (Harm & Seidenberg, 2004) and is difficult to explain. It is of worth to further investigate these results, seeking to replicate them, as the results of this thesis indicate a

[52] As a reminder, participants of Experiment 2 were from the University of East London. The regression analyses used reaction times from the ELexicon project, which were Z-score reaction times from samples from many universities.

semantic contribution to high frequency word reading may be possible. Previous investigations have focused on low frequency word reading in healthy adults with the English language, however future investigations should research a possible semantic contribution to high frequency word reading. Additionally, ERP measures may be an appropriate method as it is sensitive to effects early in the time course, prior to phonological completion.

10.4.2 Variables underlying "easy" and "difficult" word types

As concluded previously, the priming investigations of this thesis show evidence in support of a semantic contribution to word reading, with a contribution to low frequency word types, and also some suggestion of a contribution to high frequency word types as well. However, of interest now are the factors underlying these various words types that have shown some evidence of a semantic contribution. A factorial manipulation of frequency and spelling-to-sound correspondence created four word types, which ranged from easy to difficult to read: high frequency regular, high frequency exception, low frequency regular, and low frequency exception (Section 1.2.3). These four word types were used as target words in the series of priming experiments. However, as noted in Chapters 1, 4, and 7 variables tend to be correlated, such as frequency and age-of-acquisition, so that more than one variable can contribute to the ease with which a word is read (Balota et al., 2004). As acknowledged by other researchers (e.g., Balota et al., 2004; Cortese & Khanna, 2007; Ellis & Monaghan, 2002, Monaghan & Ellis, 2002; Strain et al., 2002), it can be difficult to obtain pure manipulation of a single variable, with age-of-acquisition being particularly difficult to control. As noted in the priming studies (Chapter 4 and Chapter 7), though stimuli were controlled as far as was possible, across low frequency regular and exception conditions there was a small difference in length in letters, and also in age-of-acquisition. Age-of-acquisition also differed across high and low frequency words (as did familiarity). However, these variables were confounded in such a way to make “difficult” words more difficult, i.e., low frequency exception words had a high (older) age-of-acquisition and were very slightly longer than their regular counterparts. High frequency words in the priming studies had a higher familiarity rating and lower (younger) age-of-acquisition rating than their low frequency counterparts. In this series of priming experiments, therefore, it is relevant to consider to what extent these other variables contribute to the priming results for “easy” and “difficult” words. Note, though, that in the behavioural regression study when these other variables, including familiarity, length, and age-of-acquisition, were statistically controlled, evidence for a semantic contribution to word reading was clear as imageability still significantly predicted word reading times. Similarly, in ERP Experiment 6, stimulus words were well matched on familiarity, length, and age-of-acquisition, amongst other factors, and there were still semantic effects early in the time course of low frequency word reading (Chapter 9).

Firstly, length effects are not likely to account for the “easy” and “difficult” word types of the priming studies. Though the word types significantly differed in length in letters, the actual numerical difference is negligible. Between low frequency regular and low frequency exception words the difference on average is slightly over half a letter, 4.45 and 5.05 letters respectively, and between low frequency and high frequency it is less than this, 4.75 and 4.45 letters respectively. The difference is small, and the words are short, therefore letter effects may not have had much consequence for these words (Damian et al., 2010; Young & Ellis, 1985).

Age-of-acquisition was confounded with frequency and spelling-to-sound correspondence in the target words of the priming experiment, i.e., word manipulations of frequency and regularity were also unintentional manipulations of age-of-acquisition. Therefore the "easy" words of the priming studies were in fact high frequency regular and low (younger) age-of-acquisition words, with age-of acquisition contributing to the easy with which these words are read. As a semantic contribution to high frequency words is possible, as suggested by the priming effect of Experiment 2, further investigations of the words that elicited this priming effect are warranted. It may be possible that this confound of age-of-acquisition could play a role in exposing semantic effects in priming paradigms. Therefore further study of easy and difficult word types is merited, to establish what effect age-of-acquisition may have in eliciting semantic effects in word reading.

10.5 Comments on semantic measures imageability and semantic features

The regression analyses and ERP experiments of this thesis provide information regarding the semantic measures of imageability and semantic features used in both experiments. The regression analyses (Chapter 8) suggested that imageability and semantic features are similar semantic measures. Imageability was a unique predictor of word reading times, whereas semantic features, when entered into an analysis with imageability failed to remain significant. This may indicate that imageability wholly subsumes semantic features, while the opposite is not true (Section 8.4.2).

The result of the ERP experiments, however, may offer differing evidence. ERP Experiment 6 in which conditions were well matched on a number of factors, including imageability and age-of-acquisition, showed semantic feature effects early in the time course of word reading. So, neurophsyiological measures offer evidence of semantic feature effects when imageability is controlled, but behavioural measures do not. The results from the two ERP studies may even indicate more about semantic features and imageability.

As has been shown previously, confounds of imageability and age-of-acquisition are difficult to control in factorial manipulations (Monaghan & Ellis, 2002, 2002; Strain et al., 1995, 2002), and, though stimuli were matched as far as was possible, the stimuli of ERP Experiment 5 were confounded on these two measures (higher imageability words were learned earlier in life). (As discussed in Chapter 9 imageability and familiarity were also confounded, but this is not discussed again here). Therefore, as mentioned in Chapter 9, it is possible that the imageability effects of Experiment 5 are attributable to age-of-acquisition. As detailed in Sections 8.4.2.1 and Section 9.4, this thesis does not support a purely non-semantic interpretation of age-of-acquisition. The regression analyses' results indicate that age-of-acquisition is at least partly semantic in nature as age-of-acquisition is significantly correlated with semantic features and imageability, and in an additional analysis age-of-acquisition subsumed the contribution of semantic features. Therefore, imageability effects with age-of-acquisition confounds early in the time-course of word reading may also be indicative of a semantic contribution to word reading.

Though both ERP Experiments 5 and 6 show semantic effects early in the time-course of word reading either prior to or simultaneous with the phonological processing marker, the effects occur in different time-windows in the two Experiments (Chapter 9). Experiment 5 shows an imageability effect and no interaction with regularity, in the

150ms-190ms time-window, whereas Experiment 6 showed a semantic features effect in the 190ms-250ms time-window, slightly later than the effect of Experiment 5. Also the semantic effect in the 350ms-450ms time-window, as was present in Experiment 5, was not found in Experiment 6. It is possible that imageability is active earlier in the word reading time-course than semantic features because of the type of semantic information it measures, such as inherent meaning or connections to other parts of the brain (Section 2.3.1). Another possibility is that the earlier effect in Experiment 5 may also be due to the age-of-acquisition confound, and that if imageability conditions were matched on age-of-acquisition (and familiarity) then perhaps the effect might be in line with the effect of Experiment 6 at 190ms-250ms. Therefore a replication of Experiment 5 is needed while also controlling age-of-acquisition. However it is also possible that the difference in time-window for the two semantic effects provides evidence that imageability does not fully subsume semantic features as a measure.

As described in Chapter 9, ERP measures are sensitive to time-course and may provide detailed information of cortical processes during word reading that are not provided by behavioural measures. Contrary to previous behavioural evidence in the regression analyses that imageability subsumes semantic feature effects, using the more sensitive neurophsyiological ERP measure, evidence from Experiment 6 alone, and also possibly from a comparison of Experiments 5 and 6, indicate that imageability may not completely account for semantic feature effects. Therefore, these two measures represent different aspects of semantic knowledge.

10.6 General Conclusions

This thesis has explored orthography-to-phonology computation in healthy adults, and in particular whether semantic information contributes to this process. This aim has been investigated using a variety of techniques. The studies used behavioural and ERP measures, factorial and regression methods, with tasks of overt reading and covert reading, and semantic priming and single word reading paradigms. The investigations of this thesis have provided information concerning the semantic measures of imageability and semantic features, and support a semantic interpretation of the measure age-of-acquisition. Additional information about a semantic contribution to various word types has also been provided. Evidence of a semantic contribution to both low frequency regular and exception low frequency word types was provided by Experiments 1, 2, 4, 5, and 6 and the regression investigations. There is also evidence of a semantic contribution to high frequency words as provided by Experiment 2, and not eliminated by the regression analyses, as this contained some high frequency words as well. In addition to frequency and regularity, the potential of other relevant variables types, such as age-of-acquisition, to contribute to the forming of “easy” and “difficult” words was considered. The results of this thesis have also provided comment on computational models of word reading. It concluded that the results would be best accounted for by the connectionist triangle model of word reading that includes and implements semantic memory in the process of orthography-to-phonology computation. Future research has also been noted, including further investigations into intervening filler items in semantic priming designs (in relation to distributed memory accounts), semantic effects using ERP measures while including a marker of phonological processing within the experiment itself, and a semantic contribution to high frequency word reading. In conclusion, the investigations of this thesis found evidence of a semantic contribution to orthography-to-phonology computations.

References

Abdullaev, Y. G., & Posner, M. I. (1998). Event-related brain potential imaging of semantic encoding during processing single words. *Neuroimage, 7*(1), 1-13.

Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent. *Memory and Cognition, 10*(6), 565-575.

Andrews, S., Woollams, A., & Bond, R. (2005). Spelling-sound typicality only affects words with digraphs: Further qualifications to the generality of the regularity effect on word naming. *Journal of Memory and Language, 53*(4), 567-593.

Ashby, J., Sanders, L. D., & Kingston, J. (2009). Skilled readers begin processing sub-phonemic features by 80 ms during visual word recognition: Evidence from ERPs. *Biological Psychology, 80*(1), 84-94.

Baayen, R.H., Piepenbrock R., & Gulikers,L. (1995). The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning Memory and Cognition, 14*(4), 579-589.

Bajo, M.-T., Canas, J. J. (1989). Phonetic and semantic activation during picture and word naming. *Acta Psychologica, 72*(2), 105-115.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word frequency effects in the pronunciation task : Lexical access and or production. *Journal of Memory and Language, 24*(1), 89-106.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283-316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews, 53*(1), 98-123.

Baron, J., & Strawson, C. (1976). Use of orthographic and word specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 386-393.

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1059-1082.

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: evidence for a lexical-semantic competition hypothesis. *Cognition, 96*(2), 45-54.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology Section A, 58*(4), 667-692.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11*(3), 235-260.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods Instruments & Computers, 33*(1), 73-79.

Blazely, A. M., Coltheart, M., & Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: Implications for models of reading. *Cognitive Neuropsychology, 22*(6), 695-717.

Boniface, D. R.(1995). *Experiment design and statistical methods for behavioural and social research*. London: Chapman and Hall.

Brown, G. D. A. (1984). A frequency count of 190,000 words in the *London-Lund Corpus of English Conversation. Journal of Behavior Research Methods, Instruments, & Computers,* 16, 502-532.

Brown, G. D. A., & Watson, F. L. (1987). 1st in, 1st out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition, 15*(3), 208-216.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7-8), 992-1011.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104*(2), 215-226.

Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory and Cognition, 7*(2), 68-76.

Carreiras, M., Mechelli, A., Estevez, A., & Price, C. J. (2007). Brain activation for lexical decision and reading aloud: two sides of the same coin? *Journal of Cognitive Neuroscience, 19*(3), 433-444.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 757-777.

Coles, M. G. H., & Rugg, M. D. (1996). Event-related brain potentials: An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of min:d. Event-related brain potentials and cognition* (pp. 1-23). New York: Oxford University Press.

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

Coltheart, M. (2000). Deep dyslexia is right-hemisphere reading. *Brain and Language, 71*(2), 299-309.

Coltheart, M. (2004). Are there lexicons? *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 57*(7), 1153-1171.

Coltheart, M. (2006a). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology, 23*(1), 96-109.

Coltheart, M. (2006b). Dual route and connectionist models of reading: An overview. *London Review of Education, 4*(1), 5-17.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589-608.

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

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*(1), 204-256.

Coltheart, M., Saunders, S. J., & Tree, J. J. (2010). Computational modelling of the effects of semantic dementia on visual word recognition. *Cognitive Neuropsychology, 27*(2), 101-114.

Coltheart, M., Tree, J. J., & Saunders, S. J. (2010). Computational modelling of reading in semantic dementia: comment on Woollams, Lambon Ralph, Plaut, and Patterson (2007). *Psychological Review, 117*(1), 256-271; discussion 271-252.

Cortese, M., & Fugett, A. (2004). Imageability Ratings for 3,000 Monosyllabic Words. *Behavior Research Methods, Instruments, and Computers*, 36, 384–387.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology, 60*(8), 1072-1082.

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods, 40*(3), 791-794.

Cortese, M. J., & Simpson, G. B. (2000). Regularity effects in word naming: What are they? *Memory and Cognition, 28*(8), 1269-1276.

Cortese, M. J., Simpson, G. B., & Woolsey, S. (1997). Effects of association and imageability on phonological mapping. *Psychonomic Bulletin and Review, 4*(2), 226-231.

Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1372-1384.

Damian, M. F., Bowers, J. S., Stadthagen-Gonzalez, H., & Spalek, K. (2010). Does word length affect speech onset latencies when producing single words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(4), 892-905.

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

Deacon, D., Grose-Fifer, J., Hewitt, S., Nagata, M., Shelley-Tremblay, J., & Yang, C. M. (2004). Physiological evidence that a masked unrelated intervening item disrupts semantic priming: Implications for theories of semantic representation and retrieval models of semantic priming. *Brain and Language, 89*(1), 38-46.

Deacon, D., Hewitt, S., & Tamny, T. (1998). Event-related potential indices of semantic priming following an unrelated intervening item. *Cognitive Brain Research, 6*(3), 219-225.

DeGroot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning Memory and Cognition, 15*(5), 824-845.

Dien, J. (2009). The neurocognitive basis of reading single words as seen through early latency ERPs: A model of converging pathways. *Biological Psychology, 80*(1), 10-22.

Ellis, A. W., Burani, C., Izura, C., Bromiley, A., & Venneri, A. (2006). Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *Neuroimage, 33*(3), 958-968.

Ellis, A. W., & Monaghan, J. (2002). Reply to Strain, Patterson, and Seidenberg (2002). *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(1), 215-220.

Eulitz, C., Hauk, O., & Cohen, R. (2000). Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming. *Clinical Neurophysiology, 111*(11), 2088-2097.

Fabiani, M., Gratton, G., & Coles, M. G. (2000). Event-related potentials. *Handbook of psychophysiology*, 2, 53-84.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Field, A. P. (2000). *Discovering statistics using SPSS for Windows: Advanced techniques for the beginner*. London, UK: Sage.

Ford, M., Marslen-Wilson, W. D., Davis, M. H., 2003. Morphology and frequency: Contrasting methodologies. In Baayen, H., Schreuder, R. (Eds.), *Morphological Structure in Language Processing.* Mouton de Gruyter, Berlin.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 361-379.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 104-115.

Fushimi, T., Ijuin, M., Patterson, K., & Tatsumi, I. F. (1999). Consistency, frequency, and lexicality effects in naming Japanese Kanji. *Journal of Experimental Psychology: Human Perception and Performance, 25*(2), 382-407.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-Acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 Words. *Behavior Research Methods & Instrumentation, 12*(4), 395-427.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in stroop-like word and picture-processing. *Journal of Experimental Psychology-General, 118*(1), 13-42.

Glushko, R. J. (1979). Organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*(4), 674-691.

Graham, K. S., Hodges, J. R., & Patterson, K. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia, 32*(3), 299-316.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science, 17*(12), 1021-1026.

Handy, T. C. (2005). *Event-related potentials: A methods handbook*. Cambridge, MA: MIT Press.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*(3), 491-528.

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*(3), 662-720.

Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383-1400.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermuller, F., & Rogers, T. T. (2006). [Q:] When would you prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience, 18*(5), 818-832.

Hauk, O., & Pulvermuller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115*(5), 1090-1103.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638-650.

Hines, D., Czerwinski, M., Sawyer, P. K., Dwyer,M. (1986). Automatic semantic priming: Effect of category exemplar level and word association level. *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 370-379.

Hinojosa, J. A., Martin-Loeches, M., & Rubia, F. J. (2001). Event-related potentials and semantics: An overview and an integrative proposal. *Brain and Language, 78*(1), 128-139.

Hinojosa, J. A., Martin-Loeches, M., Munoz, F., Casado, P., & Pozo, M. A. (2004). Electrophysiological evidence of automatic early semantic processing. *Brain and Language, 88*(1), 39-46.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98*(1), 74-95.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia - progressive fluent aphasia with temporal-lobe atrophy. *Brain, 115*, 1783-1806.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100*(3), 464-482.

Howell, D. C. (1992). *Statistical Methods for Psychology* (3rd ed.). Belmont, CA: Duxbury Press.

Howell, D. C. (1997). *Statistical Methods for Psychology* (4th ed.). Belmont, CA: Duxbury Press.

Humphreys, G. W., Lamote, C., & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory, 3*(3-4), 535-586.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5*(1), 67-103.

Jared, D., Mcrae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*(6), 687-715.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36*(4), 505-529.

Jefferies, E., Ralph, M. A. L., Jones, R., Bateman, D., & Patterson, K. (2004). Surface dyslexia in semantic dementia: A comparison of the influence of consistency and regularity. *Neurocase, 10*(4), 290-299.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13*(7-8), 789-845.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning Memory and Cognition, 23*(5), 1083-1105.

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(3), 483-491.

Kello, C. T. (2006). Considering the junction model of lexical processing. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 50–75). New York, NY: Psychology Press.

Kello, C.T. and Plaut, D.C. (2003) Strategic control over rate of processing in word reading: a computational investigation. *Journal of Memory and Language, 48*, 207–232.

Khateb, A., Annoni, J. M., Landis, T., Pegna, A. J., Custodi, M. C., Fonteneau, E., et al. (1999). Spatio-temporal analysis of electric brain activity during semantic and phonological word processing. *International Journal of Psychophysiology, 32*(3), 215-231.

Kinnear, P. & Gray, C.(2000) *SPSS for Windows made simple: Release 10*. New York, NY: Psychology Press.

Kounios, J., Green, D. L., Payne, L., Fleck, J. I., Grondin, R., & McRae, K. (2009). Semantic richness and the activation of concepts in semantic memory: evidence from event-related potentials. *Brain Research, 1282*, 95-102.

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI, Brown University Press.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Kutas, M. and Hillyard, S. A. (1984). Brain potentials reflect word expectancy and semantic association during reading. *Nature, 307*, 161-163.

Lambon-Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition, 13*(7-8), 928-948.

Lambon-Ralph, M. A. , McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological abstract evidence and a computational model. *Journal of Cognitive Neuroscience, 13*(3), 341-356.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language, 102*(1), 30-45.

Le Voi, M. (2005). Connectionism. In: Braisby, N. (Ed.) *Cognitive psychology: A methods companion.* Oxford, UK: Oxford University Press, pp.25–68.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*, 122–142.

Lee,M.-E. & Williams, J. N. (2001). Lexical access in spoken word production by bilinguals: Evidence from the semantic competitor priming paradigm. *Bilingualism: Language and Cognition, 4*(3)*,* 233-248.

Luck, S. J. (2005a). *An introduction to the event-related potential technique.* Cambridge, MA: MIT.

Luck, S. J. (2005b). Ten simple rules for designing ERP experiments. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook (pp. 33-56).* Cambridge, MA: MIT Press.

Lupker, S. J. (1988). Picture naming: an investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning Memory and Cognition, 14*(3), 444-455.

Lupker, S.J., Brown, P., & Colombo, L. (1997) Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, *23 (3),* 570 – 590.

Magnum, G. R. & Hillyard, S. A. (1996). Mechanisms and models of selective attention. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind. Event-related brain potentials and cognition* (pp. 40-85). New York: Oxford University Press.

Mari-Beffa, P., Valdes, B., Cullen, D. J., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Cognitive Brain Research, 23*(2-3), 293-305.

Martin-Loeches, M., Hinojosa, J.A., Fernandez-Frias, C., Rubia, F.J., 2001.Functional differences in the semantic processing of concrete and abstract words. *Neuropsychologia, 39, 1086–1096.*

Masson, M. E. J. (1991). A distributed memory model of context effects in word identification. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 233-263). Hillsdale, NJ: Erlbaum.

Masson, M. E. J. (1995). A distributed-memory model of semantic priming. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(1), 3-23.

McKone, E. (1998). The decay of short-term implicit memory: Unpacking lag. *Memory and Cognition, 26*(6), 1173-1186.

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning Memory and Cognition, 18*(6), 1155-1172.

Mcnamara, T. P. (1992). Theories of Priming : I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1173-1190.

McNamara, T. P. (2005). **Essays in cognitive psychology. Semantic priming: Perspectives from memory and word recognition.** ***Hove, UK: Psychology Press.***

McRae, K., & Cree, G. (2002). Factors underlying category-specific deficits. In E. M. E. Forde & G. Humphreys (Eds.), *Category specificity in brain and mind.* Hove, U.K.: Psychology 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*(4), 547-559.

Medler, D.A., & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal Experimental Psychology, 90*(2), 227-234.

Mechelli, A., Friston, K. J., & Price, C. J. (2000). The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI. *Journal of Cognitive Neuroscience, 12,* 145–156.

Miller, G. A. (1990). WordNet: An on-line lexical database. *International Journal of Lexiography, 3, 235-312*.

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*(1), 65-79.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(1), 183-206.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*(3), 134-140.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18(3)*, 452 -467.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 50*(3), 528-559.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(1), 116-133.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibition-less spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226-254.

Neely, J. H. (1991) Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Besner, D., Humphreys, G. W. (Eds.), *Basic processes in reading: Visual word recognition*., (pp. 264-336). Hillsdale, NJ: Erlbaum.

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/

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods Instruments & Computers, 36*(3), 402-407.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*(2), 227-252.

Otten, L. J., & Rugg, M. D. (2005). Interpreting event-related brain potentials. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook (pp. 3-16).* Cambridge, MA: MIT Press.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53*(1), 13-24.

Patterson, K., & Hodges, J. R. (1992). ***Deterioration of word meaning: Implications for reading*****.** *Neuropsychologia*, *30*(12), 1025-1040.

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*(12), 976-987.

Patterson, K., Ralph, M. A. L., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., et al. (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience, 18*(2), 169-183.

Paivio, A. (1971). Imagery and deep structure in recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior, 10*(1), 1-12.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*(2), 273-315.

Perry, C., Ziegler, J. C., Braun, M., & Zorzi, M. (2010a). Rules versus statistics in reading aloud: New evidence on an old debate. *European Journal of Cognitive Psychology, 22*(5), 798-812.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010b). Beyond single syllables: Large-scale modelling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61*(2), 106-151.

Pexman, P. M., Lupker, S. J., & Hino, Y. (2002). The impact of feedback semantics in visual word recognition: Number-of-features effects in lexical decision and naming tasks. *Psychonomic Bulletin & Review, 9*(3), 542-549.

Plaut, D. C. (1991). *Connectionist neuropsychology: The breakdown and recovery of behavior in lesioned attractor networks.* Doctoral dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA. Available as Technical Report CMU-CS-91-185.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In **Proceedings of the 17th Annual Conference of the Cognitive Science Society**, pages 37-42. Hillsdale, NJ: Lawrence Erlbaum Associates.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Price, C. J., Moore, C. J., Humphreys, G. W., & Wise, R. J. S. (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience, 9*(6), 727-733.

Proverbio, A. M., Vecchi, L., & Zani, A. (2004). From orthography to phonetics: ERP measures of grapheme-to-phoneme conversion mechanisms in reading. *Journal of Cognitive Neuroscience, 16*(2), 301-317.

Pulvermuller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences, 5*(12), 517-524.

Pulvermuller, F., Cooper-Pye, E., Dine, C., Hauk, O., Nestor, P. J., & Patterson, K. (2010). The word processing deficit in semantic dementia: All categories are equal, but some categories are more equal than others. *Journal of Cognitive Neuroscience, 22*(9), 2027-2041.

Pylkkanen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences, 7*(5), 187-189.

Rastle, K., Croot, K. P., Harrington, J. M., & Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 1083-1095.

Rastle, K., & Davis, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 307-314.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95*(3), 385-408.

Riddoch, M. J., Humphreys, G. W., Coltheart, M., & Funnell, E. (1988). Semantic systems of system: Neuropsychological evidence re-examined. *Cognitive Neuropsychology, 5*(1), 3-25.

Rugg, M. D. & Allan, K. (2000). Event-related potential studies of memory. *The Oxford handbook of memory*, 521-537.

Rugg, M. D. & Coles, M. G. H. (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 1-23). New York: Oxford University Press.

Schvaneveldt, R. W., Meyer, D. E., & Becker, C. A. (1976). Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*(2), 243-256.

Seidenberg, M.S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T.G. Waller, & E.M. MacKinnon (eds.), *Reading research: Advances in theory and practice* (p. 199-252). New York: Academic Press.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Seidenberg, M. S., & Plaut, D. C. (2006). Progress in understanding word reading: Data fitting versus theory building. In S. Andrews. (Ed.), *From inkmarks to ideas: Current issues in lexical processing*. Hove, UK: Psychology Press.

Seifert, L.S. (1997). Activating representations in permanent memory: Different benefits for pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1106-1121.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition. *Journal of Verbal Learning and Verbal Behavior, 23*(3), 383-404.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport, 9*(10), 2195-2200.

Shibahara, N., Zorzi, M., Hill, M. P., Wydell, T., & Butterworth, B. (2003). Semantic effects in word naming: Evidence from English and Japanese Kanji. *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 56*(2), 263-286.

Slotnick, S. D. (2005). Source localisation of ERP generators. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook (pp. 149-166).* Cambridge, MA: MIT Press.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General, 109*(4), 373-392.

Solomon, R. L., & Howes, D. H. (1951). Word frequency, personal values, and visual duration thresholds. *Psychological Review, 58*(4), 256-270.

Spencer, K. A. (2009). Feedforward, -backward, and neutral transparency measures for British English. *Behavior Research Methods, 41*(1), 220-227.

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging, 15*(2), 225-231.

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38*(4), 598-605.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53*(4), 347-359.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(5), 1140-1154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(1), 207-214.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250.

Tabachnick, B.G. & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Pearson.

Thorndike, E. L., &. Lorge, I. (1944). The *teacher's word book of 30,000 words.* New York: Teachers College Press, Columbia University.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*.Hillsdale, NJ: Erlbaum.

Tree, J. J., & Hirsh, K. W. (2003). Sometimes faster, sometimes slower: Associative and competitor priming in picture naming with young and elderly participants. *Journal of Neurolinguistics, 16*(6), 489-514.

Tse, C. S., & Neely, J. H. (2007). Semantic priming from letter-searched primes occurs for low- but not high-frequency targets: Automatic semantic access may not be a myth. *Journal of Experimental Psychology: Learning Memory and Cognition, 33*(6), 1143-1161.

Tulving, E. (1972). Episodic and semantic memory. In E Tulving, & W Donaldson (Eds.) *Organization of Memory.* pp. 381–403. New York: Academic.

Tyler, L. K., Voice, J. K., & Moss, H. E. (2000). The interaction of meaning and sound in spoken word recognition. *Psychonomic Bulletin and Review, 7*(2), 320-326.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*(3), 289-335.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*(6597), 254-256.

Venezky,R. L. (1970). *The Structure of English Orthography.* The Hague: Mouton.

Vitkovitch, M., & Cooper, E. (2012). My word! Interference from reading object names implies a role for competition during picture name retrieval. *Quarterly Journal of Experimental Psychology, 65*(6), 1229-1240.

Vitkovitch, M., Cooper-Pye, E., & Leadbetter, A.G. (2006). Semantic priming over unrelated trials: Evidence for different effects in word and picture naming. *Memory and Cognition, 34*(3), 715–725.

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(4), 664-680.

Vitkovitch, M., Humphreys, G. W., & Lloydjones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning Memory and Cognition, 19*(2), 243-259.

Vitkovitch, M., & Rutter, C. (2000). The effects of response stimuli interval on error priming in sequential object naming. *Visual Cognition, 7*(5), 645-670.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*(4), 635-657.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 50*(2), 439-456.

Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology Section a: Human Experimental Psychology, 44*(4), 723-761.

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language, 33*(3), 332-356.

Wilson, M. D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioral Research Methods, Instruments and Computers*, 20, 6–11.

Woollams, A. M. (2005). Imageability and ambiguity effects in speeded naming: convergence and divergence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 878-890.

Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: a doubly typical signature of semantic dementia. *Neuropsychologia, 46*(10), 2503-2514.

Woollams, A. M., Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: on the association between semantic dementia and surface dyslexia. *Psychological Review, 114*(2), 316-339.

Wydell, T. N., Butterworth, B., & Patterson, K. (1995). The inconsistency of consistency effects in reading : The case of Japanese Kanji. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(5), 1155-1168.

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to Meaning for Kanji: Is a Rows still a Rose in Japanese Kanji. *Journal of Experimental Psychology: Learning Memory and Cognition, 19*(3), 491-514.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60*(4), 502-529.

Yates, M., Friend, J., & Ploetz, D. M. (2008). The effect of phonological neighborhood density on eye movements during reading. *Cognition, 107*, 685–692.

Young, A. W., & Ellis, A. W. (1985). Different methods of lexical access for words presented in the left and right visual hemifields. *Brain and Language, 24*(2), 326-358.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*(1), 1-29.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory and Cognition, 32*(1), 31-38.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods Instruments and Computers, 28*(4), 504-515.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for vertical bar u vertical bar? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods Instruments and Computers, 29*(4), 600-618.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1131-1161.

Appendices

Appendix A

Below the stimulus lists for Experiment 1 are presented in picture-prime-filler-filler-word-target quadruplets as a function of target word type and list. Primes were in picture modality and targets were in word modality as specified in column headings. Filler modality is specified within cell; fillers presented in picture modality are in uppercase letters, filler presented in word modality are in lower letters.

Low Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Mushroom	Pitchfork	camera	dice	shovel
Eel	Ring	skis	GLOBE	brooch
Jug	Scarf	WELL	blanket	hood
Scarf	Hairbrush	ARROW	TAIL	comb
Owl	Rope	chalk	smoke	thread
Rope	Tiger	diamond	BOX	leopard
tomato	Owl	BOTTLE	wheat	pigeon
Ring	Eel	UMBRELLA	CROW	salmon
Cannon	Tomato	socket	flag	radish
Hairbrush	Hamburger	torch	BIN	steak
Pitchfork	Vicar	ASHTRAY	city	monk
hamburger	Jug	KEY	BALL	vase
Vicar	Cannon	mirror	KITE	sword
tiger	Mushroom	BARBEQUE	magnet	onion

Low Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Harp	Spanner	coffin	jade	axe
Camel	Hat	badge	MOUNTAIN	glove
Submarine	Tie	CLOWN	rug	lapel
Seal	Submarine	PENCIL	RACKET	canoe
Squirrel	Harp	feather	bricks	cello
Needle	Squirrel	fairy	IGLOO	monkey
Spanner	Seal	ROAD	trophy	porpoise
Cake	Camel	BULB	WEB	sloth
paper	Cake	shell	bubble	dough
Ladle	Boot	acorn	WING	shoe
Hat	Apple	TWEEZERS	rock	pear
Tie	Ladle	CANDLE	PRAM	sieve
Apple	Paper	towel	HOOVER	scroll
Boot	Needle	FENCE	CLIFF	wool

Low Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Strawberry	Hoe	cloud	jar	rake
Glass	Necklace	coin	SLIDE	locket
Whale	Trousers	HANDCUFFS	cage	vest
Seahorse	Arm	PHONE	ROBOT	thigh
Gun	Fox	tulip	doll	hare
Necklace	Whale	ice	CURTAIN	shark
Chair	Peacock	ANCHOR	whip	stork
Trousers	Seahorse	PRESENT	TAP	crab
Hoe	Chair	wig	battery	settee
Arm	Strawberry	cork	BINOCULARS	peach
Sandwich	Cactus	FIRE	pool	willow
Peacock	Glass	PLASTER	TIN	mug
Cactus	Gun	balcony	RADIO	spear
Fox	Sandwich	TEETH	medal	Pie

Low Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Peas	Saw	pillow	beach	hammer
Fly	Belt	clamp	FAN	sock
Saw	Wheelbarrow	PARACHUTE	barrel	cart
House	Peas	BRIDGE	ANTLERS	bean
Carrot	Trumpet	package	soap	drum
Grapes	Broom	pillar	RAZOR	mop
Jacket	Fly	Iron	mask	ant
Rain	Sheep	FOUNTAIN	BALLOON	goat
Belt	Carrot	flask	hay	leek
sheep	Handbag	cigar	PUZZLE	purse
Broom	Grapes	MATCH	scale	cherry
Handbag	Rain	light switch	BATHTUB	sleet
Wheelbarrow	Jacket	hinge	LADDER	sleeve
Trumpet	House	GRAVE	picture	shack

Appendix B

Below is the participant information sheet completed by the experimenter that included the short post-experiment questionnaire for Experiment 1. A similar questionnaire was used in all priming experiments of this thesis. The changes to the questionnaire for subsequent experiments were to tailor questions to the number of stimuli for the specific Experiment.

Participant #: _________ Counter Balance Number: ______________

Age: ________ Gender: **M / F**

Do you wear glasses? ______ Are you wearing them today? ___________

What language did you learn first? What language did you learn as a baby?
____

Do you speak a second language? **Y / N**

If yes, what is your second language? __________________________

What age did you acquire this second language?________________

Do you speak these languages equally as well? ________________

I am going to ask you some questions about the experiment now. Please answer if you thought about these things <u>while</u> you were naming.

Did you notice if any of the items were related to each other?
____

Did you notice a pattern as to how the related items came up; did you notice it was in 4's? ____

Did you try to predict what specifically the target items?________________________

<u>List of Specific Naming Errors</u> **[(1)=correct; (2)= VK; (3)=incorrect]**

Appendix C

Below the stimulus lists for Experiment 2 are presented in picture-prime-word-filler-word-target triplets as a function of target word type and list. Stimuli modality is specified in column headings.

High Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Cup	Lion	Pipe	Bear
Salt	Lips	Iron	Eye
Skeleton	Mountain	Ruler	Field
Mountain	Skeleton	Wire	Heart
Lips	Salt	Bucket	Sugar
Policeman	Hair	Cigar	Blood
King	Tree	Rug	Bush
Hair	Cup	Lake	Stein
Lion	Policeman	Bin	Soldier
Tree	King	Scale	Chief

High Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Roof	Plate	Moth	Bowl
Neck	Moon	Towel	Earth
Window	thumb	Lamp	Foot
Moon	Neck	Bottle	Head
Pig	Leg	Torch	Shoulder
Tear	Crackers	Statue	Bread
Crackers	Tear	Robot	Sweat
Plate	Roof	Basket	Floor
Leg	Window	Shell	Door
thumb	Pig	Flower	Cow

High Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Skirt	Desk	whistle	Table
Cap	Plane	Mirror	Boat
Duck	Finger	Milk	Hand
Plane	cap	Spoon	shirt
Volcano	Dresser	Pin	Bed
Desk	Duck	Barrel	Chicken
Dresser	Bone	Plug	Skin
Finger	Ladder	Beach	stairs
Bone	Skirt	Paint	Coat
Ladder	Volcano	Book	Hill

High Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Rabbit	Car	Umbrella	Bus
Jumper	Nose	Hinge	Ear
Sink	Bird	Tent	Fish
Chest	Sink	Leaf	Bath
Car	pillow	Nut	Sheet
Horse	Jumper	Fairy	Dress
Nose	Cloud	Wine	Sun
Bird	Chest	Road	Face
Cloud	Horse	Scissors	Dog
Pillow	Rabbit	Anchor	Cat

Low Frequency Exception List1

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Ring	Pitchfork	Dice	Shovel
Tiger	Ring	Key	Brooch
Jug	Hairbrush	Box	Comb
Owl	Rope	Smoke	Thread
Rope	Tiger	Arrow	Leopard
Cannon	Owl	Wheat	Pigeon
Hairbrush	Hamburger	City	Steak
Pitchfork	Vicar	Ball	Monk
Hamburger	Jug	Kite	Vase
Vicar	Cannon	Globe	Sword

Low Frequency Exception List2

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Squirrel	Spanner	Jade	Axe
Camel	Hat	Fence	Glove
Paper	Tie	Clown	Lapel
Spanner	Squirrel	Bulb	Monkey
Apple	Camel	Web	Sloth
Ladle	Cake	Bubble	Dough
Cake	Boot	Rock	Shoe
Hat	Apple	Tweezers	Pear
Tie	Ladle	Pram	Sieve
Boot	Paper	Candle	Scroll

Low Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Strawberry	Trousers	Cage	Vest
Seahorse	Arm	Phone	Thigh
Glass	Whale	Doll	Shark
Chair	Peacock	Curtain	Stork
Trousers	Seahorse	Whip	Crab
Whale	Chair	Tin	Settee
Arm	Strawberry	Fire	Peach
Sandwich	Glass	Coin	Mug
Peacock	Gun	Wig	Spear
Gun	Sandwich	Medal	Pie

Low Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Target Words
Peas	Saw	Beach	Hammer
Fly	Belt	Clamp	Sock
House	Peas	Match	Bean
Jacket	Broom	Radio	Mop
Saw	Fly	Cork	Ant
Handbag	Sheep	Flask	Goat
Belt	Carrot	Soap	Leek
Sheep	Handbag	Razor	Purse
Broom	Jacket	Hay	Sleeve
Carrot	House	Balloon	Shack

Appendix D

Below are the sixteen counterbalance orders of stimulus lists used in priming Experiments 2, 3, and 4. Counterbalance (CB) are specified per row left to right with each participant of the experiments assigned to one of these counterbalances. Related and unrelated refer to the condition the primes appeared in for the particular stimulus list for that counterbalance order.

	High Frequency				**Low Frequency**			
	Exception		**Regular**		**Exception**		**Regular**	
	List 1	**List 2**	**List 3**	**List 4**	**List 1**	**List 2**	**List 3**	**List 4**
CB 1	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
CB 2	Related	Unrelated	Unrelated	Related	Related	Unrelated	Unrelated	Related
CB 3	Unrelated	Related	Related	Unrelated	Unrelated	Related	Related	Unrelated
CB 4	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
	Regular		**Exception**		**Regular**		**Exception**	
	List 3	**List 4**	**List 1**	**List 2**	**List 3**	**List 4**	**List 1**	**List 2**
CB 5	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
CB 6	Unrelated	Related	Related	Unrelated	Unrelated	Related	Related	Unrelated
CB 7	Related	Unrelated	Unrelated	Related	Related	Unrelated	Unrelated	Related
CB 8	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
	Low Frequency				**High Frequency**			
	Exception		**Regular**		**Exception**		**Regular**	
	List 1	**List 2**	**List 3**	**List 4**	**List 1**	**List 2**	**List 3**	**List 4**
CB 9	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
CB 10	Related	Unrelated	Unrelated	Related	Related	Unrelated	Unrelated	Related
CB 11	Unrelated	Related	Related	Unrelated	Unrelated	Related	Related	Unrelated
CB 12	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
	Regular		**Exception**		**Regular**		**Exception**	
	List 3	**List 4**	**List 1**	**List 2**	**List 3**	**List 4**	**List 1**	**List 2**
CB 13	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
CB 14	Unrelated	Related	Related	Unrelated	Unrelated	Related	Related	Unrelated
CB 15	Related	Unrelated	Unrelated	Related	Related	Unrelated	Unrelated	Related
CB 16	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related

Appendix E

Below the stimulus lists for Experiment 3 are presented in picture-prime-word-filler-word-filler-word-target quadruplets as a function of target word type and list. Stimuli modality is specified in column headings.

High Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Cup	Lion	Wagon	Ashtray	Bear
Salt	Lips	Bride	Hose	Eye
Skeleton	Mountain	Wasp	Beard	Field
Mountain	Skeleton	Crown	Iron	Heart
Lips	Salt	Violin	Chalk	Sugar
Policeman	Hair	Anchor	Swing	Blood
King	Tree	Net	Doll	Bush
Hair	Cup	Dice	Tweezers	Stein
Lion	Policeman	Tin	Knot	Soldier
Tree	King	Globe	Pin	Chief

High Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Roof	Plate	Diamond	Grave	Bowl
Neck	Moon	Flea	Witch	Earth
Window	Thumb	Nut	Mask	Foot
Moon	Neck	Clock	Steam	Head
Pig	leg	Horn	Fairy	Shoulder
tear	Crackers	Phone	Wing	Bread
crackers	Tear	Pyramid	Ball	Sweat
Plate	Roof	Arrow	Whip	Floor
Leg	Window	Torch	Rock	Door
thumb	Pig	Flask	Bag	Cow

High Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Skirt	Desk	Clown	Fence	Table
Cap	Plane	Doll	Match	Boat
Duck	Finger	Spider	Camera	Hand
Plane	Cap	Tulip	Radio	Shirt
Volcano	Dresser	Trumpet	Curtain	Bed
Desk	Duck	Barrel	Tent	Chicken
Dresser	Bone	Pillar	Wall	Skin
Finger	Ladder	Umbrella	Blanket	Stairs
Bone	Skirt	Hammock	Wheel	Coat
Ladder	Volcano	Pipe	Balloon	Hill

High Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Rabbit	Car	Flute	Plug	Bus
Jumper	Nose	Wine	Chimney	Ear
Sink	Bird	Corn	Ghost	Fish
Chest	Sink	Moth	Hay	Bath
Car	Pillow	Beetle	Racket	Sheet
horse	Jumper	Box	Cork	Dress
Nose	Cloud	Plaster	Ribbon	Sun
Bird	Chest	Knot	Pencil	Face
Cloud	Horse	Scale	Basket	Dog
pillow	Rabbit	Well	Statue	Cat

Low Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Ring	Pitchfork	Drum	Coin	Shovel
Tiger	Ring	Flower	Pool	Brooch
Jug	Hairbrush	Train	Scissors	Comb
Owl	Rope	Nail	Button	Thread
Rope	Tiger	Spoon	Cigar	Leopard
Cannon	Owl	Ice	Battery	Pigeon
Hairbrush	Hamburger	rain	Kite	Steak
Pitchfork	Vicar	City	Skis	Monk
Hamburger	Jug	Tap	Soap	Vase
Vicar	Cannon	flag	River	Sword

Low Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Squirrel	Spanner	Bell	Pram	Axe
Camel	Hat	Milk	Smoke	Glove
paper	Tie	Hinge	Beach	Lapel
spanner	Squirrel	Cane	Bottle	Monkey
apple	Camel	Rose	Lake	Sloth
ladle	Cake	Wire	Tail	Dough
cake	Boot	Road	Mirror	Shoe
Hat	Apple	Key	Badge	Pear
Tie	Ladle	Bubble	Chain	Sieve
Boot	Paper	Razor	Cliff	Scroll

Low Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Strawberry	Trousers	Bicycle	Robot	Vest
Seahorse	Arm	Guitar	Bin	Thigh
Glass	Whale	Cactus	Towel	Shark
Chair	Peacock	Wheat	Bridge	Stork
Trousers	Seahorse	Lamp	Paint	Crab
Whale	Chair	Magnet	Fan	Settee
Arm	Strawberry	Coffin	Thimble	Peach
Sandwich	Glass	Web	Funnel	Mug
Peacock	Gun	Fountain	Antlers	Spear
Gun	Sandwich	hook	candle	Pie

Low Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Filler Words	Filler Words	Target Words
Peas	Saw	Jade	Bulb	Hammer
Fly	Belt	Lemon	Trophy	Sock
House	Peas	Rug	Shell	bean
jacket	Broom	Paino	Slide	mop
saw	Fly	Bomb	Needle	Ant
Handbag	Sheep	Medal	Leaf	Goat
Belt	Carrot	Saddle	Feather	Leek
Sheep	Handbag	Wig	Cage	Purse
broom	Jacket	Fire	Envelope	Sleeve
carrot	House	puzzle	bucket	Shack

Appendix F

Below the stimulus lists for Experiment 4 are presented in picture-prime-filler-filler-word-target quadruplets as a function of target word type and list. Primes were in picture modality and targets were in word modality as specified in column headings. Filler modality is specified within cell; fillers presented in picture modality are in uppercase letters, filler presented in word modality are in lower letters.

High Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Cup	Lion	wagon	ASHTRAY	Bear
Salt	Lips	BRIDE	HOSE	Eye
Skeleton	Mountain	BEE	BEARD	Field
Mountain	Skeleton	CROWN	iron	Heart
Lips	Salt	VIOLIN	chalk	Sugar
Policeman	Hair	anchor	swing	Blood
King	Tree	NET	CONE	Bush
Hair	Cup	dice	TWEEZERS	Stein
Lion	Policeman	tin	KNOT	Soldier
Tree	King	GLOBE	pin	Chief

High Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Roof	Plate	diamond	GRAVE	Bowl
Neck	Moon	flea	WITCH	Earth
Window	Thumb	NUT	mask	Foot
Moon	Neck	CLOCK	steam	Head
Pig	leg	horn	fairy	Shoulder
tear	Crackers	phone	wing	Bread
crackers	Tear	PYRAMID	BALL	Sweat
Plate	Roof	ARROW	WHIP	Floor
Leg	Window	torch	ROCK	Door
thumb	Pig	FLASK	bag	Cow

High Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Skirt	Desk	CLOWN	FENCE	Table
Cap	Plane	DOLL	MATCH	Boat
Duck	Finger	spider	CAMERA	Hand
Plane	Cap	tulip	RADIO	Shirt
Volcano	Dresser	trumpet	curtain	Bed
Desk	Duck	barrel	tent	Chicken
Dresser	Bone	PILLAR	wall	Skin
Finger	Ladder	UMBRELLA	blanket	Stairs
Bone	Skirt	hammock	WHEEL	Coat
Ladder	Volcano	PIPE	balloon	Hill

High Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Rabbit	Car	FLUTE	PLUG	Bus
Jumper	Nose	WINE	CHIMNEY	Ear
Sink	Bird	corn	ghost	Fish
Chest	Sink	moth	HAY	Bath
Car	Pillow	beetle	RACKET	Sheet
horse	Jumper	BOX	cork	Dress
Nose	Cloud	PLASTER	ribbon	Sun
Bird	Chest	scale	pencil	Face
Cloud	Horse	tassle	BASKET	Dog
pillow	Rabbit	WELL	statue	Cat

Low Frequency Exception List 1

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Ring	Pitchfork	DRUM	COIN	Shovel
Tiger	Ring	FLOWER	POOL	Brooch
Jug	Hairbrush	train	SCISSORS	Comb
Owl	Rope	nail	BUTTON	Thread
Rope	Tiger	SPOON	cigar	Leopard
Cannon	Owl	ice	battery	Pigeon
Hairbrush	Hamburger	rain	kite	Steak
Pitchfork	Vicar	city	SKIS	Monk
Hamburger	Jug	TAP	soap	Vase
Vicar	Cannon	FLAG	river	Sword

Low Frequency Exception List 2

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Squirrel	Spanner	bell	pram	Axe
Camel	Hat	milk	smoke	Glove
paper	Tie	HINGE	beach	Lapel
spanner	Squirrel	CANE	BOTTLE	Monkey
apple	Camel	ROSE	lake	Sloth
ladle	Cake	wire	TAIL	Dough
cake	Boot	road	MIRROR	Shoe
Hat	Apple	KEY	BADGE	Pear
Tie	Ladle	bubble	CHAIN	Sieve
Boot	Paper	HOOK	cliff	Scroll

Low Frequency Regular List 3

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Strawberry	Trousers	BICYCLE	ROBOT	Vest
Seahorse	Arm	GUITAR	bin	Thigh
Glass	Whale	CACTUS	TOWEL	Shark
Chair	Peacock	wheat	BRIDGE	Stork
Trousers	Seahorse	lamp	paint	Crab
Whale	Chair	MAGNET	FAN	Settee
Arm	Strawberry	coffin	THIMBLE	Peach
Sandwich	Glass	web	FUNNEL	Mug
Peacock	Gun	fountain	ANTLERS	Spear
Gun	Sandwich	CANDLE	razor	Pie

Low Frequency Regular List 4

Unrelated Prime Pictures	Related Prime Pictures	Fillers	Fillers	Target Words
Peas	Saw	RUG	BULB	Hammer
Fly	Belt	lemon	TROPHY	Sock
House	Peas	jade	shell	bean
jacket	Broom	PIANO	slide	mop
saw	Fly	BOMB	NEEDLE	Ant
Handbag	Sheep	MEDLA	LEAF	Goat
Belt	Carrot	saddle	feather	Leek
Sheep	Handbag	wig	cage	Purse
broom	Jacket	FIRE	envelope	Sleeve
carrot	House	puzzle	PAIL	Shack

Appendix G

Below are the stimulus lists for Experiment 5.

Low Frequency			
Exception		Regular	
Low Imageability	High Imageability	Low Imageability	High Imageability
ache	axe	blessing	ant
cache	boulder	blunder	bean
caste	brooch	blunt	blister
chasm	comb	deed	corpse
chute	crow	dell	crab
dread	dough	dumb	duck
fatigue	dove	figment	goat
Feat	glove	fraud	hammer
Foul	lapel	gait	lamb
lease	leopard	graft	leek
lunge	limb	hoarse	mop
mischief	monk	keel	mug
pint	monkey	madness	peach
plead	pear	mince	pickle
pork	pigeon	mister	pie
scarce	scroll	peep	pup
sew	shoe	rude	purse
shone	shovel	sage	settee
soot	sieve	salve	shack
stingy	sloth	sane	shark
swarm	steak	scorn	sleeve
ton	swamp	slang	sock
tow	sword	traitor	spear
wand	thread	tuck	stork
warn	tomb	vain	thigh
womb	vase	wail	trumpet
wrath	wolf	whoop	vest

Appendix H

Below are the stimulus lists for Experiment 6.

Low in Number of Semantic Features	High in Number of Semantic Features
bagpipe	apple
beehive	balloon
bin	bra
bouquet	canoe
brick	cheese
bucket	chicken
buckle	couch
clam	cougar
cork	deer
doorknob	dolphin
eggplant	drapes
garage	fawn
harp	freezer
hook	fridge
mackerel	garlic
menu	goat
mixer	grape
moth	grapefruit
otter	hare
paintbrush	nylons
pheasant	penguin
sack	plate
saddle	prune
shell	rooster
spade	screws
tongs	sofa
toy	toad
tripod	turkey
trolley	turtle
turnip	vulture