Children's online processing of written language Inferences from eye movements

Joseph, Holly

Children's online processing of written language: Inferences from eye movements

Holly Joseph

Thesis submitted for the Degree of Doctor of Philosophy

Durham University,

Department of Psychology

2009

Holly Joseph

Children's online processing of written language: Inferences from eye movements

PhD, 2009

Abstract

This thesis reports seven experiments that investigated children's online processing of written language. A variety of visual and linguistic factors, for which there are well-documented effects in adults, were manipulated in order to examine children's and adults' eye movements as they read sentences containing these manipulations. Experiment 1 investigated saccadic targeting of long and short words and showed that adults and children are generally alike in where they target their initial saccade to a word, and how they use parafoveal word length information to skip words. It also showed that while the length of a word directly influences the eye movement behaviour of both adults and children during text reading, the magnitude of these effects is greater in children. Experiments 2 and 3 showed that reliable word frequency effects are observed in adults and children when age-appropriate texts are used to index frequency counts. Experiment 3 also showed that word frequency effects occur in children, even when the age at which words were acquired is held constant. In Experiment 4 lexical processing of semantically ambiguous words was examined and the data were suggestive of there being a cost associated with processing words with more than one meaning for older children. However, the effects were not robust in adults or younger children. Experiments 5 and 6 investigated syntactic parsing and showed that children are slightly delayed relative to adults in their detection of initial syntactic misanalysis, but that they appear to have a similar sentence-parsing mechanism in place as adults. Finally, Experiment 7 investigated thematic processing of anomalous and implausible sentences, and showed that while there is no difference in the time course of thematic anomaly detection in adults and children; children are delayed in their detection of thematic implausibility as compared to adults. Overall, the data show that adults and children appear to have similar mechanisms in place for processing written language visually, lexically, syntactically and thematically. They also show that the magnitude of disruption associated with these effects is greater in children than in adults, that the time course of children's syntactic processing is slightly delayed relative to that of adults, and further, that children are delayed in the efficiency with which they are able to integrate pragmatic and real world knowledge into the discourse representation. The thesis also makes a number of methodological points that have implications for conducting future research with developmental populations.

Declaration

I hereby declare that this thesis has been composed by myself and that the research reported herein has been conducted by myself.

January 2009

Holly Joseph

Experiment 1 has been invited for resubmission in the following manuscript:

Joseph, H.S.S.L, Liversedge, S.P., Blythe, H.I., White & Rayner, K. Word length and landing position effects during reading in children and adults. *Vision Research*.

Experiments 2 and 3 are in preparation for submission to the following manuscript:

Joseph, H.S.S.L, Blythe, H.I. & Liversedge, S.P. Differential effects of word frequency in adults and children during reading. *Psychonomic Bulletin and Review*.

Experiment 7 is reported in:

Joseph, H.S.S.L, Liversedge, S.P., Blythe, H.I., White, S.J., Gathercole, S.E. & Rayner, K. (2008). Children's and adults processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, 61 (5), 708-723.

Material from Chapters 1 and 2 has been included in an invited review article to be published in:

Blythe, H.I. & Joseph, H.S.S.L. (2009) Children's eye movements during reading. In S. P. Liversedge & I. D. Gilchrist & S. Everling (Eds.) *The Oxford Handbook of Eye Movements*. Oxford: Oxford University Press.

Statement of Copyright

Acknowledgements

First, I would like to thank my supervisor, Simon Liversedge, who helped to make the research process stimulating, challenging and very enjoyable. Many thanks also to John Findlay for his support and advice. Thank you to Keith Rayner and Sue Gathercole for their expert advice on funding applications and submitted manuscripts.

Thank you also to Sarah White for her tireless assistance with software, hardware and landing position analyses, and to Elaine Stanton and Bob Metcalf for their programming expertise. I would like to thank colleagues in the psychology department, for their advice and assistance, and in particular, my fellow PhD students, especially Sue Fletcher-Watson and Hazel Blythe, for making our working environment an interesting and pleasant place to be.

A big thank you to all the parents and children who took the time to come to the department to take part in my experiments, many on several occasions. And thank you to the head teachers and teachers who supported my research and allowed me to go into their schools to test their pupils despite the disruption. In particular, I would like to thank head teachers Mr. Sharland, Mrs. Whitton, and Mr. Pritchard who were so enthusiastic and supportive of my research project. It goes without saying that without your help, my PhD would have come to nothing! Thanks also to all the adults who took part in my experiments.

Finally, a huge thank you to my family, Keith, Zinzi and Milo, for bearing with me through it all, and for providing a frequent and happy distraction from work.

Table of Contents

List of Tables

Table of Illustrations

Chapter 1 : Language processing

Reading is the extraction of meaning from visual linguistic stimuli, and it is a process which must be learned and developed. While much is known about adults' online reading behaviour, and as a consequence, the psychological mechanisms that adult readers have in place with which they can extract meaning from written text, as yet relatively little is known about how children carry out this process. This thesis examines children's online processing of written language using eye movement recording methodology. More specifically, it uses linguistic manipulations which have produced well-documented, robust effects in adult readers to investigate whether children have the same psychological mechanisms in place for processing various linguistic phenomena.

Eye movements are important in reading because they mediate the sequence of cognitive processes that are required for the visual and linguistic processing of text. Moreover, the length of time that the eyes remain fixated on a word reflects the ease or difficulty with which that word is being processed (Just & Carpenter, 1980), and this close correspondence between eye gaze and cognitive processing provides a basis for much eye movement research. There is a large literature that has used eye movement methodology to investigate adults' online language processing over the last 30 years, and this has provided a starting point for conducting research with children: we already know of certain robust processing preferences that adult readers have, and we can use the linguistic manipulations that have revealed these preferences as a diagnostic tool to examine whether children have these same preferences in relation to a variety of linguistic phenomena.

This first introductory chapter of this thesis sets out to examine in some detail what is involved in language comprehension, in particular written language processing (i.e. reading) in adults. It is important to understand the mature language processing system in order to begin to investigate how language processing develops in children, as adult performance can be considered to be a benchmark from which to compare children's reading behaviour. Section 1.1 will provide a general overview of what is involved in language processing and describe the sub-processes which necessarily occur for language comprehension to be successful. Section 1.2 will describe the visual processing of text. Section 1.3 will describe lexical processing and outline some of the factors which can affect the ease or difficulty associated with identifying a word. Section 1.4 will review some key research investigating syntactic processing. Section 1.5 will describe semantic processing and focus on a key aspect of semantic processing relevant to the thesis: thematic processing. Section 1.6 will introduce the field of eye movements during reading, outline some of the key issues in this area of research, and examine the relationship that exists between oculomotor behaviour and language comprehension. Finally, Section 1.7 will draw conclusions and summarise what is known about written language processing and how this relates to the thesis.

1.1 How do we process (written) language?

Written language is a systematic arrangement of abstract symbols that conveys meaning and therefore provides a basis for communication. These symbols (letters) are combined to make words which represent units of information. Words are combined according to a set of conventions known as a grammar, which stipulates which combinations of words are legal within a language. Bound by these conventions, words can be combined productively and creatively to convey new information. Language processing has traditionally been separated into different components (or sub-processes), which can be investigated individually before trying to determine how they interact with and depend on one another (Garrod & Pickering, 1999). These subcomponents are: lexical processing (the process of identifying a word); syntactic processing (the process of building structure in a sentence); and semantic processing (the processing of sentential meaning). The goal of this thesis to is understand what these processors must do in order that children, as developing readers, are able to form a semantic representation of the text they are reading.

Although written language comprehension has much in common with spoken language comprehension, there are also some aspects of written language which make it unique. When we read, we use visual rather than acoustic information to extract meaning and make sense of what we have read, and there are consequences of this which are very useful to experimenters in the field of written language processing. First, while spoken words are distributed through time from beginning to end, and are therefore not heard all at once, written words are distributed over space and are generally (if they are not very long) available to readers as a whole. Second, there is no physical record that can be directly consulted in spoken language, but in printed text there is usually an opportunity to re-read words if required. Finally, there is no reliable cue in speech to mark word boundaries; while in printed text, spaces between words unambiguously mark the beginnings and endings of words. These properties of written language make it ideally suited to controlled experimental research as words (or groups of words) can be easily isolated and manipulated. The processing of written as opposed to spoken language will be the focus of this chapter, as it is of the thesis as a whole.

A reader comprehends a sentence by interpreting it incrementally; that is to say, she develops a representation of the meaning of the sentence roughly on a word-by-word basis as each word of the sentence is processed (Crocker, 1999). In order to do this, a series of extremely efficient, highly automated processes must exist, and there are temporal dependencies between these processes such that certain of them must occur before others may take place (Frazier & Rayner, 1982). When the reader is faced with a written text, she usually moves her eyes in order to fixate the top left-hand corner of a portion of the text, and will usually move her eyes in a rightwards direction as she reads[1]. Her eyes do not move smoothly across the page but rather make a series of fixations during which the eye remains stable. Between fixations are saccades during which the eye jumps a few letters forward (and sometimes backwards) in order to bring the next portion of text into view. At fixation onset, visual information is encoded automatically. Individual letters are detected through an analysis of their visual features which are then bound together into a unitary orthographic representation. This process of identifying letters is known as orthographic encoding, and a reader is able to identify letters of a word which is not yet

[1] Of course this is not true of all languages: for example in Hebrew, readers will initially fixate the right hand side of the page and move their eyes in a leftwards direction. Note that all research reported in this thesis refers to reading in the English language unless specified otherwise

fixated, as well as the letters of the word which is currently fixated, making the reading process more efficient. Following orthographic encoding of a word, lexical identification processes can begin.

As each word of the sentence is read, it is necessary to lexically access (i.e. locate the representation of the word within the mental lexicon) and identify the word. Upon lexical identification of the word, information about that word, including its syntactic category and its meaning, become available. Once the word's syntactic category is available it is then possible to carry out syntactic processing (or parsing), whereby the reader computes the structural relationships that exist between the words of the sentence. This allows the reader, roughly speaking, to work out who or what did what to whom. Finally, on the basis of the individual word meanings and the structural relations that exist between the words, the meaning of the sentence as a whole may be computed. The semantic representation of the sentence is also constructed incrementally, being elaborated roughly as each new word of the sentence is read (Pickering & Traxler, 1998; Traxler, Bybee, & Pickering, 1997). Following this brief overview of the processes involved in the comprehension of a written sentence, a detailed explanation of what is involved in each stage of processing (visual, lexical, syntactic and semantic) will now be given.

1.2 Visual processing

Before we begin linguistic processing of a word, we must first process it visually. When we directly fixate a word, visual features of the word are projected from the central region of the retina known as the fovea to the visual cortex. The fovea subtends about two degrees of visual angle around the fixation point and visual acuity is greatest at this point. It is for this reason that it is vital to move the eyes frequently during reading: discriminating the fine details of letters and words necessary for efficient reading is only possible at the centre of vision, and the further the word centre is from the fovea, the longer the time needed to encode it (Reichle, Rayner, & Pollatsek, 2004).

The light falling on the retina stimulates the receptors (rods and cones) in the retina in order to convert the light stimulus to an electrical signal which travels through the optic nerves to the optic tract, which projects onto the lateral geniculate nucleus (LGN). After the neurons of the LGN are stimulated they send axonal projections to the primary visual cortex in the occipital lobe, where the visual information is processed. This process takes place remarkably quickly in adult readers: research has shown that if text is masked (Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981) or disappears (Liversedge et al., 2004; Rayner, Liversedge, White, & Vergilino-Perez, 2003) after only 50 or 60ms, reading behaviour and comprehension proceed as normal, indicating that the visual information necessary for linguistic processing to commence is extracted very rapidly at the beginning of a fixation.

Orthography is clearly important in visual word recognition as written words are, of course, made up of letters and (at least some of) those individual letters must be processed in order that processing of the word as a whole can begin. Letters are detected through the analysis of their visual features (e.g. horizontal lines, edges, and corners) and are coordinated to generate unitary abstract letter codes (Clifton, Staub, & Rayner, 2007; Rayner, McConkie, & Zola, 1980; Rayner & Pollatsek, 1989). These features must necessarily be abstract in order that we are able to recognise the same word in lower case, UPPER CASE, mIxEd CaSe, as well as in different typefaces and handwriting. Furthermore, letters are thought to be processed in parallel (rather than serially) during word identification (Paap, Newsome, McDonald, & Schvaneveldt, 1982; Rayner & Pollatsek, 1989; Reicher, 1969), and letters are identified more accurately within words than as single letters (Reicher, 1969): a phenomenon known as the word-superiority effect.

Letter order is important and beginning letters are especially important for lexical identification (Lima & Pollatsek, 1983; White, Johnson, Liversedge, & Rayner, 2008), perhaps because word-initial letters constrain the number of lexical candidates during lexical access. Transposing letters (e.g. *found* becomes *fuond*) within a word increases processing time, although letter transpositions have much less of an effect than letter substitutions within words. External transpositions and substitutions (i.e. the transposition or substitution occurs on the first or last letters) disrupt processing more than internal transpositions and substitutions (Johnson, Perea, & Rayner, 2007;

Perea & Lupker, 2003; Rayner & Kaiser, 1975; Rayner, White, Johnson, & Liversedge, 2006; White et al., 2008) and transpositions affect low-frequency more than high-frequency words (White et al., 2008). These results show that specific letter identities, as well as letter positioning, are crucial to successful word identification, although importantly, words are generally successfully recognised despite transpositions (depending on the nature of the transposition). Several models of visual word recognition have proposed explanations of these results, in terms of position-specific coding (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; McClelland & Rumelhart, 1981; Paap et al., 1982), contextual coding (Rumelhart & McClelland, 1986), open-bigram coding (Whitney, 2001), and spatial coding (Davis, 1999). For a full review of these models, see Davis and Bowers (2006), and White et al. (2008).

At a sentence level, early visual processing enables the reader to obtain word-boundary information that is needed to programme subsequent saccades, as well as being the first stage of word identification. The visual characteristics of a word therefore have an impact on the ease with which a reader is able to process that word. For example, it is well-documented that word length reliably influences eye movement behaviour during reading in adults. Specifically, adult readers are more likely to fixate (Rayner & McConkie, 1976) and refixate (Vitu, O'Regan, & Mittau, 1990) a long as compared to a short word, presumably because a long word extends further across the visual field. Readers also fixate a long word for relatively longer than a short word (Just & Carpenter, 1980; Rayner, Sereno, & Raney, 1996). Word length effects in adults and children will be examined in Experiment 1 (Chapter 3).

Furthermore, *where* a reader fixates within a word during reading also has a direct impact on the ease with which it is processed. Adult readers tend to fixate the same location in a word – a little to the left of the word centre known as the *Preferred Viewing Location* (PVL: McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979), and the duration of their fixations, as well as the probability of making an intra-word refixation, is modulated by their initial fixation location (O'Regan & Lévy-Schoen, 1987; Vitu, McConkie, Kerr, & O'Regan, 2001). These important findings will be directly investigated and discussed in detail in Chapter 3. Finally, readers are able to extract word length information parafoveally, from up to 15 characters to the right of fixation, and four characters to the left (McConkie & Rayner, 1975, 1976; Rayner, 1986), an area known as the perceptual span (see Chapter 2; Section 1.6 for a more detailed account of the perceptual span).

1.3 Lexical processing

Lexical processing refers to the process by which individual words are identified. Words are the smallest independent meaningful elements in language, and many researchers in the field of language processing agree with Balota (1994, p. 303) that the word "is as central to psycholinguists as the cell is to biologists." This has meant that the word as a unit has received much attention in the field of (written) language processing, although it should be noted that it is not the only possible unit in terms of which the lexicon could be organised[2]. However, because the word has received so much attention in the literature, and because it is the easiest unit to manipulate and control experimentally, the focus of this review will also be the word. During reading, before it is possible to compute structural relationships between words; or to interpret the overall meaning of a sentence, individual words must first be identified. First, we must access the representation that corresponds to the word which is stored in memory (lexical access), and this is followed by the process of lexical identification whereby the information that relates to the word (e.g. its meaning and syntactic category) is made available.

In order to access the representation of a word, it is necessary to locate the representation that corresponds to the perceived word within our mental lexicon, and this process occurs quickly and with remarkably few errors (Rayner & Pollatsek, 1989). Whether this process is done by searching through items in the lexicon serially (e.g. Forster, 1976) or by accessing an item directly (often activating possible candidates in parallel until one candidate reaches an activation threshold, e.g. McClelland & Rumelhart, 1981; Seidenberg & McClelland, 1989) is a matter of

[2] Alternative sub-lexical units are morphemes, syllables, sub-syllables, or even phonemes, the last three being particularly suited to the study of spoken language but which can also be applied to the study of written language. While there is some evidence that the lexicon is organised morphemically rather than in word units (e.g. Marslen-Wilson, Tyler, Waksler, & Older, 2002), most researchers in the field of reading use the word as the basic unit of language.

debate and has resulted in the development of different models of lexical access. For the purposes of this introductory chapter, it is not necessary to provide an exhaustive description of the numerous models which have been put forward. However, it is important to bear in mind that we must have a very efficient system in place which allows us to recognise and comprehend words at a rate of three or four per second (Rayner & Pollatsek, 1989).

Words can be broken down into graphemes (letters), phonemes (sounds) and morphemes (units of meaning) and although in this thesis the emphasis is on the word as a basic unit of language, it may be that the recognition of phonemes, graphemes and morphemes must necessarily precede word recognition. Examining how readers process these sub-units can yield useful information regarding word recognition processes generally, and help us to understand how the lexical processor is structured. Each of these lexical sub-units will now be considered in turn.

1.3.1 Orthographic processing

Visual processing of letters and words has already been discussed in the section above. This section describes the influence of orthography on the process of lexical identification. Orthographic neighbourhood size, defined as the number of words which can be constructed by changing just one letter of the target word (Coltheart, Davelaar, Jonasson, & Besner, 1977) so that the word *marsh* has two neighbours, *harsh* and *march* (Forster & Shen, 1996), plays an important role in word recognition. During lexical access, we must select the single correct lexical item from a pool of possible candidates (a candidate set), eventually discriminating the correct lexical item from its orthographic neighbours. Many models of lexical access (Davis, 1999; Forster, 1976; Grainger & Jacobs, 1996; McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) propose that a visual word activates not only its own memory representation but also those of words which are orthographically similar. The presence, as well as the frequency, of neighbours has been shown to have both facilitative and inhibitory effects on lexical access, depending on the frequency of the target word as well as the task employed (Andrews, 1989, 1992; Balota, Paul, & Spieler, 1999; Perea & Pollatsek, 1998), and indeed prior processing of a word's neighbour has a subsequent inhibitory effect on lexical identification

(Paterson, Liversedge, & Davis, 2008), while a parafoveal preview of a word's neighbour facilitates lexical identification (Williams, Perea, Pollatsek, & Rayner, 2006).

It is clear that orthographic encoding is critical to the lexical identification process, and the orthographic composition of a word can have a significant influence on the ease and speed with which a word is identified. Section 1.2 above discussed research showing that the first letter of a word is particularly important, and that the specific letters which make up a word are more important than the order in which they appear. Neighbourhood effects are complex and whether a word is high or low frequency will have an impact on whether having many neighbours is a help or a hindrance. In general terms, words which are orthographically similar to a word which is to be accessed modulate this process showing that the orthographic properties of a word are used to access items in the lexicon during reading.

1.3.2 Phonological processing

Although it is almost certain that we generate phonological as well as orthographic representations of printed words when we read, there has been some debate as to what extent these phonological representations are used in initial stages of lexical processing to derive word meaning, or whether they become available only after word meaning has been activated along with syntactic and semantic information. Importantly, whether phonological codes are used to access lexical items addresses the wider issue of whether there is one direct (visual) route to lexical access during reading or whether an indirect (phonological) route is also used. Although most researchers now agree that there are two possible routes (Coltheart, Curtis, Atkins, & Haller, 1993), findings have been somewhat inconsistent as to what extent the phonological route is used during normal reading in adults (Jared, Levy, & Rayner, 1999).

Two eye-movement studies by Daneman and colleagues (Daneman & Reingold, 1993, 1995; Daneman, Reingold, & Davidson, 1995) showed that participants did not exhibit longer reading times on homophones (e.g. *board* and *bored*) as compared to non-homophonic spelling control words (e.g. *beard*), showing that adult readers initially use orthographic rather than phonological representations to access lexical items during normal reading, and that phonological codes do not play an early important role in word identification, but rather are activated at a post-lexical stage. Other research suggests that phonology does play an important role in the activation of word meanings, especially in the cases of poor readers, in processing low frequency words (Jared et al., 1999), in processing phonologically ambiguous words such as *tear* or *wind* (Carpenter & Just, 1981), in integrating information across saccades (Pollatsek, Lesch, Morris, & Rayner, 1992) and in highly-predictable contexts (Daneman & Stainton, 1991; Rayner, Pollatsek, & Binder, 1998). Overall, it appears that phonological codes may be accessed early in the encoding of a word (K Rayner et al., 1998), but it may be that phonology is preferentially used (as compared to using the orthographic route) to activate meanings only in particular circumstances, such as when identifying low frequency words.

1.3.3 Morphological processing

A larger sub-unit of the word (compared to phonemes and graphemes) is the morpheme. Morphemes are the smallest unit of meaning in language and while some words contain only one morpheme (such as *happy*), many longer words contain two or more (e.g. *undecided* contains three: *un - decide - ed*). There has been a fair amount of research investigating the role of morphology in lexical access; in particular, whether morphemes in multi-morphemic words (especially compound words such as *cowboy* or *blackboard*) are accessed individually, or whether there is an entry in the lexicon for the whole word. In addition, there is a question as to whether morphemes are stored in memory with possible affixes and suffixes which could be attached to them. Hyönä and Pollatsek (1998) monitored the eye movements of participants as they read sentences containing compound words made up of two morphemes (in Finnish), the first of which was manipulated for frequency. They found that the frequency of the first morpheme influenced initial fixation durations (participants looked longer at low, as compared to high, frequency morphemes) but that only later fixations were influenced by the frequency of the second morpheme (see also Andrews, Miller, & Rayner, 2004; Juhasz, Starr, Inhoff, & Placke, 2003; Pollatsek, Hyönä, & Bertram, 2000). This result shows that readers decompose words into the constituent morphemes and suggests that the activation of the beginning morpheme in a compound word precedes the activation of the end morpheme (and the word as a whole). More recent research has shown that morphemic constituents are activated en-route to the retrieval of whole compound words, and that readers do represent compound words at a whole-word level in their mental lexicon (Juhasz, 2008).

1.3.4 Factors which influence lexical processing

There are several lexical characteristics which are known to influence the ease or difficulty with which a word is recognised. One of the most robust findings in the eye-movement literature (as well as in other areas of language processing) is that word frequency (i.e. how often a word is encountered in natural language, indexed by corpus data) has a direct impact on how long it takes a reader to process a word (e.g. Henderson & Ferreira, 1990; Inhoff, 1984; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner, 1977; Rayner & Duffy, 1986; Rayner, Liversedge et al., 2003; Rayner & Raney, 1996) and the wealth of data supporting the role of word frequency in lexical access suggests that it is a fundamental characteristic in the organisation of the mental lexicon. The role of word frequency has been incorporated in models of lexical access. In search models of lexical access, high-frequency words are searched prior to low frequency words; (Forster, 1976). In interactive-activation models, high and low frequency words have differing activation thresholds meaning that high frequency words will reach threshold more quickly than low frequency words (McClelland & Rumelhart, 1981). Finally, in activation-verification models, the order of verification is determined, in part, by word frequency (e.g. Paap et al., 1982). Note that, as discussed above, morphemic frequency has also been shown to influence how long readers fixate a word during reading (Hyönä & Pollatsek, 1998; Juhasz, 2008). Experiments 1, 2, and 3 investigate word frequency effects in adults and children.

Word recognition processes have also been shown to be influenced by Age-of-Acquisition (i.e. the age at which the word was learnt: Juhasz & Rayner, 2006; Juhasz, 2005; Juhasz & Rayner, 2003), concreteness (how concrete or abstract a word is: Juhasz & Rayner, 2003), familiarity (measure of the frequency of exposure to a word, highly correlated with word frequency: Juhasz & Rayner, 2003; Williams & Morris, 2004), and lexical ambiguity (whether a word has more than one possible meaning such as bark: Rayner & Duffy, 1986). Finally, words which are predictable from the preceding context are fixated for less time than words which are not

(Balota, Pollatsek, & Rayner, 1985; Ehrlich & Rayner, 1981) and are also skipped more often (Rayner & Well, 1996), suggesting that context can speed lexical access (see Rayner, 1998 for a full review).

Lexical ambiguity effects, which will be discussed in detail in Chapter 5, have been particularly informative with regard to our knowledge of lexical processing. Lexical ambiguity research has addressed how a reader determines the contextually appropriate meaning of a word which has more than one meaning. Using a cross-modal priming paradigm, Swinney (1979) showed that both meanings of an ambiguous word are accessed briefly before the contextually appropriate meaning is selected. A number of eye-movement studies by Duffy, Rayner and colleagues (Dopkins, Morris, & Rayner, 1992; Duffy, Morris, & Rayner, 1988; Rayner, Cook, Juhasz, & Frazier, 2006; Rayner & Duffy, 1986; Rayner & Frazier, 1989; Sereno, O'Donnell, & Rayner, 2006; Sereno, Pacht, & Rayner, 1992) have shown that the frequency of each meaning of an ambiguous word has an impact on processing. Specifically, readers take longer to process *balanced* (both meanings are approximately equal in frequency) as compared to *biased* (one meaning is more frequent than the other) ambiguous words presented in a neutral context, showing that only the dominant meaning of a biased word is maintained for lexical selection, but both meanings of balanced words are maintained. Furthermore, contextual information has an effect on the selection of meaning of an ambiguous word: in a biasing context, balanced words receive equally long fixations as compared to unambiguous words, but if the context disambiguates in favour of the subordinate meaning, biased words will receive longer fixations. From the evidence available, it appears that while both meanings of an ambiguous word are initially activated, both frequency and context constrain the availability of those meanings so that only the correct meaning is maintained (although only when both words are in the same syntactic category: see Folk & Morris, 2003).

The available evidence on lexical processing suggests that there are two aspects of lexical processing which can be viewed as separate: lexical access whereby an item is retrieved from long term memory which corresponds to the visual stimulus that is being perceived, and lexical identification: the so-called 'magic moment' (Balota, 1990), whereby information that is stored with this lexical item such as its meaning and syntactic category are made available. Lexical factors such as word frequency, as well as neighbourhood density and frequency, affect the length of time needed to retrieve the lexical item from memory, and the word can be accessed through both phonological and orthographic representations although it is still not clear if these two pathways are used in concert, or whether they function independently in adults. Top-down factors such as context can also have an effect on how long lexical identification takes, particularly when choosing between two or more meanings of an ambiguous word. Importantly, the final stage of lexical processing (lexical identification), during which semantic and syntactic information about a word becomes available, is required for syntactic processing, discussed in the following section, during which the relationships between individual words can be computed.

1.4 Syntactic processing

When a word is identified during sentence processing, its syntactic category (e.g. noun, verb, adjective, determiner) becomes available, and with this information, combined with the application of grammatic conventions, it is possible for the reader to compute the structural relations between the constituents of a sentence. Words in a sentence are sequenced according to the rules of grammar in this way so that the structural relationships between different elements of the sentence can be understood by the reader (or listener). How the reader computes the structure of a sentence is known as syntactic processing or parsing.

The syntactic parsing mechanism (the parser) builds structure using the order in which elements in a sentence are arranged and according to parsing principles which guide it. However, there are potentially many alternative syntactic analyses of a sentence. When faced with ambiguity in this way, the parser must either construct all possible alternative syntactic structures and then eliminate those ruled out when an incoming word shows them to be impossible (the constraint-based processing hypothesis: MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994), choose one possible syntactic structure according to systematic decision preferences and go back to revise this original analysis should it prove to be incorrect (serial models of processing: Frazier, 1978; Frazier, 1987; Frazier & Rayner, 1982; Kimball, 1973; Rayner, Carlson, & Frazier, 1983), or delay interpretation until disambiguating information specifies the correct analysis (minimal commitment hypothesis - Weinberg, 1993). In this review of the literature relevant to syntactic processing, the first two types of model will be discussed in some detail. However, the minimal commitment hypothesis has received much less attention in the literature due to the considerable memory requirements of the model, and will therefore not be discussed in detail here.

Arguably, the most influential model of syntactic processing - the Garden Path model (Frazier, 1978; Frazier & Rayner, 1982) - falls into the second category: there exist basic principles which the parser adheres to in preferentially constructing one structure over possible alternatives. Consider sentence (1) from Bever (1970) below:

(1) The horse raced past the barn fell.

Sentences such as (1) are known as garden-path sentences as they are (syntactically) ambiguous and the parser is initially "led down the garden path" to the wrong structure. In sentence (1), the parser initially interprets the sentence as a simple active construction, when in fact *raced past the barn* modifies *horse* (*raced* being a past participle rather than the simple past tense) and *fell* is in fact the main verb of the sentence. When the reader fixates the disambiguating word *fell*, processing is disrupted, as the parser's initial interpretation of a simple active construction cannot incorporate the main verb and so the parser detects a misparse and must then reanalyse the sentence.

The Garden Path model proposes that words in a sentence are assigned an initial syntactic analysis on the basis of just two general principles: Minimal Attachment and Late Closure. Minimal Attachment states that readers attach incoming material to the phrase being constructed using the fewest levels of syntactic nodes as possible (see Chapter 6, Section 6.1.1 for a more detailed explanation).

(2a) The girl knew the answer by heart.

(2b) The girl knew the answer was wrong.

For example, in sentences (2a) and (2b) the parser would interpret the phrase *the answer* as being the direct object of the verb *knew* even though in sentence (2b) this is not the correct interpretation as *the answer* is actually the subject of a new clause. The analysis of sentence (2b) requires an additional syntactic level as compared to sentence (2a) and thus violates the principle of Minimal Attachment.

The principle of Late Closure states that readers attach new items to the phrase or clause currently being developed rather than starting a new phrase or clause, if this is grammatically permissible.

(3a) Since Jay jogs a mile this seems like a short distance to him.

(3b) Since Jay jogs a mile seems like a short distance to him.

According to the principle of Late Closure, sentence (3a) will be parsed correctly but sentence (3b) will be initially parsed incorrectly as the reader will initially interpret the phrase *a mile* as a direct object of the verb *jogs*, rather than as the subject of the following clause.

An important and controversial question in relation to syntactic processing is whether the parser is autonomous. While proponents of the Garden Path model (e.g. Ferreira & Clifton, 1986; Frazier, 1978; Frazier, 1987; Frazier & Rayner, 1982; Rayner et al., 1983) argue that the parser operates solely according to the principles of Minimal Attachment and Late Closure during the construction of the initial analysis, other theorists (Altmann & Steedman, 1988; Crain & Steedman, 1985; MacDonald et al., 1994; Marslen-Wilson, 1975; Tanenhaus, Carlson, & Trueswell, 1989; Tyler & Marslen-Wilson, 1977) have argued that parsing is interactive and makes use of information from post-syntactic stages of processing (in particular the semantic processor) when choosing which analysis to favour: constraint-based processing. Whether or not syntactic processing is interactive has important consequences for our understanding of the basic structure of the language processor, most notably in terms of whether it conforms to the assumptions of modularity (Fodor, 1983).

Fodor (1983) claimed that certain mental faculties are organised into two distinct types of systems: a number of input systems (modules) and a central processing system. Modules are informationally encapsulated and operate independently without access to information from other modules. The Garden Path theory applies Fodor's theory to language processing in its claims that the parser is independent from higher-level and non-linguistic knowledge sources.

The central issue with respect to the debate concerning autonomy and interaction calls into question the fundamental architecture of the language processor: while no-one would deny that semantic and real world knowledge are involved in sentence processing, whether these sources of knowledge are employed during initial processing determines whether syntactic processing works in a modular fashion or not.

The debate regarding the autonomy of the parser is best characterised by two competing positions: a serial autonomous model of syntactic processing (the Garden Path model) and a constraint-based model in which all levels of representation interact freely. Numerous experimental studies have investigated the issue of the autonomy with respect to syntactic processing and how this relates to the modularity of the language system. While a comprehensive review of this literature is beyond the scope of this thesis, those studies which best summarize the development of theoretical understanding in this field, and which are most relevant to the issues raised in this thesis, will be detailed. For clarity, the studies will be outlined roughly in chronological order.

Importantly, Frazier, Rayner, and colleagues (Ferreira & Clifton, 1986; Frazier, 1978; Frazier & Rayner, 1982; Rayner et al., 1983) provided strong confirmatory experimental evidence for the Garden Path model and for its two principles of Minimal Attachment and Late Closure. Such data also served as evidence against constraint-based (interactive) processing (and minimal commitment) hypotheses. In their 1982 paper, Frazier and Rayner monitored participants' eye movements as they read Minimal Attachment and Late Closure sentences (like sentences (2a-b) and (3a-b) respectively) and found that readers looked longer at syntactically disambiguating words (such as *seems* in sentence (3b) and *was* in sentence (2b)) in sentences which violated the principles of Late Closure and Minimal Attachment as compared to sentences which were in accordance with those principles. They argued that the parser, therefore, exhibits a processing preference for particular syntactic structures, and this preference is captured by the principles of Late Closure and Minimal Attachment. In this way, the parser does not compute all possible syntactic structures but rather only one, and if that isn't appropriate then the system must go back and recompute.

In their 1983 paper, Rayner et al. further investigated the possibility that semantic information might influence initial parsing decisions. In Experiment 1, they showed that real world knowledge (in this case the relative plausibility of two possible real world events) did not influence the syntactic processor's initial choice of analysis. In Experiment 2, they monitored adults' eye movements as they read sentences such as (4a) and (4b) below;

(4a) The spy saw the cop with the binoculars but the cop didn't see him.

(4b) The spy saw the cop with the revolver but the cop didn't see him.

They found that participants took longer to read sentences such as (4b), where the non-Minimal Attachment interpretation was more plausible, than sentences like (4a) where the Minimal Attachment interpretation (i.e. the spy used the binoculars as an instrument to see the cop) was more plausible. They interpreted these results as evidence that the syntactic and semantic processors operate independently and that the syntactic processor initially computes only the structurally preferred analysis of a sentence, after which the thematic processor assesses the plausibility of that interpretation on the basis of real world knowledge. According to this argument, modularity is not violated.

On the basis of this evidence, there was support for the Garden Path theory, and indeed for the argument that the syntactic processor operated in a modular fashion, independently of post-syntactic sources of knowledge. However, following Frazier and Rayner's early work, and based on an early study by Marslen-Wilson (1975), a number of studies in support of a constraint-based model of language processing began to emerge. The idea that the semantic content could influence decisions regarding syntactic processing was appealing as it seemed to fit well with research into lexical ambiguity resolution (see Chapter 5). Indeed, Macdonald, Pearlmutter and Seidenberg's (1994) constraint-based theory was a system for ambiguity resolution which was implemented within a lexicalist framework. Moreover, this position was in direct opposition to the Garden Path model, in that it claimed that language processing is highly interactive. Proponents of this position (e.g. McRae, Spivey-Knowlton, & Tanenhaus, 1998; Trueswell et al., 1994) argued that multiple syntactic analyses are activated in parallel and weighted on the basis of how compatible they are with a range of constraints (including semantic cues), and that syntactic representations and constraints interact freely with other levels of representations. In this way, parsing is a one-stage process with no distinction between initial and later stages of processing (see also Tyler & Marslen-Wilson 1977).

One of the first studies to challenge the Garden Path model with empirical evidence was by Crain and Steedman (1985) who used a word-by-word, fixed presentation rate reading task to investigate whether referential contexts could induce or nullify garden path effects by placing Minimal Attachment sentences in different contexts. In previous experiments, they argued, sentences had been used in isolation (the null-context), and the presentation of sentences in null-contexts were the cause of garden path effects, rather than the principles of Minimal Attachment and Late Closure, as stipulated by the Garden Path theory. Crain and Steedman argued that there were no intrinsically garden-pathing structures, but rather, for any given sentence, there were certain contexts which could induce garden paths, and certain others which would not. In this way, it was the referential suppositions associated with the preceding context which determined the way a reader initially interprets a syntactically ambiguous sentence, and once context was controlled appropriately, no residual effects of structural mechanisms would be apparent. The primary responsibility for the resolution of syntactic ambiguities then, they argued, rests with the immediate almost word-by-word interaction between syntax and reference to the (semantic) context.

Crain and Steedman (see also Altmann & Steedman, 1988) proposed a referential theory of parsing according to which alternative analyses are constructed in parallel by the parser and contextual information is used to adjudicate between them immediately. They described their account of processing as *weakly interactive* in that, unlike *strongly interactive* accounts, the semantic processor does not influence which syntactic entities are constructed in the first place, but rather decides whether analyses which have already been constructed should be abandoned. Importantly, this distinction meant that Fodor's (1983) modularity hypothesis was not compromised.

Following Steedman and colleagues' (Altmann & Steedman, 1988; Crain & Steedman, 1985) study which found evidence against the Garden Path model, Ferreira and Clifton (1986) provided further support for serial syntactic processing. They criticised Crain and Steedman's methodology, arguing that only an online method such as eye-movement recording could differentiate between initial and eventual use of context. As mentioned, there is no dispute as to whether contextual information is made use of during language processing, but rather whether it is called upon during the initial stages of syntactic processing. Ferreira and Clifton conducted a series of experiments in which they manipulated the animacy of noun phrases (*the defendant examined* versus *the evidence examined*) and included a biasing semantic context before target sentences. They showed that the principle of Minimal Attachment holds even in natural language parsing settings such as discourse; and even when there are thematic cues (i.e. animacy information) to help the reader to decide between possible syntactic analyses, readers do not make use of these but rather blindly follow generic syntactic principles.

The Rayner et al. (1983) study was also criticised methodologically, calling into question the stimuli used. Taraban and McClelland (1988) used Rayner et al.'s stimuli and also created their own which they controlled for frequency and pre-screened in detail. Using a self-paced listening task, they replicated Rayner et al.'s results using the Rayner et al. stimuli, but found the opposite pattern of effects (i.e. readers exhibited longer reading times on the non-Minimal Attachment sentences) using their own materials, showing not only that Rayner et al.'s materials were not as tightly controlled as they could have been, but also that when materials were properly controlled, semantic factors guided initial attachment preferences. They argued that readers generate content-based expectations during sentence processing and use these to guide parsing, in line with constraint-based models of syntactic processing.

Trueswell, Tannenhaus and Garnsey (1994) criticised Ferreira and Clifton's (1986) materials which, although manipulated for animacy, used examples of objects which could perform actions (e.g. *the car towed*) or which had ergative readings (*the trash smelled*). Trueswell et al. adapted and improved these materials and found that participants did use thematic cues to resolve syntactic ambiguities, providing further evidence against the Garden Path model (although see Clifton, Kennison, & Albrecht, 1997 for evidence to the contrary). Like Taraban and McClelland (1988), they argued that their results fitted well with a interactive constraint-based model of syntactic processing in which "the process of constraining ambiguity in one domain is accomplished by recruiting information from other relevant domains" (MacDonald et al., 1994, pp 308).

Finally, Clifton, Traxler, Mohamed, Williams, Morris and Rayner (2003) used the materials from Trueswell et al. (1994) and conducted exhaustive eye movement analyses to investigate whether semantic cues had an effect on initial parsing preferences. Although like Trueswell et al., first pass reading times were reduced by the semantic bias of an inanimate subject (*the evidence examined* as compared to *the defendant examined*), later measures revealed significant disruption to processing in both inanimate-subject and animate-subject sentences, showing that there was a cost associated with processing a non-minimal attachment construction, independent of semantic cues.

Overall, conclusive empirical evidence which distinguishes between serial and constraint-based architectures in parsing has not yet been found. Serial models are appealing in that they are conceptually simpler, and have fewer memory requirements (i.e. they are also computationally simpler), while parallel models, although more complex and less specified, can arguably provide a more unified account of language comprehension (MacDonald et al., 1994). The fact that 'high-level' linguistic variables that affect sentence comprehension processes (as compared to lower-level lexical factors) are very complex, both in their definition and effects, means that as yet a fully comprehensive and explicit theory of syntactic processing has not yet succeeded in explaining all the empirical data available (Clifton et al., 2007).

In terms of relevance to the thesis, the issue of serial versus constraint-based processing and how it relates to modularity is not as important as the psycholinguistic manipulations employed in the studies outlined above and what they have told us regarding adults' syntactic processing preferences. It is a robust finding that adult readers exhibit disruption to processing in non-Minimal Attachment sentences (in the null context) and therefore the psycholinguistic manipulations that were used to reveal these effects can be used a diagnostic tool in order to investigate for the first time whether children also exhibit these preferences. Whether the effects are accounted for by a serial or parallel model of processing, and whether the findings have implications for the assumptions of modularity are not as important in relation to the aims of this thesis as the empirical evidence that children behave in a certain way in response to the psycholinguistic manipulation employed. Experiments 5 and 6 of the thesis use two manipulations, which have generated robust syntactic processing preferences in adults, in order to examine how children process these same kinds of syntactic structure online.

1.5 Semantic processing

Up to this point we have seen how the identification of words necessarily precedes the building of syntactic structure, and how both these processes are critical to sentence processing. However, the primary function of language is to express meaningful content, and while word identification and parsing operate in service of this function, it is semantic processing which actually results in a representation of sentential meaning. There has been a great deal of research investigating many different aspects of semantic processing during sentence reading in adults. However, following a brief overview of semantic processing, the focus of this section will be on one aspect of (shallow) semantic processing which will be directly investigated in Chapter 7: thematic processing.

Following lexical identification, there is a process of integration during which the meaning of the word is incorporated into the overall representation of the sentence meaning and this is done roughly on a word-by-word basis (Traxler et al., 1997). The objective of reading is to achieve a coherent mental representation of what is being communicated and theories of discourse processing have sought to specify how this coherence is achieved. Faced with a discourse, a reader must do several things: (1) she must determine the referent of referring anaphoric expressions (such as pronouns and noun phrases) and whether the discourse is introducing a new entity or whether it is referring to an existing one; (2) she must determine how any assertion made in a sentence is related to the previous discourse; (3) she may need to make inferences when assertions are not explicitly stated; and (4) she must create a non-linguistic representation of the content of what she is reading. Clearly, these processes are complex and real world knowledge plays a significant role in understanding the meaning of a sentence. As such, not all semantic processing is derived from the text itself.

For the purposes of this thesis, the focus within semantic processing will be entirely on thematic processing which refers to processing that results in a very shallow representation of meaning. Thematic processing has been argued to be the intermediary between relatively low-level processes (lexical identification and syntactic processing) and higher level discourse parsing (Liversedge, 2003). When a reader encounters a verb, the verb's subcategorisation frame and thematic grid become available. While the subcategorisation frame specifies the syntactic categories for the arguments of the verb, the thematic grid (the abstract specification of the thematic role possibilities for each predicate) specifies the meaning of those same categories. A thematic role of a verb argument refers to the part played by an entity (as denoted by a noun) in an event (as denoted by a verb). For example, in sentence (5) below, the verb *cut* takes three arguments: an agent (e.g. a *butcher*), an instrument (e.g. a *knife*) and a patient/theme (e.g. a *steak*).

(5) The butcher used the sharp knife to cut through the steak.

Understanding sentence (5) above requires knowledge about cutting events, butchers, steaks, knives, and their interrelationships to know that *butcher* is the agent, *steak* is the patient, and *knife* is the instrument. Verb arguments are an intrinsic part of the verb's meaning and are central to understanding not only the verb but also the sentence as a whole. There are a finite number of entities which can plausibly be assigned thematic roles given a particular verb, and not all the thematic roles for a particular verb must necessarily be explicitly filled (Mauner, Tannenhaus & Carlson, 1995), although some verbs, such as *put* do require this (i.e. it is not legal to say, *He put the cup*). Irrespective of this, if a thematic role is filled, it must be plausible with regard to the reader's knowledge of the kinds of events denoted by the verb if the reader is not to experience disruption to processing.

In a relatively early study, Marslen-Wilson, Brown and Tyler (1988) participants had to detect a target word while listening to sentences such a 6a-6c below. Results showed longer monitoring latencies when listening to sentences such as (6b) and (6c) as compared to control sentence (6a).

(6a) John carried the guitar. (control)

(6b) John buried the guitar. (pragmatic anomaly)

(6c) John drank the guitar. (semantic anomaly)

Sentence (6b) is anomalous in a pragmatic sense and inferences about the real world are necessary in order to recognise the anomaly. While it is certainly unusual for people to bury guitars, it is quite possible to find a real world context in which such an action would be perfectly acceptable. In contrast, sentence (6c) does not require knowledge about the real word, only a concept of 'something drinkable' and of a guitar as 'something solid' (and therefore non-drinkable) to detect the anomaly. The longer monitoring latencies in response to both pragmatic and in particular semantic anomalies, as compared to control sentences, suggest that both types of anomaly had disruptive effects on processing and that non-linguistic domains of interpretation and inference are quickly integrated with thematic properties of verb argument frames. Note, though, that data from spoken language comprehension may not generalise to reading.

Rayner, Warren, Juhasz and Liversedge (2004) investigated the effects of thematic plausibility on adults' reading behaviour (see also Braze, Shankweiler, Ni, & Palumbo, 2002; Filik, 2008; Ni, Fodor, Crain, & Shankweiler, 1998; Warren & McConnell, 2007; Warren, McConnell, & Rayner, 2008). Rayner et al. used sentences that described events in which an individual performed an action with an instrument. In each case, the verb had three thematic roles (see sentences 7a-c): an agent (*John*), an instrument (*knife, axe or pump*) and a patient/theme (*carrots*).

(7a) John used a knife to chop the large carrots for dinner.

(7b) John used an axe to chop the large carrots for dinner.

(7c) John used a pump to inflate the large carrots for dinner.

Rayner et al. found that disruption to processing occurred earlier when the sentences were anomalous (7c) rather than implausible (7b). They suggested that the differential effects may be due to either the severity of the violation (i.e. how implausible it is perceived to be), or because, in most cases, anomalous violations can be detected on the basis of lexical information alone (a verb argument violation). Implausible violations, on the other hand, can only be detected at a later stage of processing after the semantic evaluation of the combination of a verb and the objects involved in the event it denotes. The delay in disruption for implausible compared with anomalous thematic roles has since been replicated (Joseph et al., 2008; Warren & McConnell, 2007), and indeed even contextual information does not eliminate the immediacy of anomaly effects (Filik, 2008; Warren et al., 2008). Despite these now robust effects of anomaly and implausibility in thematic processing, it is not yet known whether this same pattern of effects is observed in children during reading. The rôle of plausibility in thematic role assignment will be directly investigated in the final experiment of this thesis (Chapter 7).

1.6 Eye movements during reading

While much has been learned about the acquisition of spoken language and the development of reading skills using offline methodologies, the use of online methods such as eye tracking can provide unique insights into the moment-to-moment comprehension of written language, rather than assessing the product of comprehension, as offline methods necessarily do. It is for this reason that this thesis focuses on eye movement recording to examine children's online language processing during reading. In this section, an overview of the field of eye movements will be given which will include an explanation of how the eyes move during reading, important issues in eye movements during reading, and finally, a brief description of the dominant models of eye movements during reading.

When we read text, our eye movements are not smooth but rather are made up of a sequence of stable fixations (usually 200-250ms in adults) which are separated by brief ballistic eye movements, known as saccades (typically 7-9 character spaces, though there is much variability, Rayner, 1998). Saccades are made in order to bring a new region of text into foveal vision (the central 2° of vision, usually corresponding to 6-8 characters) because it is difficult or impossible to read text presented only in the parafovea (Rayner & Bertera, 1979) However, readers are able to shallowly process text which is not being directly fixated during normal reading (this issue is discussed further below).

While saccades are needed to bring different portions of text into view, it is during fixations that visual and linguistic information is extracted. The duration of a fixation reflects the difficulty the reader is experiencing in processing the fixated portion of text. This assumption, that there is a close correspondence between eye gaze and cognitive processing (Just & Carpenter, 1980), provides the basis for much eye movement research into reading. If we assume that the eye remains fixated on a word as long as that word is being processed, then it is possible to make meaningful inferences about underlying linguistic processes. These assumptions are warranted by the large body of evidence which shows that linguistic characteristics of a text, such as those outlined in Chapter 1, directly influence how long it takes to read that same text. Short fixations generally indicate that a word has been easily processed while long fixations indicate relative difficulty of processing. Likewise, most saccades are forward saccades (from left to right in English) and are indicative of successful processing while 15% of saccades are (leftward) regressive saccades (Rayner, 1998). Short within-word regressions may be due to oculomotor error, or else may indicate that a reader is experiencing problems processing the fixated word. Longer regressions (ten characters or more) tend to suggest that the reader is experiencing difficulty with their ongoing comprehension of the text.

While most words are fixated during reading, a significant minority of words are skipped (i.e. not fixated), in particular short words and function words. As word length increases, the probability of fixating a word increases (Rayner & McConkie, 1976). Other words, particularly long, content words, and words which are difficult to process (such as low frequency or unpredictable words) are refixated, that is fixated more than once. In the field of eye movements during reading, in addition to taking measures of reading time (i.e. fixation durations), whether a word is skipped, fixated once, or refixated, is calculated in order to make further inferences about a reader's ongoing cognitive processing of written text.

As mentioned, our eyes move roughly from one word to the next during reading in order to bring a new portion of text into foveal vision because the amount of information that can be extracted from a single fixation is limited. The perceptual span refers to the 'size of the effective vision' (Rayner, 1986, pp 212), that is, how many characters of text are visually available to a reader during a single fixation. Rayner and colleagues (McConkie & Rayner, 1975, 1976; Rayner & Bertera, 1979; Rayner, Well, & Pollatsek, 1980) used the moving window technique whereby a "window" of unaltered letters was visible to the reader but which moved contingent on the reader's gaze as they proceeded through the sentence. All letter spaces outside the window were replaced by a mask, thereby concealing letter identities, and sometimes word boundaries. In this way, the amount of information available to the reader could be strictly controlled on a fixation-by-fixation basis to examine the point at which reading was disrupted. Rayner and colleagues found that adult readers had perceptual spans of 3-4 characters to the left of fixation, and 14-15 characters to the right, showing that no information is extracted beyond this limited area.

One issue which has provoked controversy in the field of eye movements during reading is how much information can be acquired parafoveally, that is, to the right of fixation. Partial-word information, as well as word length information, from the word to the right of fixation is known to be used by readers to facilitate fluent reading (Rayner, 1998). Indeed words are skipped because sufficient parafoveal pre-processing of the word to the right of fixation makes this possible. However, the extent and depth to which parafoveal words are processed prior to direct fixation has received considerable attention in the field of eye movement research in recent years, largely because the question of whether parafoveal words are lexically identified prior to fixation has direct implications for whether words are identified serially or in parallel, which could potentially help to discriminate between competing models of eye movements during reading (this will be discussed further below). Previous research has found that orthographic (Lima & Inhoff, 1985; Pynte, Kennedy, & Ducrot, 2004; Vitu, Brysbaert, & Lancelin, 2004) and phonological (Pollatsek et al., 1992) information can be extracted from a parafoveal word. However, manipulating lexical-level aspects of a word has yielded inconsistent results (Henderson & Ferreira, 1993; Hyönä & Bertram, 2004; Inhoff, Radach, Starr, & Greenberg, 2000; Inhoff & Rayner, 1986; Inhoff, Starr, & Shindler, 2000; Kennedy, 2000; Kennedy, Murray, & Boissiere, 2004; Murray & Rowan, 1998; Murray, 1998; Rayner, White, Kambe, Miller, & Liversedge, 2003). This issue of to what degree a parafoveal word can be pre-processed remains contentious and is discussed further in Chapter 7.

The programming of eye movements during reading can be categorised into two classes of decisions: the where decision (fixation location) and the when decision (fixation duration) and there is substantial evidence to show that these two decisions are made independently (Pollatsek, Reichle, & Rayner, 2006; Rayner & McConkie, 1976; Rayner & Pollatsek, 1981). Decisions concerning where to fixate appear to be determined by low-level visual aspects of the text such as word length (although lexical-level characteristics of the text can influence word skipping, Drieghe, Rayner, & Pollatsek, 2005), as well as the location from which the saccade was launched (McConkie et al., 1988), and, as discussed in Chapter 1 (Section 1.2), adult readers tend to initially fixate halfway between the beginning and the middle of a word (McConkie et al., 1988; Rayner, 1979): the so-called *Preferred Viewing Location*. The PVL has been contrasted with the *Optimal Viewing Position* (OVP) which is defined as the location within a word at which recognition time is minimised, and is located close to the word centre, a little to the right of the PVL. O'Regan and Levy-Schoen (1987) investigated fixation locations on words presented in isolation and discovered two main consequences of not fixating the OVP. First, there is a refixation cost in that the further the eyes are from the OVP, the more likely the reader is to refixate. Second, there is a processing cost in that for every letter that a reader's fixation deviates from the OVP, there is a processing cost of approx. 20ms (i.e. the reader will require an additional 20ms to identify the word). While the processing cost does not hold in normal text reading (indeed an inverted OVP curve is observed: 2001), the refixation cost does, suggesting that where a reader fixates a word directly affects the ease with which it can be processed.

In contrast to the 'where' decision, the 'when' decision appears to be governed by the ease or difficulty associated with processing a word (e.g. the word frequency effect - Henderson & Ferreira, 1990; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980), although there is evidence that low-level non-linguistic factors also influence fixation durations (Vitu et al., 2001). As reviewed in Chapter 1, many linguistic factors have been shown to influence how long readers look at a word, including word frequency and predictability. The majority of experiments described in this thesis will focus on the 'when' decision and how different linguistic manipulations influence fixation durations during reading. However, Experiment 1 (Chapter 3) will examine the 'where' decision in the context of landing position effects in adults and children.

There are two fundamental issues which have caused much controversy in the eye movement literature. The first is whether eye movements are driven primarily by cognitive (i.e. linguistic) or oculomotor factors; and the second is whether attention is distributed serially or in parallel (according to a gradient) during reading. Two broad classes of models are divided regarding which factors drive eye movements during reading: *cognitive* (or processing) models and *oculomotor* models. Cognitive models are based on the assumption that eye movements are driven by ongoing cognitive processing while oculomotor models maintain that eye movements are mainly controlled by low-level visuomotor processes and are only indirectly related to language processing. Cognitive models focus primarily on the temporal aspect of eye fixations (although they also take account of the spatial aspect) while oculomotor models focus predominantly on the spatial aspect of eye fixations. Although this thesis does not set out to assess these different types of models, nor to discriminate between them, it is useful to provide here a brief outline of two of the most prominent models and how they address this issue.

The E-Z Reader model (Pollatsek, Reichle, & Rayner, 2003; Pollatsek et al., 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle et al., 2004) is, arguably, at present the most advanced cognitive model. Its central assumptions are: 1) that a stage of word identification is a signal to move the eyes (i.e. it is a linguistic factor that triggers an eye movement), 2) attention is allocated from one word to the next on a strictly serial basis (although saccades can be programmed in parallel). It does not claim to account for higher-level processes that influence reading such as syntactic or semantic processing, but rather describes 'default' processing that accounts for the forward progression of the eyes through text when disruption to comprehension does not occur (although see Reichle, Liversedge, Pollatsek, & Rayner, 2009; Reichle et al., 2004). On fixating a word, three stages of processing occur. First, the word is visually processed, and following this, there are two lexical processing stages which culminate in the word's meaning becoming available for further linguistic processing. The completion of the first stage of lexical processing (L1) signals the oculomotor system to program a saccade to the next word, and the completion of the second lexical processing stage (L2) signals the attention system to shift attention to the next word. It is clear from this description that, according to this model, when an eye movement is made is determined by the linguistic processing of the text.

The Saccade generation With Inhibition by Foveal Targets (SWIFT) model (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Kliegl & Engbert, 2003) also assumes that eye movements are driven by word recognition, although unlike E-Z Reader, the signal to move the eyes is generated by a random timer rather than the completion of a stage of word processing (but inspection times can be extended for difficult-to-process words). Like E-Z Reader, words are identified in two stages and saccades are programmed in two stages, although they are autonomously generated. However, the SWIFT model differs from E-Z Reader in how attention is allocated: this issue will be discussed below.

In contrast to cognitive models, in oculomotor models, (e.g. O'Regan, 1992; O'Regan, 1990; O'Regan & Lévy-Schoen, 1987; Yang & McConkie, 2004; Yang & McConkie, 2001), it is visuo-oculomotor factors that are prominent in controlling the eyes during reading, and research in this tradition has sought to identify non-cognitive factors which influence when and where the eyes move. In particular, it has been shown: that the eye's initial landing position determines the length of the fixation and where the following fixation will be made (McConkie et al., 1988; O'Regan & Lévy-Schoen, 1987; Vitu et al., 2001); that saccade length varies with eye position on a screen (central versus peripheral: Vitu, Kapoula, Lancelin, & Lavigne, 2004); and that regressions are more likely following longer forwards saccades (Vitu & McConkie, 2000). Furthermore, many of these effects remain even when words are replaced by Z-strings (Vitu, O'Regan, Inhoff, & Topolski, 1995).

In addition, according to one influential oculomotor theory proposed by O'Regan (e.g. O'Regan, 1992, O'Regan, 1990, O'Regan & Levy-Schoen, 1987), readers employ within-word tactics which are based on low-level non-lexical information (e.g. word length) obtained early in a fixation. These tactics explain Optimal Viewing Position (OVP) effects, discussed in detail in Chapter 3, whereby fixation durations are shortest and the probability of refixating a word is lowest, when the optimal location (close to the word centre) of a word (presented in isolation) is fixated. In this way, it is oculomotor constraints that determine fixation durations and refixation probabilities, not cognitive factors. Linguistic factors do influence the duration of long single fixations, and the second of two fixations on a word, but do not drive all eye movements. However, while oculomotor models can easily explain word length effects and other effects related to the visual processing of words, they have more difficulty explaining the robust effect of word frequency, as well as higher-level processing effects.

The second issue which is central to research in eye movements during reading concerns the way in which attention can be allocated during word processing. On this issue, models can be divided into *serial-attention-shift* (SAS) models, in which attention is allocated sequentially to only one word at a time, and *guidance-by-attention-gradient* (GAG) models, in which attention is distributed as a gradient of processing that usually encompasses more than one word at a time. The E-Z Reader model (Pollatsek et al., 2006; Reichle et al., 2004) is currently the most fully developed SAS model and assumes that attention is allocated serially and is word-based, and proponents of this model argue that by shifting the focus of attention from one word to another, readers are able to process each word in its correct order (although see Kennedy & Pynte, 2008; Pollatsek et al., 2006). If this assumption of serial processing is correct, then although it is perfectly possible (and indeed typical) for some shallow orthographic processing of the word to the right of fixation ($word_{n+1}$) to take place while fixating $word_n$ (once the saccade to the next word has been programmed), parafoveal pre-processing of a word should not occur at the same time as the processing of the fixated word, and properties of $word_{n+1}$ should not affect fixation times on $word_n$ (known as parafoveal-on-foveal effects).

The most advanced GAG model is the SWIFT model (Engbert et al., 2002; Engbert et al., 2005; Kliegl & Engbert, 2003). Unlike E-Z Reader, SWIFT assumes that attention is distributed continuously over a gradient thus allowing for the simultaneous lexical processing and identification of more than one word. Although proponents of serial processing models argue that this is an untenable position (e.g. Reichle et al., 2009), the assumption that parallel processing of words is possible means that the processing of fixated $word_n$ and parafoveal $word_{n+1}$ in parallel is permitted and thus inflated fixations on $word_n$ when $word_{n+1}$ is difficult to process can be explained by GAG models, although as noted above, evidence in support of higher-level (e.g. semantic) pre-processing is somewhat inconsistent (Henderson & Ferreira, 1993; Kennedy, 1998; Murray, 1998; Rayner, Fischer, & Pollatsek, 1998). The issue of parafoveal-on-foveal effects will be specifically addressed in Chapter 7 and has profound implications for the issue of serial versus parallel allocation of attention.

There are other examples of both cognitive and oculomotor models, as well as alternative SAS and GAG models, but a full review would be beyond the scope of this thesis. For the purposes of the current work, it is important to contrast models with serial versus parallel allocation of attention, and models in which linguistic factors primarily do or do not drive eye movements. The results from some the experiments in this thesis will have implications for these distinctions and will be discussed in relation to the models outlined above, where relevant. The distinctions are also important in terms of understanding children's eye movements during reading and taking steps towards incorporating child data into general models of eye movement control during reading (although again this is well beyond the scope of the current work).

1.7 Summary and conclusions

This chapter has provided a brief overview of some aspects of the vast literature on written language processing in adult readers. Language processing can be seen not as one single process but as the culmination of several sub-processes. First, a word must be visually encoded in order for word recognition processes to begin. Word recognition consists of first accessing a lexical item from memory and then accessing information about that item in order that it can be identified uniquely in the lexicon. Once lexical identification has taken place, syntactic structure can be built. Finally when the structural relationships between the words in the sentence have been computed, a representation of semantic meaning of the sentence can be obtained.

A large number of linguistic factors have been shown to influence eye movement behaviour during reading, including the frequency with which a word is encountered, the age at which a word was acquired, the number of phonological or orthographic neighbours that a word has, whether a word has more than one meaning, the syntactic structure of a sentence, and the plausibility of a sentence. While language processing is extremely complex, the wealth of data that has come from the numerous studies in this area means that we now have a reasonably good understanding of the mechanisms that adult readers have in place to process written language online.

It is from this point that we can begin to examine how children process written language, as the robust effects found in adult readers can be used diagnostically to assess whether children exhibit these same preferences, and whether the time course and magnitude of the effects observed is the same in adults and children. While this is an as yet unexplored area of research, there is a large literature pertaining to children's language acquisition and reading development, albeit using offline methods. The next chapter will review the literature that has examined children's language development and what is known about how children learn to read, as well as giving an overview of the field of eye movements during reading.

Chapter 2 : Children's language development and eye movements during reading

Chapter 1 gave an overview of the mature language processing system and what is known about visual, lexical, syntactic and semantic processing in adult readers. However, this thesis is primarily concerned with children's language processing, and therefore it is important to consider in some detail what is known about children's language development. While relatively little is known about children's online processing of written language, there is a large literature which has investigated language acquisition; as well a substantial literature examining children's reading development using offline methods. This chapter will review this work and describe in detail the few studies which have used eye movements to examine children's reading behaviour. The chapter will also provide an introduction to the field of eye movements during reading, as this is the primary method of investigation used in the thesis.

Chapter 2 will be structured as follows. In Section 2.1 a brief overview of children's language acquisition will be given in order to build a context for the following section (Section 2.2) which will review the literature on children's reading development and address how beginning readers start to make associations between visual stimuli and meaning. Section 2.3 will provide a review of research examining children's online processing of written language using eye movement methodology. Finally, Section 2.4 will address how the issues raised in this introduction are relevant to the current work and will provide a rationale and structure for the remainder of the thesis.

2.1 Children's language acquisition

While the focus of this thesis is on reading development, a brief review of the literature pertaining to the acquisition of language from infancy through to school age is necessary in order to put reading development in context. In almost all cases, language development follows a predictable sequence, although there is a lot of variability in the time course of development. Of course auditory language comprehension develops much earlier than written language comprehension, and this section focuses on how infants and young children first understand and then produce spoken language.

While young infants (6 months) are able to discriminate between phonetic units used in many different languages, older infants (12 months) lose this ability and are able to distinguish only the sounds in their own language (Werker & Tees, 1984, 1999). From 6-8 months infants start to make extended sounds that are interrupted rhythmically by oral articulations into syllable-like sequences, known as babbling (Oller, 2000). They are also able to successfully segment words from fluent speech by the age of 8 months, by detecting consistencies in patterns of syllables which are predictable when part of the same word, but not predictable when they span word boundaries (Saffran, Aslin, & Newport, 1996). Prosodic cues can also help an infant to identify word candidates in natural speech (Kuhl, 2004) as 90% of multisyllabic words in English are stressed on the first syllable. When an infant produces his/her first words around the age of 12 months, their phonological output is simplified, with a general tendency to produce shortened strings of words, to omit the final consonant, to reduce consonant clusters, and to substitute easier sounds (such as those in their babbling repertoire) for more difficult sounds (Smith, 1973).

From approximately 12 months onwards, infants begin to utter recognisable words, usually in the context of naming, and early words tend to refer to either things that move (e.g. people, animals, vehicles), or things that can be moved. Children's semantic development is, of course, dependent on their conceptual development in that they can only map meanings onto concepts available to them at that time. First words emerge out of shared reference, and *motherese* (child-directed speech which contains more repetition, references to the here and now, exaggerated prosody, exaggeration of vowel sounds, and is syntactically and phonologically simplified) enables children to attend to the stressed parts of speech they hear, thereby emphasising the label of the referent being spoken of. Children preferentially treat novel words as labels for objects, seeking taxonomic, rather than thematic relations (Markman & Hutchinson, 1984), and furthermore assume that such labels and objects are mutually exclusive (Markman & Wachtel, 1988).

Children may use syntactic cues, such as whether a verb is transitive or intransitive (Naigles, 1996), in order to glean information about word meaning (Gleitman, 1990), as well as making cognitive inferences using knowledge of real world contingencies (Pinker, 1994). Nouns are acquired more easily than verbs (Gentner, 1983), which may be due to verbs' increased complexity as compared to nouns, but also their acquisition being dependent on the prior acquisition of some nouns, as well as a rudimentary grasp of some aspects of syntax.

From the age of 18 months, children pass into the *two-word* or *telegraphic* stage, in which words are combined to form two-word mini sentences with simple semantic relationships (e.g. *daddy gone*), in which grammatical morphemes are usually omitted. From 24 months onwards, words are combined using grammatical rules of syntax, although they may appear 'ungrammatical' to adult speakers. How children learn syntactic categories such as noun, verb and adjective is an issue of contention and while some theorists believe that such categories are innate (e.g. Pinker, 1984), increasingly research has shown how they are learnt either through semantics (Gleitman, 1981) or independent of semantics by looking for syntactic regularities in language input (e.g. Rumelhart & McClelland, 1986). Children will often over-regularise rules, using regular plural and past tense forms to exceptional cases, such as *mouses* instead of *mice*, and *singed* instead of *sang*. By three years, most children have large vocabularies, and are using complex sentences containing relative clauses, and by the age of four or five, use of language is usually comparable to that of an adult. It is at this advanced stage of language competency that many children begin the process of learning to read.

2.2 Children's reading development

While there has been remarkably little work investigating children's online comprehension of written language (though see Blythe et al., 2006; Buswell, 1922; Häikiö, Bertram, Hyönä, & Niemi, 2008; Hyönä & Olson, 1995; Joseph, Blythe, & Liversedge, 2009; Joseph et al., 2008; Joseph, Liversedge, Blythe, White, & Rayner, 2009; McConkie et al., 1991; Rayner, 1986; Taylor, 1965 for notable exceptions), there has been much research investigating children's reading acquisition and development, mostly using offline methods. Before reviewing the small literature on children's online written language comprehension, a brief overview of how offline methods have informed understanding of the process of learning to read will be given.

2.2.1 Precursors to reading development

Unlike learning to speak, learning to read (almost always) requires formal and explicit instruction and there are a number of prerequisites to successful reading acquisition. Perhaps most importantly, a child must understand the alphabetic principle: that is, the idea that there are systematic correspondences between the spoken and written forms of words. Phonemes are represented by letters in alphabetic orthographies, but in a non-transparent, irregular language such as English, these correspondences are far from straightforward as grapheme-phoneme explicitness is often sacrificed for the sake of symbol economy (e.g. the written letter *a* has many different pronunciations such as in the words *car*, *care*, *cat*, *cake*), as well as for the sake of morphological transparency (e.g. the past tense affix *–ed* has three different pronunciations). Despite these complexities, a beginning reader must understand the concept that abstract phonemes are associated with specific graphemes and use this knowledge to decode new words.

Rayner and Pollatsek (1989) cite five additional prerequisites to reading: (1) recognising letters; (2) discriminating left and right; (3) gaining control of eye movements; (4) becoming conscious of the word as a unit of language; and (5) developing phonological awareness. Recognising letters requires identifying the discriminating features of letters such as straight lines, curves, intersections etc (see Chapter 1, Section 1.2). This skill is clearly essential at the start of reading acquisition and indeed letter knowledge is the best predictor of early reading ability (Chall, 1967). Discriminating left from right, or more generally, acquiring directional skills is another important skill required for learning to read (Clay, 1970). Research has shown there to be developmental changes in the programming of visually guided saccades, such that saccadic latencies tend to be increased in children compared to adults although their saccade targeting is as accurate as that of adults (Cohen & Ross, 1978). However, more significant challenges in learning to read are Rayner and Pollatsek's final two pre-requisites: becoming conscious of the word as a basic unit of language and phonological awareness.

Word consciousness refers to the knowledge necessary for children (or adults) to learn and effectively use words. While it is often assumed that words are natural units of language and children view them as such, research has shown that this is not necessarily the case. As mentioned, in spoken language (the only form of language a child will be familiar with before they begin the process of learning to read) there are no clear acoustic boundaries between words as there are in written language; and Ehri (1975) and others have shown that beginning readers are not able to segment words effectively, and moreover that pre-readers tend to confuse syllables and words. There is also much evidence to suggest that pre-readers have very little interest in printed words prior to beginning reading instruction (Evans & Saint-Aubin, 2005), although they do have limited knowledge of their own names (Treiman, Cohen, Mulqueeny, Kessler, & Schechtman, 2007). Clearly an early part of the process of learning to read must involve the acquisition of the concept that words are units of meaning in language.

One of the most robust predictors of reading achievement is phonological awareness, which has been shown to be critical to successful reading acquisition. Phonological awareness describes one's own awareness of, and access to, the phonology of one's language (Wagner & Torgesen, 1987) and is demonstrated by performance on various tasks such as tapping out the number of sounds in a word, putting sounds together to form a word, or segmenting words into phonemes. Bradley and Bryant (1983) found that phonological awareness at 4 years predicted reading and spelling achievement at 8 years, even when IQ, memory and social class were controlled (see also Adams, 1990; Share, 1995), and indeed it is now widely accepted that dyslexic readers have impaired phonological awareness (Snowling, 2000; Stanovich & Siegel, 1994). However, it should be noted that the relationship between phonological awareness and reading is reciprocal and increased competence in reading leads, in turn, to improved phonological skills, both in children (Wagner & Torgesen, 1987), and in adults (Morais, Cary, Alegria, & Bertelson, 1979).

2.2.2 Stages of reading development

Having considered some of the prerequisites of reading acquisition, the process of reading development itself will now be discussed. There is some general agreement over the processes that occur as a child begins to learn to read, although it is now widely accepted that strict stage theories of reading development, popular in the 1980s, cannot encompass the variability and complexities involved in reading acquisition (Snowling, 2000). While there have been many different models put forward to explain and describe the processes involved in reading acquisition, Frith's model (1985) will be the one outlined here as it is straightforward, and in addition, typifies many of the other models (Chall, 1983; Ehri, 2002; Ehri & Wilce, 1987; Marsh, Friedman, Welch, & Desberg, 1981; Mason, 1980). In general, there is substantial agreement among alternative models of reading development as to what constitute the distinct stages involved in learning to read, and each phase is characterised by the predominant type of connection that connects written words with their other identities (e.g. visual/phonological) in memory (Ehri, 2005).

Very generally, it is thought that there are three broad (overlapping) stages of reading, starting with the logographic stage whereby beginning readers use their visual skills to read words globally by partial cues and by accessing semantic memory, much in the same way as pictures are processed (Ehri, 1987). Associations between visual aspects of a letter string and the child's lexical representation of the word are purely arbitrary (e.g. *dog* has a 'tail' at the end) and unconnected to the sound of the word. While this stage is functional for as long as only a few words are needed, there soon comes a point at which confusion between words is apparent, as the emphasis is purely on graphemic cues.

Children then pass into an alphabetic stage, during which alphabetic skills are transferred back to reading which allows the child to read novel words by using orthographic rule-based cues. This stage can only begin once the child knows the shapes, names or sounds of the letters of the alphabet and signals the beginnings of letter-to-sound mapping (although usually only word-initial and word-final letters: Rayner, 1976; Rayner & Hagelberg, 1975). This stage differs from the first stage in that there is a systematic, rather than an arbitrary, association between spellings and pronunciations. The limitation of this stage is that children are not yet able to fully decode novel (irregular) words, but rather must rely on their memories to access words which have been previously encountered.

Finally as reading skills become more automatic, children begin to rely on orthographic relationships such as morphological spelling patterns which go beyond grapheme-phoneme correspondences (Snowling & Frith, 1981), and to use analogies to read unknown words (Goswami, 1986). This orthographic stage of reading requires a more complete knowledge of the orthographic system whereby children are able to segment words into phonemes which crucially enables them to decode unfamiliar words. This stage requires instruction, in addition to extensive experience of learning letter-to-sound correspondences, and usually takes place between the ages of eight and ten. Note that the children that took part in the experiments in this thesis were beyond the first two of Frith's stages and were probably able to decode novel words fairly well. It is likely then, that the majority of children were in Frith's orthographic stage and that their reading had become relatively automatic. This was important as it was children's online comprehension processes which were of interest rather than their ability to decode individual words. A measure of reading ability was administered to each child in order to assess their reading skill, and no child whose reading ability was below the normal range took part in the experiments.

While stage models have been helpful in that they have provided a good descriptive account of reading development, they have little explanatory power and do not supply mechanisms of change that trigger movement from one stage to the next. More recently, there has been a move to viewing reading development as continuous rather than stage-like (e.g. Nation, in press), and in addition, research has focussed on the role of certain skills (such as phonological and orthographic skills) in learning to read, as well as predictors and correlates of reading ability. The question has become not so much how reading progresses, but rather what skills/abilities are necessary for successful reading acquisition to take place.

The correlation between non-word reading (a measure of phonological ability) and *irregular* word reading is much weaker than that between non-word reading and *regular* word reading (Baron & Treiman, 1980), indicating that the transition to skilled word reading may be modulated, at least in part, by other factors (Castles & Nation, 2006). Likewise, while phonological factors provide an essential substrate to decoding, other aspects of oral language such as vocabulary and listening comprehension are important for reading comprehension (Muter, Hulme, Snowling, &

Stevenson, 2004; Oakhill, Cain, & Bryant, 2003), and oral language development is a reliable predictor of word recognition abilities during reading (Nation & Snowling, 2004).

Chall (1967) showed that the best predictor of early reading ability was pre-readers' knowledge of letter names. Nation and Snowling (2004) found that children's oral language proficiency influenced reading development. Castles and Nation (2006) cite orthographic processing skills, exposure to print, and semantic knowledge as possible contributors to skilled word reading, in addition to phonological skills, and Nation and Cocksey (2008) showed that semantic information is activated during visual word recognition in children as young as seven. It is clear then, that there are many more factors than phonological skills that contribute to successful reading acquisition, and these factors should not be underestimated in favour of focussing heavily on decoding skills. What we do not yet know is how these factors and others constrain online reading behaviour in children, and how they change over time and contribute to overall reading development. Although no PhD thesis could hope to address all these issues, the present work will build on what is already known in order to provide a picture of how semantic, syntactic and lexical, and pre-lexical information constrains children's eye movement behaviour during text reading, and to provide direct comparisons between adults' and children's processing of these linguistic phenomena.

2.3 Children's online language processing

There have only been a handful of studies that have used eye movements to investigate children's online language processing, presumably because until recently the eye tracking equipment available was not conducive to use with developmental populations. Buswell (1922) and Taylor, Frackenpohl and Pettee (1960) conducted some early studies into children's eye movements during reading, which were later built on by Rayner (1986) and McConkie et al. (1991). These studies emphasized children's general oculomotor behaviour during reading and did not manipulate the content of the text used. In addition, there have been a small number of studies investigating specific linguistic aspects of online language comprehension in children (Aghababian & Nazir, 2000; Arnold, 2008; Arnold, Brown-Schmidt, & Trueswell, 2007; Bijeljac-Babic, Millogo, Farioli, & Grainger, 2004; Felser, Marinis, & Clahsen, 2003; Häikiö et al., 2008; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Meroni & Crain, 2003; Nation, Marshall, & Altmann, 2003; Snedeker, Thorpe, & Trueswell, 2001; Traxler, 2002; Trueswell, Sekerina, Hill, & Logrip, 1999), two of which have used eye movements in natural reading (Blythe et al., 2006; Hyönä & Olson, 1995). Section 2.4.1 will review the four studies which have examined children's oculomotor behaviour, and Section 2.4.2 will review only two studies in detail (Aghababian & Nazir, 2000; Hyönä & Olson, 1995) which are particularly relevant to the current work (for a review of the other studies, see Joseph et al., 2008).

2.3.1 Children' oculomotor behaviour during reading

Buswell (1922) monitored the eye movements of 186 children aged 6 to 18 years as well as a group of adults, as they read written texts. He deliberately excluded particularly good or poor readers and used the same stimuli for all participants, with the aim of excluding the material as a confounding source of variance, although this meant that the stimuli were not age-appropriate in terms of difficulty. Buswell found that the most significant development in reading skill occurred up to the age of nine, after which improvement began to level off. The mean number of fixations and regressions made per line of text, as well as the mean fixation duration, decreased with age, with the steepest curves from ages six to nine. Interestingly, unlike number of, and duration of, fixations which flattened out at age nine, the number of regressions per line continued to change across all ages. However, given that the same reading materials were used for or all ages, it is not clear whether these developmental changes were due to oculomotor differences between older and younger children per se, or whether they would have disappeared had the text been age-appropriate for all children. In a sub-group of participants aged 8-12 years, Buswell also found that the mean number of fixations and regressions per line, as well as the duration of fixations, decreased as comprehension of the texts increased. That is to say, the more difficult the children found the texts to understand, the longer and more often they looked at them. This finding suggests that for children as well as adults, the linguistic properties of a text and the difficulty in cognitive processing that the text produces have a profound influence on reading behaviour.

Taylor (1965) reported data from a large study in which photographic recordings were used to record the eye movements of 12,143 participants from 6-years-old to college students (Taylor, Frackenphol & Pettee, 1960). It is important to note that Taylor did not provide details of the reading materials used by participants and therefore it is impossible to know whether the texts were age-appropriate or not. Nevertheless, in line with the Buswell study, it was observed that as age increases, the mean number of fixations, number of regressions, and fixation durations decrease. In addition it was found that both the mean recognition span (i.e. the area of text that is visually available to the reader and from which s/he can recognise words and letters) and reading rate increased with age. These data therefore clearly corroborate those from the Buswell study and show children's eye movements during reading show a clear developmental trajectory. However, it is not clear from these data whether the developmental differences observed were due to improvements in language processing or basic oculomotor control. Indeed, non-reading research has shown that although children have longer saccadic latencies, the accuracy and velocity of saccades do not vary systematically with age (Cohen & Ross, 1978; Fukushima, Hatta, & Fukushima, 2000; Groll & Ross, 1982).

Rayner (1986) and McConkie et al. (1991) continued with this line of research some decades later. While Rayner compared the eye movements of 7-, 9- and 11-year-old children with those of adults, McConkie et al. examined children's eye movements only, but over a slightly different age range (ages 6-10, and testing all ages between these two limits). Overall, both studies showed that beginning readers, as compared to older readers and adults, exhibited longer fixation durations and shorter saccades. While an adult's average fixation duration is between 200 and 250ms (Rayner, 1998), children beginning the reading process exhibited fixation durations of more than 300ms (Rayner, 1986). Likewise, saccades which are about 8 characters for adult readers, were between 3.6 and 6.3 characters in children between ages six and ten (McConkie et al., 1991). Blythe et al. (2006) also found that children aged 7-11 years made longer fixations and shorter saccades as compared to adult readers.

There were also some inconsistencies between the two studies: while Rayner found that beginning readers made more frequent regressions than their more highly-skilled counterparts (as did Buswell, 1922; Taylor, 1965), McConkie et al. found that the frequency of regressions remained fairly constant between the ages of six and ten although adults made fewer regressions as compared to children. A second inconsistency between the two studies was that while Rayner found that the frequency of (forward) fixations increased with age, McConkie et al. did not find this to be the case. However, these two discrepancies can be attributed to the difficulty of the text. In the Rayner study, like the Buswell study (1922), all participants read the same text which made it relatively easier for the older readers. However, in the McConkie et al. study, texts were age-appropriate and this almost certainly accounted for the stability in the number of fixations made (both progressive and regressive) across age groups. These results show very clearly that eye movements reflect the processing difficulty of the reader, as argued by proponents of cognitive models of eye movements during reading, such as E-Z Reader.

McConkie et al. also found that beginning readers were more likely to refixate words as compared to more highly skilled readers. While adults refixated 5-letter words on 15% of occasions, 6-year-old children refixated 5-letter words 57% of the time. In addition, the probability of skipping a word during reading increased developmentally. These findings fit well with the results already reported that children make shorter saccades and more frequent fixations (when text is not controlled for age-appropriateness). All the experiments reported in this thesis will use identical materials for both child and adult participants in order to make direct comparisons between groups. It is therefore expected that children will make more regressive and progressive saccades as compared to adults. However, what is important in the experiments reported here is whether there are qualitative differences in the eye movement patterns between adults and children in response to the psycholinguistic manipulations employed. Importantly, this is where the current work deviates from previous research outlined in this section: while previous studies have been informative as to the basic oculomotor processes occurring in child and adult readers, the experiments conducted as part of this thesis will go further to examine precisely how oculomotor behaviour is influenced by different linguistic phenomena. To this extent, the effects observed in the series of experiments reported in this thesis will be diagnostic of language processing in children.

In addition to taking eye movement measures from his participants, Rayner also used the moving window technique (described in Section 1.6) to examine the size of beginning readers' perceptual spans. Perceptual span can be seen as somewhat similar to the 'recognition span' reported in the Taylor (1965) study but the procedure used by Rayner was much more accurate. He found that, in contrast to adults who have perceptual spans of 3-4 characters to the left of fixation, and 14-15 characters to the right (McConkie & Rayner, 1976; Rayner & Bertera, 1979; Rayner, Well et al., 1980), children's perceptual spans varied from 11 characters (7 year-old children) to 14 characters (11 years old) to the right of fixation, although it is probable that it is only word length rather than letter identity information which is available further to the right of fixation, as is the case with adult readers. Younger readers were less disrupted by having only a small window of letters available to them, suggesting that beginning readers focus more of their attention on foveal word processing during a fixation as compared to more skilled readers. More recent research (Häikiö et al., 2008) has shown that letter identity span (the number of letters that can be identified during one fixation) is much smaller than the perceptual span (5-7 characters for 8-year-old readers, 7 characters for 10 year-old readers, and 9 characters for 12-year-old and adult readers), but also increases developmentally, independent of text difficulty.

To summarise what is already known about children's oculomotor behaviour during reading; while it has been documented that children make shorter saccades and longer fixations than adults, and that these changes take place gradually with age, it is unlikely that these changes are due to oculomotor development per se, but rather reflect cognitive changes taking place as reading skill develops. We have seen that the difference in the number of fixations, both forward and regressive, between beginning and more skilled readers disappears when text is age-appropriate, and this strongly suggests that processing difficulty drives children's eye movements, as it does adults'. As children gain more experience of reading and become more fluent both in decoding words and in sentence comprehension, their eye movements come to reflect this relative ease of processing that they experience. There is not a qualitative, but a quantitative difference between adults' and children's perceptual spans and this again would appear to reflect the ease with which children are able to process text that is available foveally.

2.3.2 Children's online language comprehension

There is also a small but significant literature which has used online methods to investigate specific aspects of children's language comprehension during reading, most notably word length effects and word frequency effects. In this section, only two studies will be described in detail as other studies have either used methodologies which may not generalise to natural reading (e.g. Felser et al., 2003; Nation et al., 2003; Traxler, 2002; Trueswell et al., 1999), or else emphasized non-linguistic aspects of sentence reading (Blythe et al., 2006).

Aghababian and Nazir (2000) used a single-word paradigm to determine what beginning readers can extract from a word during a single fixation. This is an interesting question given the evidence that beginning readers tend to make multiple fixations within a single word where skilled readers make just a single fixation (Rayner, 1986; McConkie et al., 1991). They found a significant word-length effect (i.e. children were more accurate at reporting short as compared to long words that had been presented to them) and this effect slowly diminished with age. Bijeljac-Babic, Millogo, Farioli and Grainger (2004) found similar results in both a naming and luminance increment (word recognition) task. Word length effects (in terms of number of letters, but also in number of syllables and number of phonemes) decreased with age from age eight to adulthood. Bijeljac-Babic et al. suggested that children may process letters of a word serially rather than in parallel; that is to say, children read sequentially from the left to the right of a word, whereas adults (and older children) are able to processes more than one letter in parallel and can therefore identify a word more efficiently in a single fixation. Aghababian and Nazir argued that reading experience mainly reduces the time needed to extract visual information from print, rather than there being qualitative differences between beginning and skilled readers. However, whether these results can be generalised to normal text reading, and whether the effects remain when multiple fixations on a word are permitted and later measures of processing are taken, is an empirical question which remains to be investigated.

Hyönä and Olson (1995) recorded the eye movements of both dyslexic children (mean age = 14.4 years) and reading-age-matched controls (mean age = 10.5 years) as they read aloud texts which contained words which were categorised as high, medium or low frequency, for each of three word lengths (short, medium, long). They found a strong effect of word length in both groups which was apparent in gaze duration (the sum of all first pass fixations), number of first pass fixations, number of regressions to the word, and total reading time. These results reflect the increased within-word refixation probabilities on long words as compared to short words, which is the pattern also seen in adult readers (Vitu et al., 1990). However, an oral reading task might bring about a more robust length effect than silent reading, especially if the participant fixates a word until its pronunciation is completed, as beginning readers tend to do (Rayner & Pollatsek, 1989). In any case, it is known that eye movements differ somewhat for reading silently versus reading aloud (Rayner, 1998).

Hyönä and Olson also found a strong effect of word frequency in both groups of readers observed in first fixation durations (the duration of the first fixation made on a word), showing that the frequency of a word has an immediate effect on processing in children, as it is known to do in adults. Importantly, the word frequencies of the target words were drawn from age-appropriate texts so that the high and low frequency words were high or low frequency for child readers, rather than for adults. While this important study suggests that the frequency of occurrence of a word has an immediate effect on the reading behaviour of children as well as adults, a more recent study by Blythe et al. (2006) failed to find a frequency effect in children aged seven to eleven years, despite finding a strong frequency effect in their adult participants in the same study (although note that this study used adult corpus data to index word frequency). Furthermore, Hyönä and Olson did not include an adult group in their study as they were primarily interested in reading performance of the dyslexic group. Experiments 1-3 (reported in Chapters 3 and 4) directly compare adults' and children's online processing of long versus short, and high frequency versus low frequency, words during natural silent reading and focus on fine grain oculomotor behaviour, in order to investigate word frequency effects in children further.

2.4 Summary and thesis outline

Having reviewed the few studies investigating children's online reading behaviour, it can be concluded that there is a significant gap in the eye-movement literature. If, as postulated, the differences observed between adults and children in their oculomotor behaviour are due to the cognitive demands of the written stimulus, then it is a valid and unanswered question as to which aspects of linguistic processing cause children relatively greater processing load as compared to adults. It may be that the differences emerge only in higher linguistic processes such as semantics, or it may be that they emerge at a lexical or pre-lexical level. Now that previous research has generated important preliminary data regarding children's basic oculomotor behaviour, what is needed is to introduce specific linguistic manipulations in order to systematically uncover differences (or similarities) between adults and children in pre-lexical, lexical, syntactic, and semantic processing during text reading. That is the aim of this PhD thesis.

The thesis will describe a series of experiments all of which employ linguistic manipulations in order to examine children's as compared to adults' online processing of a particular aspect of language, starting from the lowest level (pre-lexical) and working up towards the highest level (semantic). Chapter 3 will describe an experiment (Experiment 1) investigating basic word length and word frequency effects in adult and child readers. As mentioned, these effects are extremely robust in adults so it is important to establish whether children exhibit these same strong lexical and pre-lexical effects during reading. In addition, landing position effects are examined in order to examine any disparities in where children and adult fixate a word as this could have consequences for any differences observed in how words are recognised. Chapter 4 will further address the issue of word frequency and describe two experiments in which the frequency of target words was indexed according to both adult (Experiment 2) and child (Experiment 3) corpora. Chapter 5 will examine lexical ambiguity effects in children and adults (Experiment 4), to investigate whether children, like adults, look longer at words with two meanings. Chapter 6 will move on from lexical-level effects to report two experiments (Experiments 5 and 6) in which syntactic structure was manipulated. This will provide an opportunity to establish whether children exhibit garden-path type effects in the same way that adults do. The final experiment in Chapter 7 (Experiment 7) will investigate plausibility manipulations during thematic role assignment in reading, and this will be the only aspect of semantic interpretation examined in the thesis. Finally, Chapter 8 will discuss the theoretical and methodological conclusions that can be drawn from this series of experiments.

Chapter 3 : Visual and lexical processing

3.1 Introduction

The first experiment of the thesis is an examination of visual and lexical level effects during reading. It is important to know whether differences exist between adults and children at this relatively low level of language processing before examining higher levels of processing involving syntactic and semantic manipulations, as any differences observed at a lexical level must be taken into account when looking for differences at a higher level. Experiment 1 involves the examination of three different aspects of processing; (1) landing positions, (2) word length effects, and (3) word frequency effects. The chapter will be structured in the following way. The introduction will comprise of a brief review of the literature pertaining to landing position effects (Section 3.1.1), length effects (Section 3.1.2) and word frequency effects (Section 3.1.3). This will be followed by a description of the methods used (Section 3.2), which will be broadly the same for all experiments in the thesis, the results of the experimental analyses (Section 3.3), and finally a discussion of the findings in terms of understanding children's processing of words (Section 3.4).

3.1.1 Landing position effects

The landing position on a word is the location that the eyes initially fixate after making a saccade onto that word. As noted in Chapter 1, Rayner (1979) first labelled the position within a word where readers typically make their initial fixation during text reading as the *Preferred Viewing Location* (PVL). The PVL is a little to the left of the word centre, (McConkie et al., 1988). McConkie et al. (1988) conducted extensive analyses in which they examined the initial locations of adults' fixations, on words which were 3-8 characters in length. They found that the PVL was actually a composite distribution of many landing site distributions, each contingent on the site from which the saccades were launched (see also Rayner et al., 1996). In this way, word length did not influence landing position per se, although landing positions have been shown to shift to the right as word length increases thereby increasing saccade length (Rayner et al., 1996), and resulting in slightly different distributions for different word lengths. McConkie et al. argued that their findings showed that the oculomotor system is not simply avoiding blank spaces (if this were the case, variation would be much greater for long words), but rather the visual pattern of a word provides the oculomotor system with a target or 'centre of gravity' which is more constrained than the space that the word occupies.

O'Regan and Lévy-Schoen (1987) subsequently made a distinction between the PVL and the *Optimal Viewing Position* (OVP) which is operationally defined as the location within a word at which recognition time is minimised. The OVP is located close to the word centre, a little to the right of the PVL. O'Regan and Lévy-Schoen investigated OVP effects in isolated word reading and found that the probability of making a refixation was reduced, and gaze durations were shorter, when the OVP was fixated. Indeed, in natural reading, refixations have been shown to be more likely when initial fixations land at the beginning or end of a word than in the middle of a word (McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Rayner et al., 1996; Vitu & O'Regan, 1995; Vitu et al., 1990), and this leads to longer fixation durations at the word centre than at the beginning or end of the word in sentence reading (Vitu, McConkie, Kerr, & O'Regan, 2001), in contrast to the inflated gaze durations at word centres observed on words presented in isolation.

The location of intra-word refixations has received less attention than the location of initial fixations in the literature although refixation location (at least on isolated letter strings) has been shown to be dependent on word length as well as initial landing position (Beauvillain, Dukic, & Vergilino, 1999; McConkie et al., 1989; O'Regan & Lévy-Schoen, 1987; Vergilino & Beauvillain, 2000) However, an interesting question is whether the refixation saccade is pre-programmed at the time the initial saccade is programmed (based mostly on word length information available at that time), or whether it is dependent on the location of the first fixation. Vergilino and Beauvillain (2000, see also Beauvillain et al., 1999; Vergilino-Perez et al., 2004) compared (progressive) refixations on letter strings which increased (9-11 letters) or decreased (11-9 letters) in length at various time intervals (0, 50, 140, and 220 ms into the fixation) following an initial saccade to the string. They found that the refixation saccade was based on initial length information presented in the parafovea (and available when programming the initial saccade to a word), rather than updated length information available once the initial fixation had been made, although modification of the refixation program was possible provided the new length information was available sufficiently early. However, recent estimates indicate that only about 20% of refixations in reading are pre-planned (McDonald & Vergilino-Perez, 2006) and so a large proportion of refixations are likely to be programmed after the word has been fixated, perhaps due to processing difficulty, visual factors such as word length, or corrective saccades following oculomotor error.

In contrast to the pre-planned refixation account given by Vergilino and colleagues, O'Regan and colleagues (O'Regan, 1990; O'Regan & Lévy-Schoen, 1987) argue that refixation location is determined by the location of the initial fixation as the eye attempts to "spread its fixations evenly over the word" (O'Regan, 1990, p. 427) and bring unidentified letters into foveal vision. In this way, if the initial fixation is close to the beginning of the word, then the refixation will be targeted towards the end of the word, and if the initial fixation is close to the end of the word, the refixation will be located closer to the beginning of the word. Importantly, this pattern of effects has been observed during normal text reading (McDonald & Shillcock, 2004; Rayner et al., 1996), as well as with isolated words (O'Regan & Lévy-Schoen, 1987). Refixation probability is also known to be modulated by linguistic processing factors such as word frequency (McConkie et al., 1989; Rayner et al., 1996) and predictability (Balota, Pollatsek, & Rayner, 1985), as well as visual factors. Other findings suggest that the location of refixations within words can be influenced by the specific orthographic characteristics of the word (White & Liversedge, 2004; 2006a; 2006b). Overall, refixations are much less understood than other aspects of eye movement behaviour during reading, and therefore merit further research. Following from previous research, it was anticipated that initial fixations at the beginning of a word would result in an increased number of refixations at the word end, and that initial fixations at the word end might also result in an increased number of refixations located at the beginning of a word. It was also expected, at least for the adult readers, that there would be very few refixations following an initial fixation at the word centre in line with previous findings.

While there has been much less research investigating saccade targeting and landing position effects in child readers than in adults, there are three studies in particular which have addressed these issues. Aghababian & Nazir {, 2000 #193, see Chapter 2, Section 2.3.2 for a full description of this study} investigated whether children's (aged 6-10 years) performance in a word recognition task would be affected by their viewing position within a word. They used a paradigm in which isolated words were briefly presented (presentation time was calculated on an individual basis to prevent children making a saccade) after two fixation dots, and the position of the word centre was laterally displaced in relation to these dots so that children fixated different locations within the target words. They found that children were much better at recognising words when they fixated the word centre as compared to the word boundaries, for different word lengths, and in all age groups tested. Furthermore, this basic pattern of performance did not change with age. These findings show that for children as young as six years old, as well as adults, word recognition is facilitated when the word centre is fixated.

However, there are clearly some important differences between the paradigm employed by Aghababian & Nazir and using eye movements during normal text reading to examine landing position effects. First, what children can extract from a single fixation on a word will not necessarily reflect what they are able to extract from a word when they are free to move their eyes to fixate and refixate a word freely. Second, reading a word in the context of a sentence is very different to reading it in isolation. Third, in a sentence reading experiment, it is not necessary to ask participants to complete an additional task such as reporting the word presented as the eye movement data are sufficient. Fourth, although children were instructed to fixate fixation dots on every trial, without measuring their eye movements, it is difficult to establish whether the children were in fact fixating the word centre, and indeed research has shown that even adult participants do not reliably fixate a fixation cue when instructed to do so (Jordan, Patching, & Milner, 1998). Fifth, Aghababian & Nazir divided their target words into five equal zones irrespective of word length, rather than by the more psycholinguistically meaningful measure of characters. Finally, landing position effects were indexed by the accuracy of word recognition, and not speed of recognition which will be measured in the experiment set out in this chapter.

A second study that examined within-word eye movement behaviour in children but using a more natural reading task was conducted by Vitu, McConkie, Kerr and O'Regan (2001). Although the main purpose of this study was not to compare adults and children, the data for adults and children were presented alongside each other so it is possible to draw comparisons. It should also be noted that the children in this study were 12 years old: an age at which previous studies (McConkie et al., 1991; Rayner, 1986) suggest that most sentence-level characteristics of eye movement behaviour are close to, or have reached, adult levels. The children did not appear to differ from the adults in the locations of the first or second fixations, although Vitu et al. did not test this specifically. These data suggest that children do not differ from adults in terms of where they target their saccades during reading, and further suggest that children are as adept as adults in targeting their saccades accurately towards their intended destination (see also Cohen & Ross, 1978). Finally, Vitu et al.'s data show that adults and children do not appear to differ in their refixation probabilities as a function of where they initially fixate a word. However, given that Vitu et al. did not conduct a comparative study of adults' and children's landing positions, and that the children were at the upper end of the age range in which (most) eye movement behaviour during reading might be thought to develop, further investigation of this issue is required.

The third study to investigate landing position effects in children was the study outlined in Chapter 2 (Section 2.3.2) by McConkie et al. (1991). McConkie et al. reported data from Grimes (Grimes, 1989), which showed that during their first year of reading instruction, children showed the same pattern of landing positions as adults, although full analyses were not given. That is, children as young as seven years old targeted their saccades towards the word centre during normal text reading. Furthermore, McConkie et al. found that, like adults, children were more likely to refixate a (five-letter) word following an initial fixation on the space before, or the first letter of, a word than if the first fixation was close to the word centre. However, importantly, these inferences were made from observing trends in the data rather than from formal statistical analyses. Experiment 1 will build on these preliminary findings to examine in more detail where children locate their initial fixations on words of different lengths, and how their refixation behaviour is influenced by their initial landing position.

Finally in relation to landing position effects in children, it is interesting to note that the empirical evidence from the three studies outlined above suggesting that children target their saccades towards the word centre during reading is supported by a modelling study in which Reichle and Laurent (2006) used reinforcement learning to allow an artificial agent to learn to move its eyes as efficiently as possible during reading. The resulting pattern of eye movements clearly and quickly (after as few as ten learning iterations) resembled those of skilled readers, further supporting the claim that the distribution of initial fixation locations emerges early in the reading acquisition process.

From the literature examining where children and adults fixate during reading, it seems clear then that (for words up to eight characters in length) adult readers target their initial fixations on a word towards the word centre and that they do so in order to maximise efficient word recognition processes. The small literature investigating this issue in children suggests that this is also the case in children, and that the tendency to fixate close to the word centre may arise very early in the reading acquisition process (as early as the first or second year). Furthermore, it is also the case that decisions regarding where to move the eyes during reading are made independently of lexical variables in adult readers (Rayner et al., 1996), and instead depend on the location from where the saccade was launched (McConkie et al., 1988; McConkie et al., 1989; O'Regan, 1990), as well as the length of the targeted word (e.g. Vergilino-Perez, Collins, & Dore-Mazars, 2004).

3.1.2 Word length effects

The length of a word is known to affect the length and frequency of fixations that adults make on that word during reading. Whereas words of four or five letters generally only require a single fixation for skilled readers (as a fixation close to the word centre will bring all the letters into foveal view so that it can be processed easily and efficiently), a longer word of nine or ten letters may require more than one fixation in order for the reader to obtain the visual information necessary to begin word recognition processes. Word length effects are among the most robust in the adult eye movement literature and the length of a word directly affects the time taken to lexically identify it. It is well-documented in the eye-movement literature that adult readers are more likely to fixate (Rayner & McConkie, 1976) and refixate (Vitu et al., 1990) a long as compared to a short word and also that fixation durations on long words are significantly longer than those on shorter words, both when only a single fixation is made (Rayner et al., 1996), and when all first pass fixations (i.e. all fixations made as the reader moves his eyes from left to right, but excluding regressive fixations made to re-read the word) are considered (Just & Carpenter, 1980). Readers also skip short words more frequently than long words (Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner & McConkie, 1976), and denial of word length information impedes reading, especially for infrequent words (K. Rayner et al., 1998).

However, word length effects have been less studied in child readers. In the same study outlined above in Section 3.1.1 (and also in Chapter 2; Section 2.4.2), Aghababian and Nazir (2000) found a significant word-length effect which slowly diminished with age (see also Bijeljac-Babic et al., 2004 for similar results), although these results may not generalise to normal text reading. Hyönä and Olson (1995) recorded the eye movements of both dyslexic readers and reading-age-matched controls as they read aloud texts which contained words which were manipulated for length (see Chapter 2; Section 2.4.2 for a more detailed review of this study). They found a strong effect of word length in both groups. However, Hyönä and Olson did not include an adult control group as they were primarily interested in reading performance in the dyslexic group. Furthermore, the effects observed in oral reading may not generalise to the domain of silent reading.

To summarise what is known about word length and visual word recognition, it is clear that adults' eye fixations are influenced by the length of a word, and that the longer a word is, the more likely an adult reader is to fixate it and refixate it, resulting in longer gaze durations. While there have been no studies to date which have explicitly examined word length effects in adult and child readers in a way which has enabled direct comparisons of their eye movement records, the limited evidence available suggests that children's eye fixations will also be influenced by word length and indeed that word length effects may be even more pronounced in child readers.

3.1.3 Word frequency effects

Word frequency is indexed by how often a word is encountered per million words. This information is available through corpora, such as CELEX (Baayen, Piepenbrock, & Gulikers, 1995) and Kučera Francis (Kučera & Francis, 1967), which document how frequently a given word appears in a range of written texts. Previous research has shown that words which are low in frequency are fixated for longer in reading than high frequency words, both on the first fixation and on subsequent first pass fixations (e.g. Henderson & Ferreira, 1990; Inhoff, 1984; Just & Carpenter, 1980; Rayner & Duffy, 1986; Rayner & Raney, 1996). This is presumably because encountering a word more often increases the ease with which it is identified, and therefore requires fewer processing resources. Indeed, word frequency is incorporated into most models of word recognition (e.g. Forster, 1976; McClelland & Rumelhart, 1981; Paap et al., 1982). Like short words, high frequency words are also skipped more often than low frequency words when they are six letters or less in length (O'Regan, 1979: Rayner et al., 1996), and there is sometimes a spillover effect (that is, the fixation subsequent to fixating a low frequency word is inflated compared to that following a fixation on a high frequency word) from fixating low frequency words (Rayner & Duffy, 1986). Frequency effects are found in lexical decision and naming tasks (Schilling, Rayner, & Chumbley, 1998), as well as in reading.

Seidenberg (1985) found that skilled adult readers exhibited a smaller frequency effect than less skilled adult readers. This is an interesting finding in relation to this thesis as it may suggest that children will exhibit even larger frequency effects than less skilled adult readers especially those closer to the beginning of the reading acquisition process. Waters, Seidenberg & Bruck (1984) suggested that high frequency words are recognised visually whereas low frequency words are sounded out (see also Jared et al., 1999). This difference may be even more apparent in adults as compared to children, although it may be that children, in particular young children, sound out both high and low frequency words.

In the same experiment outlined in Section 3.1.2 (and Chapter 2; Section 2.4.2), Hyönä & Olson (1995) found strong effects of word frequency in both dyslexic and typically developing children, which were observed in first fixation durations. While this study suggests that word frequency has an immediate effect on reading behaviour in children as well as in adults, the same limitations of the study mentioned in Section 3.1.2 (no adult participant group, oral rather than silent reading task) apply here as well. It is worth noting at this point that in the studies outlined in this section, together with the three experiments which examine frequency effects in this thesis, word frequencies are taken from written language rather than spoken language, The relationship between these two different kinds of frequencies, and how they impact on children's online language processing is not yet known and, although beyond the scope of this thesis, would certainly be an interesting issue to investigate in future research.

A more recent eye movement study failed to find frequency effects (in first fixation or gaze duration) in children aged 7-11 years during normal sentence reading (Blythe et al., 2006), despite finding reliable frequency effects in adults in the same experiment (although the adults read different sentences). The target words in this experiment were controlled a priori for word length which may in part explain why the results conflict with those from the Hyönä and Olson study. Furthermore, the children in the Blythe et al. study were considerably younger than those in the Hyönä and Olson study and read sentences silently rather than aloud. Finally, word frequency was indexed by counts from adult corpora, whereas in the Hyönä and Olson study, counts were taken from age-appropriate texts. There are therefore several factors which could explain the discrepancy in findings and further empirical evidence is required to clarify the issue of whether children do require greater processing resources to read low- as compared to high-frequency words. The third part of the experiment outlined in this chapter will specifically address the question of whether children, like adults, exhibit frequency effects while reading identical sentences.

3.1.4 Predictions

To summarise, there were three main issues that were investigated in this first experiment. The first issue related to where adults and children target their saccades in a word during reading. First, the location of adults' and children's initial fixations on four-, six- and eight-letter words was examined. It was predicted, in line with previous research (e.g. McConkie et al., 1988; McConkie et al., 1991), that both adults and children would fixate the PVL, close to the word centre, but that the distributions would differ slightly for the three different word lengths (i.e. readers would initially fixate further to the right as word length increased). Second, the location of, and the probability of making, refixations in adults and children in relation to initial landing positions were examined. In line with previous research (e.g. McConkie et al., 1989), it was predicted that both adults and children would be more likely to make a refixation following an initial fixation at the beginning or end (as compared to the middle) of a word. It was further anticipated that both adults and children would be more likely to refixate the word end following an initial fixation at the beginning of the word, and likewise more likely to refixate the beginning of a word following an initial fixation at the word end. Finally, it was predicted that any such effects might be more pronounced for children than adults.

The second main issue that was addressed in this experiment concerned word length effects. It is known that children make longer and more frequent fixations overall as compared to adults (Blythe et al., 2006; Rayner, 1986) and that the duration and frequency of fixations decreases with age during childhood (McConkie et al., 1991: see Chapter 2; Section 2.4.1). In addition, adults have had more experience decoding words, as well as having a more established knowledge of printed letters, which means they may be better equipped to identify a word globally rather than fixating each component part of it (Bijeljac-Babic et al., 2004). Following these premises, it could be that children require a more detailed visual sample of a word in order to extract the visual information necessary to initiate lexical processing. Therefore, it was predicted that while both adults and children would exhibit inflated fixation times on long words as compared to short words, the magnitude of length effects would be greater in children than in adults, as longer words would receive relatively more, and longer, fixations as compared to short words in child as compared to adult readers.

The third and final issue investigated in the current experiment was word frequency effects. Like word length effects, word frequency effects are extremely robust in adult readers, but the small number of studies that have investigated these effects in children have yielded inconsistent results (Blythe et al., 2006; Hyönä & Olson, 1995). It is important to establish whether word frequency has an immediate and profound effect on the ease with which children access a word in their lexicon, in the same way it does in adult readers. It was predicted that children, like less skilled adult readers (Seidenberg, 1985), would exhibit stronger frequency effects as compared with adult readers, but that both groups of participants would fixate longer on low, as compared to high, frequency words.

3.2 Method

3.2.1 Participants

Twenty adults and twenty children took part in the experiment. Children were recruited from local primary schools in the Durham area, and were all between seven and eleven years old, with a mean age of ten years and four months. Adult participants were undergraduate students at Durham University. All participants were native English speakers with uncorrected vision and no known reading difficulties. Adults were paid for their participation at a rate of £5 per hour. Children received a small gift in return for taking part. Children completed the Wechsler Objective Reading Dimensions (WORD; Wechsler, 1993) which provided assessments of basic word reading, spelling and reading comprehension. All children scored less than 2 SD above or below the standardised norm for their age (mean = 100, SD = 15) on the reading test, with an average score of 111. This meant that overall children had a mean reading age of 12.2 years.

Note that in this experiment, and in all experiments reported in the thesis, there was a substantial disparity between chronological and reading age of the child participants. This was almost certainly due to the self-selection bias, in that the parents and children who chose to take part in the experiments were generally well-educated, intelligent (mean IQ was 120 for children in Experiment 3) and good at reading. It is likely that these children found the experimental sentences easier to process than children whose reading ages matched their chronological age, and that their eye movements reflected this relative ease of processing. It is not known whether the effects observed would have been qualitatively different had children been selected on the basis of their reading age rather than their chronological age. This is a question for future research (see Chapter 8).

3.2.2 Materials

There were 56 experimental items in total. 28 of which contained a word length manipulation (see sentences 1a and 1b where *medicine* and *salt* are target words), and 28 of which contained a frequency manipulation (see sentences 1c and 1d where *record* and *shield* are target words). For the landing position analyses, data for target words in all 56 experimental items were analysed. For the length analyses only the 28 items which contained a length manipulation were analysed. Of these items, 14 contained four letter words, and 14 contained eight letter words.

(1a) Her brother was really mean and put some medicine in my tea. (long)

(1b) Her brother was really mean and put some salt in my tea. (short)

(1c) Yesterday morning I found an old broken record in our shed. (high frequency)

(1d) Yesterday morning I found an old broken shield in our shed. (low frequency)

All sentences were between 50 and 60 characters long. Target words were presented either towards the beginning (50% of sentences) or towards the end of the experimental sentences. All target words were at least 10 characters from the start/end of the sentence. The mean frequency for the target words in the high frequency condition was 151 counts per million (SD = 270) and the mean frequency for the target words in the low frequency condition was 7 counts per million (SD = 3.69). This difference was statistically significant, $t(27) = 2.83$, $p < .01$, and words which were manipulated for frequency were controlled for length: all target words were six characters long. Likewise, there was no difference in word frequency between long and short target words, $t(1, 27) = 1.29$, $p = .21$. The mean frequency for long target words was 48 counts per million (SD = 43) and the mean frequency for short target words was 52 counts per million (SD= 49). The short target words were 4 characters long and the long target words were 8 characters long. Word frequencies were taken from the CELEX English word form corpus (Baayen et al., 1995). In addition there were no significant differences in token or type initial trigram frequency (either position specific or non-position specific) between the eight and four letter (or six letter) words (all $ps > .1$).

Ten adults who did not take part in the main experiment completed a cloze test in which they were asked to complete the experimental sentences with any word which made sense. Only the target word was omitted from the sentence. There were no differences in predictability between long (M = 0.03; SD = 0.07) and short (M= 0.07; SD = 0.16), or high (M = 0.07, SD = 0.19) and low (M = 0.02; SD = 0.05) frequency, target words ($t < 1.2$, $p > .2$). As in all experiments in this thesis, experimental lists were constructed according to a Latin Square such that each list contained a different version of each item and each list contained an equal number of long and short, and high and low frequency target words.

3.2.3 Apparatus

Participants' eye movements were monitored using a Fourward Technologies Dual Purkinje Image eye tracker as they read sentences from a computer monitor at a viewing distance of 100cm. Each character covered 0.19° of horizontal visual angle so that five characters equalled one degree of visual angle. The eye trackers were interfaced with a Pentium 4 computer, with all sentences presented on a 24" monitor. Sentences were presented in white in Courier New font, on a black background. Although participants read binocularly, only data from the movements of one eye were analysed.

3.2.4 Procedure

Participants sat in a customised chair in front of a computer monitor. Head movements were minimized by the use of a bite bar and forehead rests; a restraint was also secured around the back of the head. Participants undertook a calibration procedure during which they looked at each of three horizontal fixation points. Sentences were then presented one at a time. Participants were required to read the sentences normally and then press a button when they had finished reading. The button press terminated the display. In addition to the 56 experimental items, five practice items were also presented at the beginning of the experiment. Participants were asked to respond yes/no to comprehension questions after 19 of the experimental sentences by pressing a button. The experimental session lasted approximately 20 minutes.

3.2.5 Analyses

Custom-designed software was used for the data analyses. Fixations were manually identified and first fixation durations, gaze durations and total word reading times less than 80ms were deleted from the data set. In addition, first fixations more than 1200ms were deleted, as well as outliers (more than 2.5 SDs from the grand mean), and trials during which there was tracker loss or which were not completed due to fatigue (in child participants only). In total these exclusions accounted for 18% of trials. Note that while all participants were extremely cooperative, it is more difficult to run young children than adults in eye-movement experiments, and it is for this reason that a comparatively large proportion of the data was excluded due to fatigue. Participants were never put under any pressure to finish the experiment in its entirety and it was therefore not unusual for children to stop before completing all the trials.

It is worth noting at this point that there was increased variance in the child group than the adult group in all measures of eye movements. While inequality of variance between groups is a violation of the assumptions of Analysis of Variance, this violation is generally viewed as unproblematic (Judd, McClelland, & Culhane, 1995), in particular when sample sizes are the same across participant groups (Field, 2005), as was the case in the experiments reported in this thesis. For this reason, it was not considered necessary to transform the data into z-scores as is often done in cases of heterogeneity of variance.

3.3 Results

All participants scored highly on the comprehension questions. All participants scored 75% or higher; and the mean score was 98% for adults and 92% for children, showing that both group of participants understood the sentences very well. Note that in all experiments reported in this thesis, an effect is termed reliable if the significance value is less than 0.05 in both participants and items analyses; an effect is termed marginal if the significance value is more than 0.05 but less than 0.1, and an effect in termed unreliable if the significance value is more than .1. In cases where one of the two analyses is significant and the other is marginal, this is indicated in the text for the individual cases. In cases where only one of the two analyses is clearly reliable, the effect is referred to as non-significant.

3.3.1 Global measures

Before the main experimental analyses were conducted, some global analyses were first carried out in order to make direct comparisons between the oculomotor behaviour of adults and children during reading independent of the manipulations employed in the experiment. Table 1 shows that children made shorter saccades, longer fixations, more regressions and had longer total sentence reading times than adults, consistent with previous research (Blythe et al., 2006; McConkie et al., 1991; Rayner, 1986). All of the differences between the children and adults were significant (all ps $\leq .01$), except for saccade length in which the difference was reliable by items ($p < .001$) but not by participants ($p = .2$).

Table 3-1: Global reading time measures, mean saccade length and proportion of regressions for adults and children.

	Adults	Children
Saccade length (characters)	8.2 (*5.6*)	7.6 (*6.7*)
Fixation duration (ms)	235 (*104*)	283 (*141*)
Total sentence reading time (ms)	2932 (*1038*)	5381 (*2232*)
Proportion of regressions	0.22 (*0.09*)	0.29 (*0.04*)

3.3.2 Landing positions

For the landing position analyses, data for target words in all 56 experimental items were analysed. Of these items, 14 contained four-letter words, 28 contained 6-letter words, and 14 contained eight-letter words. Initial landing position, the probability of making a refixation as a function of initial landing position, the direction of refixation saccades, the location of refixations as a function of initial landing position, and intra-word saccade length were calculated.

Figure 3-1 shows landing position distributions for adults and children for each of the three word lengths. A 3 (word length: eight-, six- and four-letter words) x 2 (group: adults, children) mixed design ANOVA was conducted. Because the six-letter words were not a within-item variable, and because there were higher levels of word skipping and fewer refixations on four-letter words resulting in fewer data points per cell, only F1 (participant) analyses were carried out for all word lengths. There was a reliable effect of word length on initial landing position: consistent with previous research, landing positions were further to the right on long than short words, $F(1.5, 76) = 49.1$, $p < .001$. There was no effect of group ($ps > .7$). There was, however a reliable interaction between word length and group, $F(1.5, 76) = 5.63$, $p = .01$, allowing separate analyses for each word length and each group to be conducted.

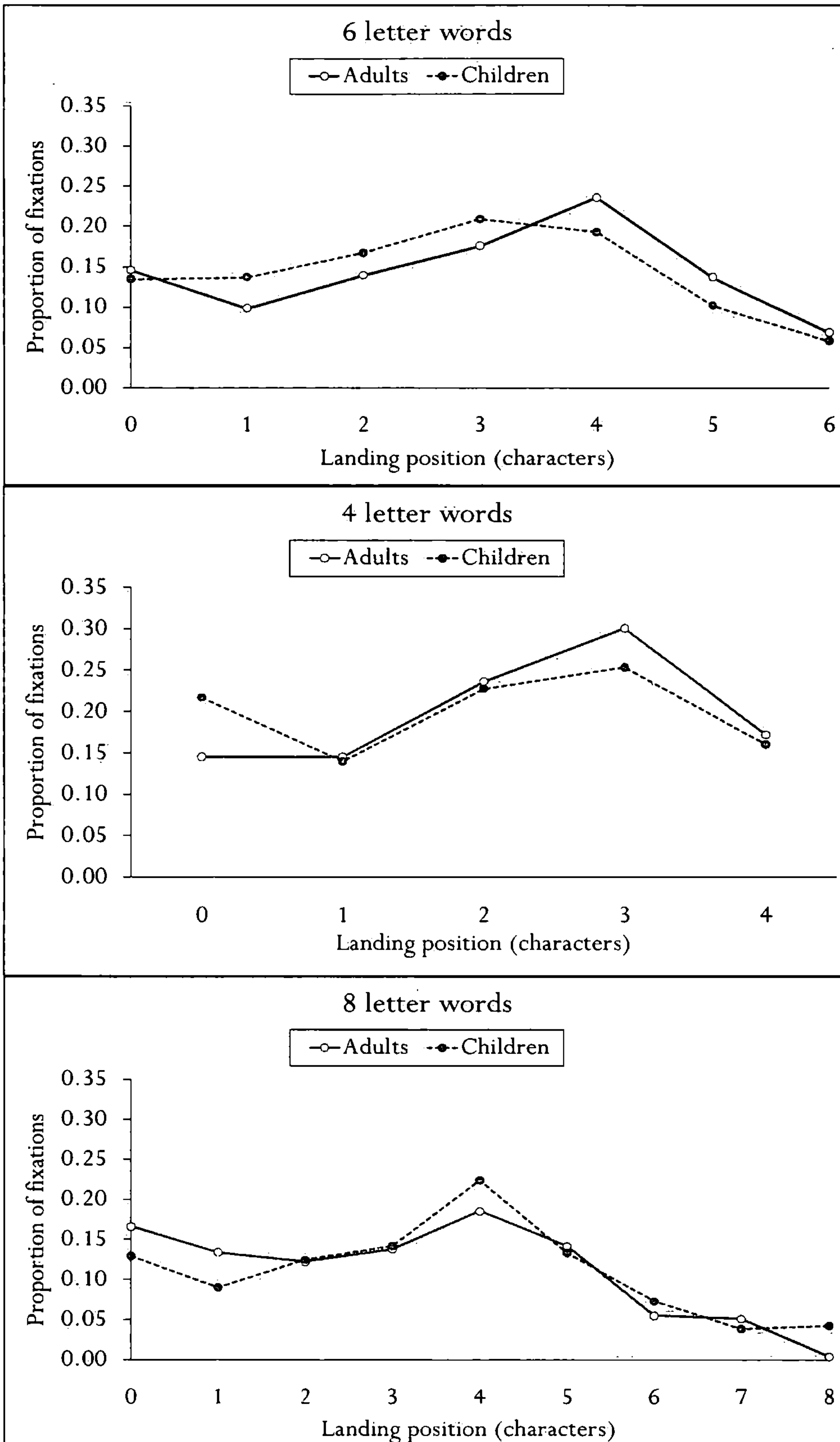

Figure 3-1: Initial landing positions on eight (top panel), six (middle panel), and four (bottom panel) letter words for adults and children.

For eight-letter words, adults had mean landing positions (in characters) reliably nearer the beginning of the word ($M = 2.97$; $SD = 2.08$) than children ($M = 3.39$; $SD = 2.15$); $t(38) = 2.17$, $p < .05$. For six-letter words, adults initially fixated very slightly further into the word ($M = 2.94$; $SD = 1.80$) than children ($M = 2.73$; $SD = 1.73$), but this difference was not reliable, $t(38) = 1.49$, $p > .1$. There was no reliable difference ($p > .07$) in initial landing positions between adults ($M = 2.21$; $SD = 1.30$) and children ($M = 2.00$; $SD = 1.38$) on four-letter words. Furthermore, when analysed separately, both adults and children showed reliable differences in landing positions between eight and six letter words ($ps \leq .005$), and between eight and four letter words ($ps < .001$), but neither group showed a reliable difference in landing position between six and four letter words ($ps \geq .1$).

An Analysis of Covariance (ANCOVA) was then conducted in which launch site (the distance from which the first saccade to the target word was launched) was entered as a covariate. Launch site has been shown to influence initial landing position in adult readers (McConkie et al., 1988). There was a reliable effect of word length on initial landing position: consistent with previous research, landing positions were further into long ($M = 3.17$, $SD = 2.13$) than short ($M = 2.10$, $SD = 1.34$) words, $F(2, 75) = 20.98.1$, $p < .001$. There was no effect of group ($ps > .5$). There was, however a marginal interaction between word length and group, $F(2, 75) = 2.84$, $p = .065$. Pairwise analyses showed no differences in initial landing position between adults and children for eight- or six-letter words ($p > .1$). However, adults (fixated slightly further into four-letter words than children, $F(1, 37) = 4.70$, $p < .05$, perhaps reflecting the overall shorter saccades made by children

There were then no reliable difference between adults and children in the landing positions when the results from both the ANOVA and the ANCOVA are considered.

Overall, the results replicate findings from previous research which has shown that adults make saccades further into a word the longer it is (e.g. McConkie et al., 1988; Rayner, 1979). They also support findings that children, as well as adults target their saccades towards the word centre during reading (McConkie et al., 1991).

The probability of making a refixation as a function of initial landing position was then examined. In order to carry out these analyses (i.e. to have enough data in each condition), each landing position for each word length was categorised as falling at the beginning, middle or end of a word. This was quite straightforward for the eight-letter words as when the space before the word was included; there were nine possible landing positions which fell neatly into three categories. Categorising the landing positions for four and six letter words was more problematic, however, and for this reason, two sets of analyses for each word length were conducted, in which the categorisation of landing positions was different. For six-letter words, in the first set of analyses, the beginning was categorised as letter positions 0 and 1 (where 0 equals the space before the word), the middle as 2, 3 and 4, and the end as 5 and 6; and in the second set of analyses the beginning category encompassed letters 0, 1 and 2, the middle encompassed letters 3 and 4, and the end encompassed letters 5 and 6. For the four-letter words, in the first set of analyses the beginning was categorised as letter position 0, the middle as 1, 2, and 3, and the end as letter 4; and in the second set of analyses the beginning category encompassed letters 0 and 1, the middle encompassed letters 2 and 3, and the end encompassed letter 4. There was no difference in the results for these two sets of analyses and consequently only the first set is reported here.

Table 3-2 shows the refixation probabilities for adults and children for each word length as a function of initial landing position. A 3 (landing position: beginning, middle or end) x 3 (word length: eight, six and four letters) x 2 (group: adults and children) mixed design ANOVA showed a reliable effect of landing position, $F(2, 4) = 8.54$, $p < .001$: readers made more refixations following initial fixations at the beginning as compared to the middle of a word, $t(39) = 4.60$, $p < .001$; and at the end as compared to the middle of the word, $t(39) = 2.34$, $p < .05$, but only marginally more refixations at the beginning compared to the end of a word, t (39) = 1.91, p = .06. This result was in line with predictions, as well as with previous research (Rayner et al., 1996) and is illustrated for each word length in Figure 3-2.

Table 3-2: Refixation probabilities as a function of initial landing position (beginning, middle or end of a word) for eight letter, six letter and four letter words for adults and children

Word length	Initial landing position	Adults	Children
8 letter words	Beginning	0.42	0.56
	Middle	0.06	0.28
	End	0.10	0.49
6 letter words	Beginning	0.31	0.39
	Middle	0.06	0.25
	End	0.13	0.33
4 letter words	Beginning	0.19	0.27
	Middle	0.09	0.12
	End	0.19	0.19

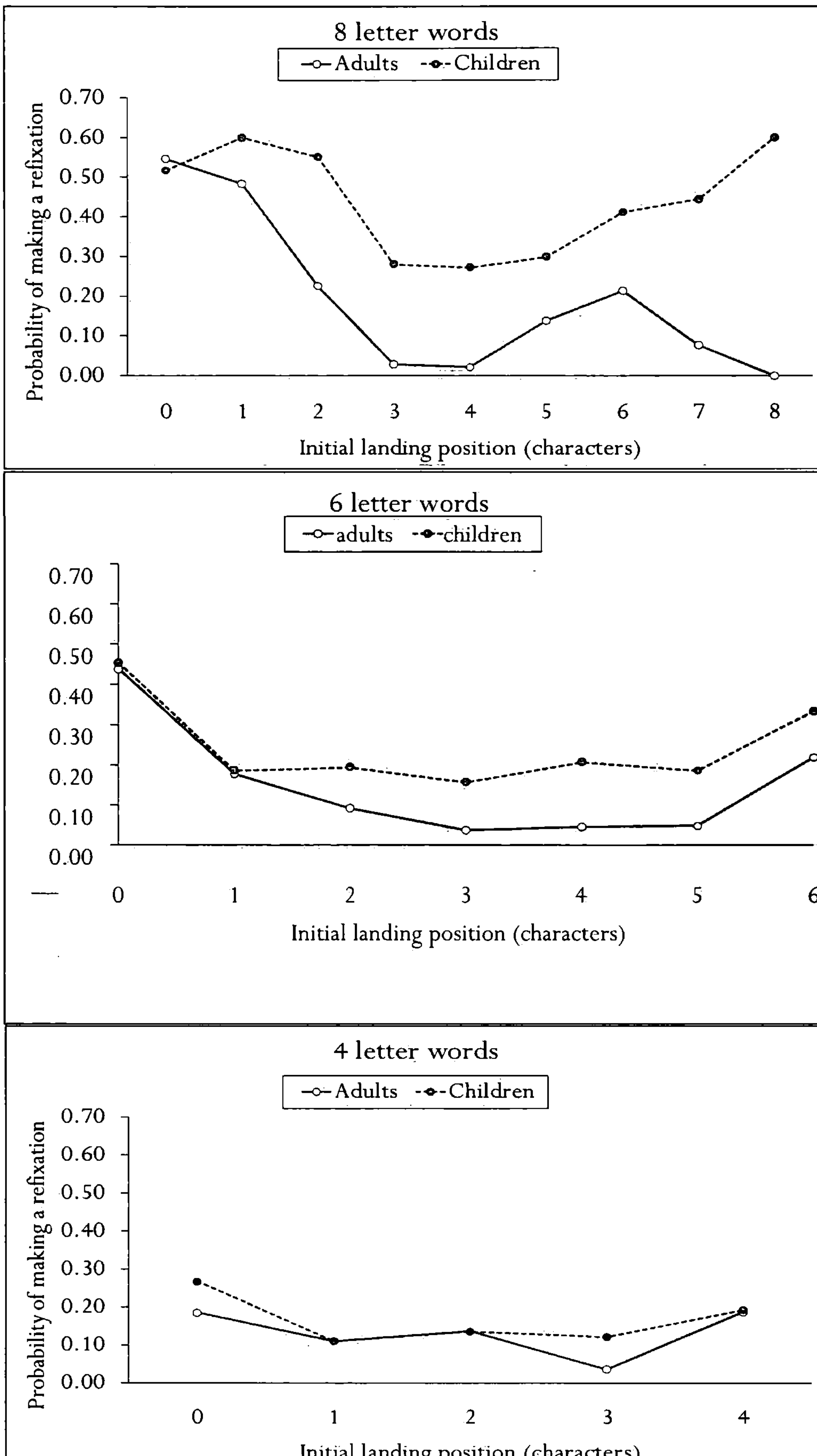
8 letter words
Adults
Children
Probability of making a refixation
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6 7 8
Initial landing position (characters)
6 letter words
adults
children
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4 5 6
Initial landing position (characters)
4 letter words
Adults
Children
Probability of making a refixation
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0 1 2 3 4
Initial landing position (characters)

Figure 3-2: Probability of making a refixation as a function of initial landing position for adults and children on eight (top panel), six (middle panel), and four (bottom panel) letter words.

There was also a reliable effect of word length on the probability of making a refixation, $F(2, 92) = 3.33$, $p < .05$: readers made significantly more refixations on eight, compared to six letter words, $t(39) = 4.05$, $p < .001$; on six, compared with four letter words, $t(39) = 2.90$, $p < .01$; and on eight, compared with four letter words, $t(39) = 5.56$, $p < .001$. There was also a main effect of group, $F(1, 92) = 11.55$, $p < .005$, with children making more refixations overall as compared to adults. Finally, there were no reliable interactions (ps > .1), showing that, contrary to predictions, adults and children did not differ in the probability of making a refixation as a function of initial landing position or of word length: both groups were more likely to make a refixation if they initially fixated the beginning or the end of the word as compared to the middle of the word, in line with the results of previous studies (e.g. Rayner et al., 1996), and more likely to refixate a long than short word.

Finally, the direction of refixation saccades, and the location of those refixations were examined. Specifically, for cases in which there were multiple first pass fixations, whether participants made a progressive (i.e. rightward) or regressive (i.e. leftward) saccade to refixate a word following an initial fixation was calculated. Table 3-3 below shows the proportion of progressive versus regressive intra-word saccades for adults and children on all word lengths.

Table 3-3: Probability of making a progressive or regressive refixation, and mean intra-word saccade length on eight letter, six letter, and four letter words, for adults and children. Standard deviations are in parentheses.

		Progressive	Regressive	Saccade length
8 letters	adults	0.88	0.12	4.82 (1.62)
	children	0.65	0.35	3.54 (*1.47*)
6 letters	adults	0.72	0.28	3.87 (*1.57*)
	children	0.55	0.45	3.06 (*1.48*)
4 letters	adults	0.45	0.55	2.68 (*1.17*)
	children	0.48	0.52	2.48 (*1.24*)

Importantly, there was a significant difference in the proportion of regressive versus progressive refixation saccades between adults and children across word lengths, $t\ (37) = 2.71$, $p < .05$, with adults making a relatively smaller proportion of regressive saccades (26%) compared to children (42%). Furthermore, there were marginal differences in the proportion of regressive versus progressive refixations for the three word lengths, $F\ (1.3, 44) = 2.96$, $p = .086$, and both adults and children made relatively fewer regressive refixation saccades as word length increased.

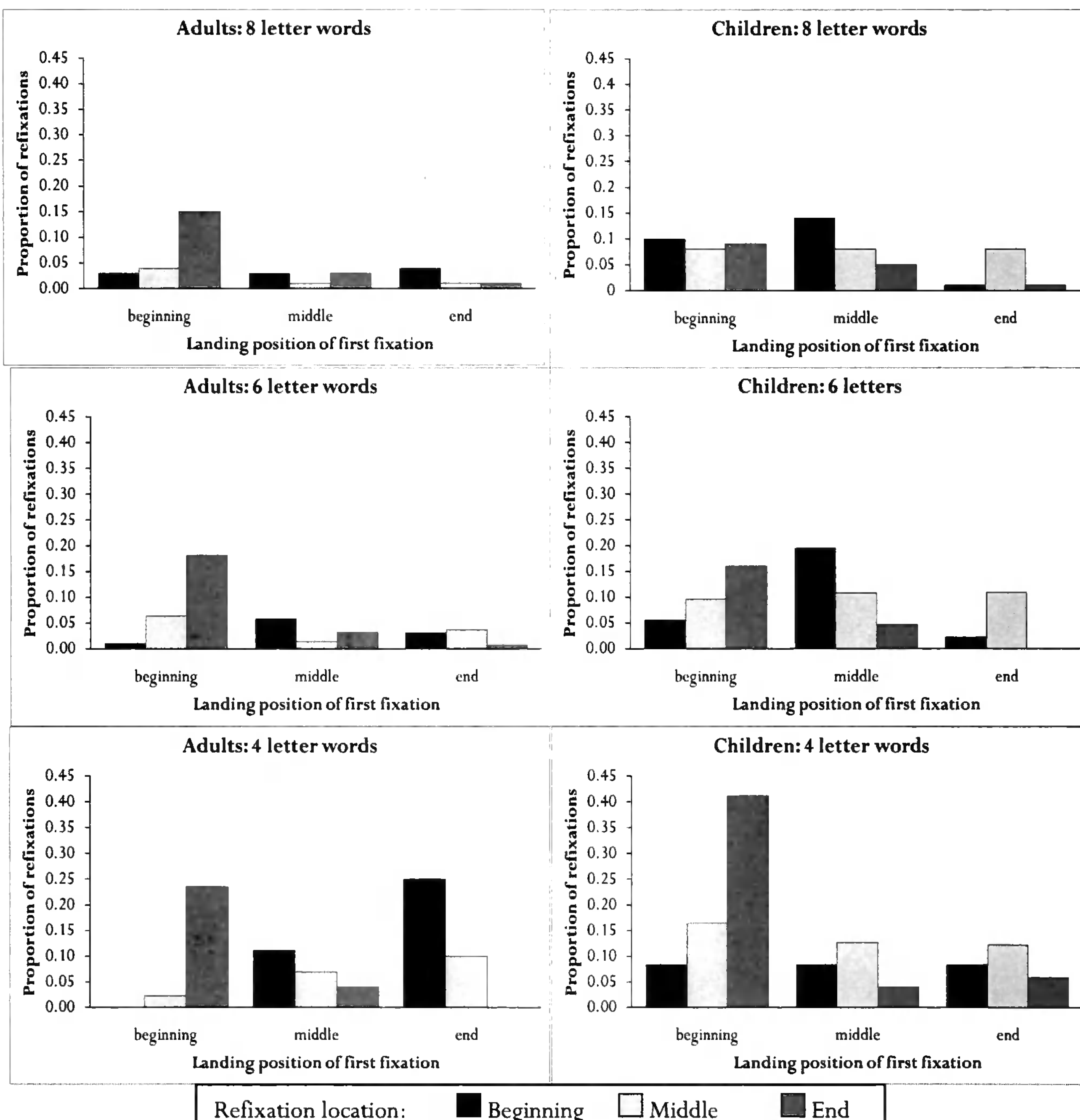

Figure 3-3: Destination of refixations as a function of landing position of initial fixations on eight letter (top row), six letter (middle row) and four letter (bottom row) words for adults (left) and children (right).

Figure 3-3 provides a graphical description of where adults and children located their refixations relative to their first fixations (it was not possible to conduct statistical analyses on these data due to insufficient data points). Figure 3-3 illustrates that adults exhibited quite a clear pattern of refixations in all word lengths, and particularly in the six- and eight-letter words. They tended to initially fixate the beginning of the word and then refixate the end of the word. Children, however, appear to have made shorter saccades than adults, from the beginning to the middle, from the middle to the beginning or from the end to the middle of a word, rather than a long saccade from the beginning to the end. Analyses of saccade length further confirm this observation (see Table 3-3). In line with previous research (Rayner, 1986), children's within-word saccades were shorter (2.87 characters) than adults' (3.71 characters) overall, $F (1, 220 = 12.57$, $p < .005$, and saccade length decreased as word length decreased, $F (1.5, 44) = 12.13$, $p < .001$. However, there was no interaction between word length and group ($ps > .3$).

These results further support the finding that children make more and smaller saccades as compared to adults (Blythe et al., 2006; Rayner, 1986), and can be explained by the fact that the amount of letter information that is available during a fixation increases with age (Häikiö et al., 2008). Children also made relatively more regressive saccades than adults, suggesting that children's saccadic targeting is less strategic than adults' in that adults appear to, on some occasions, intentionally target the beginning then the end of a long word while children consistently target the middle of the word (not always optimal for longer words) and then make a regressive and/or a progressive saccade following this (note that while there were a few three-fixation cases on a single word for children, there were no such cases for adults).

3.3.3 Word length manipulation

The data from 28 long (eight letter) and short (four letter) items were analysed using 2 (age group: adults, children) x 2 (word length: long, short) mixed design ANOVA based on participants (F1) and items (F2) variability. The following eye movement measures were calculated for all critical words (see Table 3-4): skipping probability (the probability of not fixating a word during first pass) first-fixation duration (the duration of the first fixation made in a region); gaze duration (the sum of all first pass fixations made in a region); refixation probability (the probability of making a leftward eye movement out of a region before leaving that region to the right); and total reading time (the sum of all fixations in a region).

Table 3-4: Means, standard deviations (in parentheses), and skipping probabilities of eye movement measures for adults and children on long and short words

	Adults		Children	
	Long	Short	Long	Short
Probability of skipping	0.08 (*0.14*)	0.32 (*0.13*)	0.08 (*0.11*)	0.25 (*0.12*)
First fixation duration (ms)	234 (*77*)	240 (*78*)	288 (*107*)	298 (*117*)
Gaze duration (ms)	280 (*114*)	264 (*99*)	438 (*238*)	350 (*175*)
Probability of refixation	0.23 (*0.16*)	0.12 (*0.12*)	0.41 (*0.24*)	0.17 (*0.16*)
Total reading time (ms)	335 (*151*)	288 (*130*)	663 (*402*)	588 (*404*)

There was a reliable effect of length on the probability of word skipping, with readers skipping short words (0.29) more than long words (0.08), $F1\ (1, 38) = 111$, $p < .001$; $F2\ (1, 27) = 48.89$, $p < .001$. There was no effect of group, which was reliable by items, $F2\ (1, 27) = 5.15$, $p < .05$, but not by participants ($F1 < 1.3$, $p > .2$). Finally, there was no reliable interaction ($Fs < 2.7$, $ps > .1$). Overall, both adults and children skipped short words more than long words.

There was no effect of word length on the duration of the first fixation on the critical word ($Fs < 1.3$, $ps > .2$). However, this is not surprising as it is well-documented that word length effects are generally apparent in gaze durations (Just & Carpenter, 1980; Vitu et al., 1990), rather than first fixation durations (although see Liversedge et al., 2004 for evidence of length effects in first fixation durations). Adults' first fixations were significantly shorter (46ms) than those of children overall, $F1\ (1, 38) = 16.72$, $p < .001$; $F2\ (1, 27) = 55.29$, $p < .001$, and there was no interaction between group and word length ($Fs < .2$, $ps > .6$).

As predicted, there was a highly significant effect of word length on gaze durations, $F1\ (1, 38) = 21.47$, $p < .001$; $F2\ (1, 27) = 18.71$, $p < .001$, as well as a difference between adults and children in their gaze durations overall, $F1\ (1, 38) = 18.79$, $p < .001$; $F2\ (1, 27) = 150$, $p < .001$, with adults again exhibiting shorter gaze durations than children. There was also a significant interaction between group and word length, $F1\ (1, 38) = 12.82$, $p < .005$; $F2\ (1, 27) = 12.05$, $p < .005$.

Pairwise analyses showed that although adult participants had numerically longer (16ms) gaze durations on long than on short words, this difference was not significant (ts < 1.4, ps > .2). Children did, however, exhibit significantly longer gaze durations (88ms) on long than short words, $t1\ (1, 19) = 4.49$, $p < .001$; $t2\ (1, 27) = 4.54$, $p < .001$. The increased magnitude of word length effects on gaze duration for children might be explained by children requiring longer and more visual samples of words, as suggested in the Introduction. However, the effect may also have been exacerbated by differences in text difficulty between the two groups, indicating that processing difficulty may also have a role in modulating word length effects during reading. These possibilities will be commented on further in the Discussion (Section 3.4.2).

Consistent with the null effects in first fixation durations and significant differences in gaze durations, as well as with the analyses reported in the previous section (Section 3.3.2), there was also a reliable effect of length on the probability of making an intra-word refixation, $F1\ (1, 38) = 34.2$, $p < .001$; $F2\ (1, 27) = 23.94$, $p < .001$, as well as an effect of group, $F1\ (1, 38) = 6.28$, $p < .05$; $F2\ (1, 27) = 16.57$, $p < .001$, with children making more refixations overall (29%) than adults (18%). There was also a reliable interaction, $F1\ (1, 38) = 5.53$, $p < .05$; $F2\ (1, 27) = 5.80$, $p < .05$ between word length and group. Pairwise analyses showed that both adults, $t1\ (19) = 3.26$, $p < .005$; $t2\ (27) = 2.80$, $p < .01$, and children, $t1\ (19) = 4.86$, $p < .001$; $t2\ (27) = 4.63$, $p < .001$, showed word length effects on their refixation probabilities. The magnitude of this effect was greater in children. Together, the first pass reading time and refixation probability results show clear first pass word length effects for both adults and children. Adults produced numerically longer gaze durations, significantly fewer skips, and more refixations on long compared to short words, and for children both such differences were significant. For both gaze durations and refixation probability the effects of word length were larger for children compared to adults. Finally, participants spent longer fixating long as compared to short words overall, as shown by total word reading time, F1 (1, 38) = 20.2, $p < .001$; F2 (1, 27) = 10.24, $p < .005$. Consistent with earlier measures, adults spent less time fixating words in total as compared to children, F1 (1, 38) = 29.9, $p < .001$; F2 (1, 27) = 353, $p < .001$). There was no significant interaction between group and condition (Fs < 2.6; ps > .1).

To summarise the results from the length manipulation, both adults and children skipped short words more than long words, showing that both groups of readers were able to use word-length information from the parafovea in order to program a saccade that would skip a short word. In addition, both adults and children made significantly more refixations on long as compared with short words, and for both gaze duration and refixation probability, the magnitude of the effects of word length were greater for children than adults. These findings suggest that children, more than adults, require a second fixation on a long word as, due to their relatively small perceptual spans, they may not have extracted sufficient visual information to lexically identify the word in a single fixation. The larger word length effects for children may also have been at least partially influenced by greater levels of processing difficulty for children compared to adults due to differences in reading ability. Finally, both groups spent longer overall reading long than short words.

3.3.4 Word frequency manipulation

The final manipulation in this experiment investigated word frequency effects in adults and children. The 28 items which were manipulated for frequency were analysed using a 2 (group: adults, children) x 2 (word frequency: high, low) mixed design ANOVA, based on participants (F1) and items (F2) variability.

Table 3-5: Means, standard deviations (in parentheses), and skipping probabilities of eye movement measures for adults and children on high and low frequency words.

	Adults		Children	
	High	Low	High	Low
Probability of skipping	0.15 (*0.12*)	0.16 (*0.20*)	0.13 (*0.13*)	0.20 (*0.20*)
First fixation duration (ms)	233 (*69*)	238 (*67*)	289 (*113*)	290 (*120*)
Single fixation duration (ms)	238 (69)	243 (*66*)	298 (119)	283 (106)
Gaze duration (ms)	261 (*90*)	267 (*88*)	383 (*212*)	382 (*197*)
Total reading time (ms)	291 (*120*)	307 (*131*)	612 (*364*)	679 (*432*)

There was no reliable effect of word frequency in skipping probability (ps > .09), no effect of group (ps > .5), and no interaction between group and frequency (ps > .2). In first fixation durations, there was no effect of frequency (ps < .6), but there was an effect of group, F1 (1, 38) = 13.5, p < .005; F2 (1, 26) = 77.46, p < .001, with children exhibiting longer first fixation durations as compared to adults, but no interaction (p > .8). In single fixation durations, there was no effect of frequency (ps > .2), and effect of group, F1 (1, 38) = 15.16, p < .001; F2 (1, 26) = 43.3, p < .001, and no interaction (ps > .17) Likewise, in gaze duration, while children exhibited longer gaze durations than adults overall, F1 (1, 38) = 16.68, p < .001; F2 (1, 26) = 126, p < .001, there were no significant effects of word frequency and no interaction (Fs < 1.5, ps > .2). However, participants did make marginally more first pass fixations on low as compared to high frequency words, F1 (1, 38) = 3.93, p = .055; F2 (1, 26) = 3.19, p = .086, in addition to the effect of group, F1 (1, 38) = 12.66, p < .005; F2 (1, 26) = 30.48, p <.001, but again there was no interaction between group and frequency (Fs < .7; ps > .4).

In total word reading time, there was a reliable effect of word frequency, F1 (1, 38) = 7.34, p < .05; F2 (1, 26) = 6.66, p < .05, with participants taking longer to read low as compared to high frequency words. There was also a difference between adults and children's total reading time, F1 (1, 38) = 34.17, p < .001; F2 (1, 26) = 571, p < .001. There was a marginal interaction between group and frequency, F1 (1, 38) = 3.19, p = .08; F2 (1, 26) = 3.19, p = .08. Pairwise analyses showed that adults did not exhibit a significant difference in total reading times between the two frequency conditions (ts < 1.4; ps > .19), but that children did spend longer (65ms) reading low as compared to high frequency words, $t1\ (1, 19) = 2.4, p < .05$; $t2\ (1, 26) = 2.34, p < .05$.

Overall, the results from the frequency manipulation show that while children were influenced by the frequency of the target word, the effect appeared only when second pass reading times were included, and adults showed no effect of word frequency. Given the plentiful evidence for word frequency effects in adult readers, this suggests that there was a problem with the manipulation employed. There are several reasons why this might be the case. First, the choice of target words was highly constrained due the use of the same experimental stimuli with both adult and child participants. As a result all target words were simple and acquired relatively early in life (before age eight), which meant that the frequency manipulation was necessarily less strong than would be used in an adult experiment investigating frequency effects. This might explain why the adult participants, most of whom were highly skilled readers, did not exhibit frequency effects. Second, the frequency counts were based on adult corpora which might be inappropriate for children. These two issues will be commented on further in the Discussion section (Section 3.4.3) of this chapter as they motivated the experiments carried on in Chapter 4.

3.4 Discussion

The experiment reported in this chapter investigated three issues: 1) Do children, like adults, target their initial saccades towards the word centre and does the location of this first fixation influence their refixation behaviour within the word? 2) Do children, like adults, require greater processing resources to read long, as compared to short words? 3) Do children, like adults, process high frequency, as compared to low frequency words more quickly (i.e. do they require fewer processing resources to process high than low frequency words?)? These three issues will be discussed separately.

3.4.1 Landing positions

This part of the experiment examined landing position distributions on eight, six and four letter words. Results showed that while adults and children were alike in their fixation location on all word lengths. Both adults and children initially fixated near the PVL on eight-letter words, and a little to the right of the PVL on the shorter length words, consistent with previous adult studies (Rayner, 1979; Vitu et al., 2001).

Although there has been relatively little empirical research investigating children's landing positions during text reading, these results sit well with the data from both the McConkie et al. (1991) and the Vitu et al. (2001) studies which show that children are adult-like in targeting their initial saccades in a word toward the word centre. Furthermore, these data provide empirical support for Reichle and Laurent's (2006) reinforcement learning model in which "intelligent" eye movements, including fixating close to the word centre, emerge quite quickly during learning. Finally, these data provide evidence that while adults and children do not generally differ in early oculomotor decisions regarding saccade targeting, children have yet to adopt in full the efficient strategy employed by adults according to which initial fixations are made towards the beginning, and refixations made towards the end, of longer words.

Refixation probability as a function of initial landing position was also examined. Both adults and children were more likely to refixate a word if their initial fixation was away from the word centre (i.e. at the beginning or end of the word), and while children made more refixations overall as compared to adults, there was no reliable interaction between group and the probability of making a refixation as a function of landing position. These results show that both adults and children make early decisions regarding intra-word refixations, either on the basis of initial fixation location, or perhaps even earlier while programming their initial saccade to a word (as argued above for the adults on eight letter words). Note that there may also be linguistic differences between adults and children which account for refixation and skipping behaviours, for example in their capacity to use predictability information to lexically identify a word thereby eliminating the need to refixate it, or enabling a reader to skip it entirely.

The direction of refixation saccades were also examined and it was found that adults made proportionally more progressive refixation saccades as compared to children, and that the saccades that adults made were longer than those made by children. If children require more visual samples of the word they are fixating before moving on to process the subsequent word, as suggested by the data from the word length analyses, both regressive and progressive saccades may be necessary for word recognition processes to be completed, especially given that children tend to initially fixate the word centre even when the word is long. This conclusion is in line with previous research showing that children make smaller saccades than adults during reading (Rayner, 1986) and have less letter information available to them during a fixation (Häikiö et al., 2008; Rayner, 1986). Both children and adults also made fewer regressive saccades as word length increased, perhaps because, in relation to the word centre, the initial fixation on a word tended to be further into the word for short than long words. Another possible explanation for this tendency is that, as argued, two-fixation cases on long words may be pre-planned while two-fixation cases on short words are not. As discussed, on long words, this would result in the first fixation landing relatively close to the beginning of the word, and the refixation landing close to the word end. In contrast, a relatively large proportion of two-fixation cases on shorter words may have been corrective (as readers would presumably not plan to fixate such short words more than once) due to an initial mislocated fixation close to the word end. This tendency was (numerically) more pronounced in adult readers, and given that adults also initially fixate long words closer to the word beginning than children, the data support this argument.

3.4.2 Word length

It is well-documented that adults exhibit reliable differences in their eye movement behaviour when reading long as compared to short words. In this part of the experiment, it was found, in concordance with previous research, that adults' eye movements were significantly influenced by word length, even when linguistic influences (specifically word frequency and predictability) were controlled for. Adults showed both first pass and second pass effects of word length. However, the main focus was to establish whether children showed these same robust effects. In line with predictions, it was found that children exhibited stronger effects of word length as compared to adults in gaze durations. These results suggest that children, like adults, experience an increased processing load when reading long as compared to short words, but more interestingly, that the increase in word length has a more substantial effect on child readers' ongoing processing as compared to that of adult readers. Because word frequency and predictability were carefully controlled, this difference between adults and children must have been due to the visual demands of long as compared to short words, suggesting that children require more and/or longer visual samples of words in order to reach the point at which they are able to begin processing the following word.

However, it is important to note that, unlike in previous adult studies, there was not a reliable effect of word length in gaze durations for the adult group. This may be because the 'long' words in Experiment 1 were only eight letters long, shorter than what might be expected in adult stimuli containing a word length manipulation (e.g. in Liversedge et al., 2004, long words were ten letters), and because the adult participant group (undergraduate students) were proficient and experienced readers. The reason eight letter words were used in the long condition was that very long words are less likely to be familiar to young readers. In addition, the sentences were designed to be easily comprehensible to young children making them very simple to read for adults. This constraint on the creation of linguistic stimuli is inherent in conducting developmental research of this kind. Experimental sentences will always be age-appropriate for the youngest children taking part, and these same sentences will therefore be necessarily easy for adult participants to process. However, this situation cannot be avoided if the critical premise of the thesis is to be maintained; that is, robust processing preferences documented in adult readers are used diagnostically to examine children's processing preferences in order that direct comparisons can be made between participant groups. In relation to this point, Chapter 4 directly investigates the influence of age-appropriate stimuli on reading in adults, and the methodological implications of conducting research in this way will be commented on further in Chapter 8 of the thesis.

It was also found that both children and adults were more likely to skip short words as compared to long words. The lack of a reliable interaction showed that both adults and children were able to use parafoveal information to target or reprogram a saccade in order to skip an easy-to-process (i.e. short) word. This finding is interesting in conjunction with the differences between adults and children in fixation durations on long and short words, as well as the results from the landing position analyses, as it shows that although children appear to require a more detailed visual sample of a word as they read, they do not differ from adults in early decisions regarding saccade targeting. Thus far then, the results from the first two parts of the experiment, which address visual processes during word reading, show that while children and adults are broadly alike in early oculomotor behaviour which is governed by non-linguistic factors, such as targeting saccades to land close to the word centre, or to skip a parafoveal word, differences appear to emerge during later stages of word processing, such as refixation locations when linguistic influences have begun to influence eye movement behaviour.

3.4.3 Word frequency

In the final manipulation of this experiment, word frequency effects were investigated. Like word length effects, word frequency effects are extremely robust in adult readers and so it was important to establish whether children also exhibited these effects. Results showed that adults did not exhibit effects of word frequency at all and children exhibited frequency effects only in total reading time. Given the ubiquity and robustness of the word frequency effect, it is clear that these findings cannot be interpreted as a true reflection of adults' lexical processing mechanisms, and consequently, the results from the children cannot be taken as reliable either. It must be, then, that there was a weakness in the stimuli used which failed to generate the predicted effects in word frequency.

As mentioned in the Results section, there are several possible reasons why a reliable word frequency effect was not found. First, when using the same stimuli with adults and children, the manipulation employed will always be weaker than when testing adults alone. Sentences for use with children will necessarily be simple, both in syntax and in the ideas conveyed. All words in the sentences were acquired before the age of eight, including the target words, and may have been more predictable from the context than target words in standard adult experiments. Most importantly, the range of words available for the frequency manipulation was drastically reduced as compared to in an adult experiment, as very low frequency words would not be known to young children. As a result, the manipulation was greatly constrained before it had even begun.

Second, and very importantly, while experiments which have manipulated word frequency in adult readers have used standard corpora such as CELEX (Baayen et al., 1995) or Kučera-Francis (Kučera & Francis, 1967) to index word frequency counts, this may not be appropriate for child readers. Databases such as these are based on written texts for adult readers, and therefore arguably do not provide a realistic index of frequency for young children who have little experience of reading, certainly of reading texts written for adults. In the one study which did find frequency effects in child readers (Hyönä & Olson, 1995), frequency was indexed by counts from age-appropriate texts (i.e. school text books and reading books) and so it may be that with a more suitable gauge of word frequency, a reliable effect may be observed in child readers. The two experiments outlined in Chapter 4 set out to address exactly this issue.

Chapter 4 : Word frequency effects

4.1 Introduction

In Chapter 3, reliable first pass frequency effects were not observed in adult or in child readers using a conventional index of word frequency that has been used in adult research. This surprising result raised the question of whether adult indices of word frequency such as CELEX (Baayen et al., 1995) and Kučera Francis (Kučera & Francis, 1967) databases are appropriate as sources of word frequency indices for children. As mentioned, these conventional indices of frequency are based on texts written specifically for adults and, as such, may not reflect the frequency with which children, particularly beginning readers, encounter the words listed. Indeed although there was a reliable difference in word frequency between the target words in the high and low frequency conditions in Experiment 1 when frequencies were taken from the CELEX database; when frequency counts were taken from a database which used texts written for children to index word frequency (The Children's Printed Word database: Stuart, Masterson, Dixon, & Quinlan, 2002), the difference between the conditions was far from significant ($t < 1$, $p = .4$).

On closer inspection, three of the pairs of target words, such as *rattle* and *marble*, which are low, and high frequency respectively according to the CELEX database, actually have the opposite pattern on frequencies according the Children's Printed Word Database (CPWD from hereafter), that is, *rattle* is a high frequency word (260 counts per million) and *marble* is a low, frequency word (5 counts per million). A further eleven pairs of words were simply weak manipulations according to the CPWD, such as *coffee* (32 counts per million) and *cherry* (16 counts per million), while only five word pairs (frequency counts were not available from the CPWD for the remaining nine word pairs) were strong manipulations for both adults and children such as *people* (1926 counts per million) and *tailor* (8 counts per million). It is probable, then, that the manipulation employed in Experiment 1 was not satisfactory for child readers and further experimentation was necessary to investigate frequency effects in children.

Also, as was seen in Chapter 3, the frequency manipulation did not appear to be satisfactory for adult readers. Because the range of frequencies was restricted due to the need for all target words to be familiar to children as young as seven years old, and because the syntactic structure and semantic content of the sentences was necessarily simple and easy to comprehend, it appeared likely that the low frequency words did not require sufficient additional processing resources so as to generate strong effects. It was therefore decided to introduce a stronger frequency manipulation for the adult participants in order to obtain a robust word frequency effect. These target words would be embedded in sentences which were age-appropriate for adult readers, and the same adults would read two sets of sentences: first, sentences which were age-appropriate for adults and contained a strong frequency manipulation as indexed by an adult corpus, and second, sentences which were age-appropriate for children as young a seven and which contained a frequency manipulation as indexed by a child corpus. Because the same adults would read both sets of sentences, it would be possible to tease apart the effect of these two different indices of frequency on adults, and to compare effects of the frequency manipulation indexed by a child corpus on adults and children.

Another factor that has been shown to influence the speed with which we process a word is Age-of-Acquisition (AoA). As mentioned (Chapter 1, Section 1.3.4), words that are acquired earlier in childhood are processed more quickly and accurately than words acquired later, and AoA effects are observed in many different tasks (lexical decisions, picture naming, word naming, face recognition), using different stimuli (words, pictures, faces), languages, and participant groups (adults, older adults, clinical patients, children). AoA effects are as strong as frequency effects, and have often been confounded with frequency effects in adult studies (for a full review of A-o-A effects, see Juhasz, 2005). Because AoA and word frequency are highly correlated (i.e. frequent words are usually acquired early and low frequency words acquired late, although there are some notable exceptions such as *dragon*), it is notoriously difficult to control for word frequency and manipulate AoA (or vice-versa), although both frequency and AoA have been shown to influence eye fixations independently (Juhász & Rayner, 2006).

In Experiment 1, although it was ensured that all target words were acquired before the age of eight, it was not possible to control for AoA, and so it may have been that AoA effects concealed or modulated any frequency effects that were in fact present. While it is difficult to control for AoA and manipulate word frequency using adult indices of frequency, it is entirely possible to control for AoA and obtain a very strong frequency manipulation when using child indices of frequency. Therefore in Experiment 3, target words were tightly controlled for AoA as well as being manipulated for frequency. This ensured that any frequency effects observed in children or adults were real and not confounded by AoA.

This chapter describes two experiments devised to test whether the failure to obtain frequency effects in Experiment 1 was due to the index of frequency used and/or the age at which target words were acquired. In Experiment 2, adult participants read sentences which contained a strong frequency manipulation using adult frequency counts. Note that children could not take part in Experiment 2 as the target words, and indeed the sentences as a whole, would be too difficult for them. Sentences in Experiment 2 were more syntactically and semantically complex than those in Experiment 1 (and Experiment 3), making them age-appropriate for adult readers. In Experiment 3, the same adults who took part in Experiment 2, and children, read sentences which contained target words manipulated for frequency as indexed by child frequency counts, which were additionally controlled for AoA. These sentences were syntactically and semantically simple so that they would be easily comprehended by children as young as seven, as well as the adults.

In Experiment 2, it was predicted, in line with numerous studies (Henderson & Ferreira, 1990; Inhoff, 1984; Just & Carpenter, 1980; Rayner & Duffy, 1986; Rayner & Raney, 1996), that adults would exhibit significantly longer reading times on low, as compared to high frequency words. In Experiment 3, in which frequency was indexed by child frequency counts, it was predicted that children would look longer at low, as compared to high frequency words, and that these effects would exist independent of AoA. For these same sentences, it was predicted that adults would not exhibit frequency effects as the index of frequency used was not appropriate for adult readers.

4.2 Experiment 2

4.2.1 Method

Participants. 30 adults, all from Durham University, took part in the experiment. The same conditions for taking part outlined in Experiment 1 (Chapter 3: Section 3.2.1) were also applied here (no known reading disabilities, native English speakers).

Materials. There were 24 experimental items, all of which contained a frequency manipulation (see sentences 1a and 1b below). All word frequencies were taken from the CELEX database (Baayen et al., 1995), as in Experiment 1.

(1a) It's a lovely little street and it has real character.

(1b) It's a lovely little bistro and it has real character.

The high frequency words (e.g. *street* in 1a) had a mean frequency of 344 counts per million (SD = 263), ranging from 172 to 1480 counts per million, and the low frequency words (e.g. *bistro* in 1b) had a mean frequency of 1.04 counts per million (SD = 0.20), ranging from 1 to 2 counts per million. The difference in frequency between the two conditions was highly significant, $t(23) = 6.38$, $p < .001$. Thirty pairs of sentences (both high frequency and low frequency versions of each sentence were given) were rated for plausibility by 12 adults who did not take part in the experiment. Of these 30 sentences, six items were excluded due to differential ratings for the two conditions. Of the 24 remaining items, there was no significant difference in their plausibility ($t < 1$, $p > .7$).

Apparatus. Participants' eye movements were recorded using a head-mounted Eyelink II eye tracker manufactured by SR Research (Mississauga, Canada), as they read sentences from a computer monitor at a viewing distance of approximately 100cm. The eye tracker was an infrared video-based tracking system with two cameras mounted on a headband which were placed approximately 5cm from the eyes. Head position was detected by four LEDs attached to the computer monitor, and any movements were compensated for in the eye movement records. Participants' eye movements were monitored at a rate of 500Hz to produce a sequence of fixations with start and finish times. Although participants read binocularly, only the movements of one eye were monitored.

Procedure. Participants sat in a customised chair in front of a computer monitor. The eye tracker was placed on the participant's head and secured by adjusting two headbands. Two cameras were placed in front of the eyes. Participants undertook a calibration procedure during which they looked at each of three horizontal fixation points. Participants then looked at a fixation box at the left of the screen and the sentence appeared contingent on their gaze. Participants were required to read the sentences normally and then press a button when they had finished reading. The button press terminated the display. If the participant did not press the button within 15 seconds of the sentence appearing, the display was automatically terminated. In addition to the 24 experimental items, two practice items were also presented at the beginning of the experiment. Sentences were interspersed with experimental sentences from Experiments 3, 5 and 6. Participants were asked to respond yes/no to comprehension questions after six of the sentences by pressing a button. The experimental session lasted approximately 35 minutes in total.

4.2.2 Results

Data for target words in all 24 items were analysed. First fixation durations, single fixation durations (the duration of a fixation when it is the only first pass fixation on a word) and gaze durations were calculated (see Chapter 3, Section 3.3.3 for a description of measures). Table 4.1 shows the mean reading times for high and low frequency words for these measures in adults and children.

Table 4-1: Mean fixation durations on high and low frequency target words for adults. Standard deviations in parentheses.

	High frequency	Low frequency
First fixation duration (ms)	212 (*65*)	241 (*82*)
Single fixation duration (ms)	218 (*66*)	244 (*75*)
Gaze duration (ms)	233 (*93*)	291 (*163*)

Adults made significantly longer (by 29 ms) first fixations on low than high frequency target words, t1 (29) = 6.37, p < .001; t2 (23) = 4.71, p < .001, in line with predictions. This same difference was observed both in single fixation durations, t1 (29) = 6.57, p < .001; t2 (23) = 4.27, p < .001, and in gaze durations, t1 (29) = 4.59, p < .001; t2 (23) = 3.98, p < .005, as illustrated in Figure 4.1 below. These results clearly show that by using age-appropriate materials with a frequency manipulation that is indexed by adult corpora, adults showed a strong effect of word frequency.

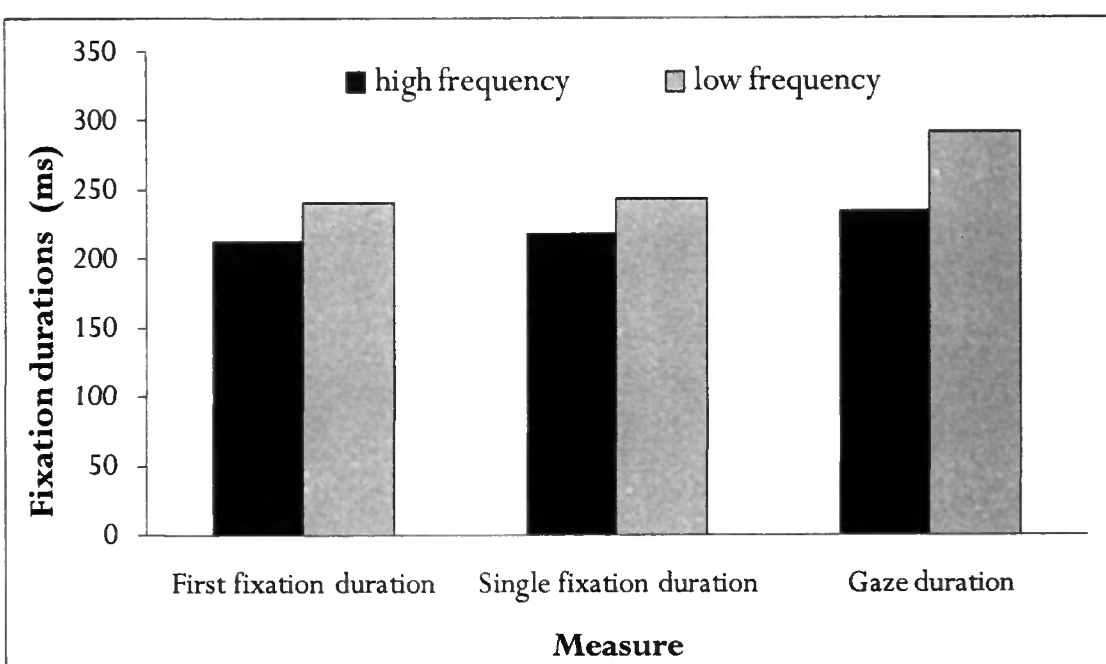

Figure 4-1: First fixation durations, single fixation durations and gaze durations on high and low frequency words for adult readers

4.2.3 Discussion

The results from Experiment 2 show that when reading sentences which are written for an adult audience, and when using a strong frequency manipulation derived from adult norms, adults exhibited a strong effect of word frequency, fixating significantly longer on low, as compared to high frequency target words on three early measures of processing time. These findings are in line with previous research which has shown that word frequency is a central characteristic of the lexicon and the frequency with which we encounter a word has a powerful and immediate influence on how long it takes to lexically identify that word. However, this finding alone is not particularly interesting given the wealth of evidence already in support of word frequency effects in adult readers (e.g. Henderson & Ferreira, 1990; Inhoff, 1984; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner, 1977; Rayner & Duffy, 1986; Rayner, Liversedge et al., 2003; Rayner & Raney, 1996). These results are important in the context of this thesis however, as they provide the basis from which to carry out the next experiment in which frequency counts were taken from child corpora. If adults show no difference in their eye movement behaviour between the high, and low frequency words using these frequency counts, while children do show a difference, then this will show that the lexicon develops and is organised in terms of frequency, and that it reflects the order in which an individual has encountered words over a lifetime (which is different in adults and children). This finding will have important implications both theoretically and methodologically for understanding of the role of word frequency in lexical processing.

4.3 Experiment 3

4.3.1 Method

Participants. The same 30 adults who took part in Experiment 2 participated in Experiment 3, as part of the same experimental session. In addition, ten children took part in the experiment. Participants were recruited as in Experiment 1 (Chapter 3: Section 3.2.1). Children were all between seven and eleven years old, with a mean age of 9.6 years. All children completed the WORD (see Chapter 3, Section 3.2.1 for further details on this measure). All children scored within 2SDs of the standardised norm for their age (mean = 100; SD = 15), with an overall mean score of 111. The mean reading age was 11.1 years. Children received a small gift in return for taking part.

Materials. There were 20 experimental items in total, all of which contained a frequency manipulation (see sentences 2a and 2b).

(2a) All the children loved the fairy because she was kind and beautiful.

(2b) All the children loved the nurse because she was kind and beautiful.

The mean frequency for the target words in the high frequency condition (*fairy* in 2a) was 202 (SD = 113) counts per million (ranging from 95 to 552 counts per million) and the mean frequency for the target words in the low frequency condition (*nurse* in 2b) was 18 (SD = 14.3) counts per million (ranging from 3 to 57 counts per million). All frequencies were taken from the Children's Printed Word database (Stuart et al., 2002). The difference in frequencies between the two conditions was statistically significant, $t(19) = 7.26$, $p < .001$ and words which were manipulated for frequency were controlled for length and for AoA. High frequency words had a mean AoA of 36 months (ranging from 20 to 52 months), and low frequency words had a mean AoA of 39 months (ranging from 17 to 61 months). AoA norms were taken from Morrison, Chappell and Ellis (1997), as they used objective measurements, rather than adult ratings, with British schoolchildren, the same population that took part in our experiment.

Fifty sentences containing a frequency manipulation were given to 17 adults, 14 children aged 7-8 years, and 14 children aged 10-11 years, none of whom took part in the main experiment. All 50 sentences were rated for plausibility. From these sentences 20 were chosen for which both the high frequency and low frequency conditions were rated as equally plausible. Pairwise comparisons showed that there was no difference in plausibility between high, and low frequency conditions for adults or children of either age group of children ($ps > .15$).

Apparatus. The apparatus used was identical to that in Experiment 2 (Section 4.2.1).

Procedure. The procedure used was identical to that in Experiment 2 (Section 4.2.1), with the addition that memory foam was attached to the inside of the headband for some of the younger children in order to make it more comfortable and secure. Experimental sentences were interspersed with sentences from Experiments 5 and 6 (and sentences from Experiment 2 for the adult participants). In addition to the experimental items, two practice items were also presented at the beginning of the experiment. The experimental session lasted approximately 25 minutes for children, and 35 minutes for the adults.

4.3.2 Results

Data for target words in all 20 items were analysed. As in Experiment 2, first fixation duration, gaze duration, single fixation duration and total reading time were calculated. Table 4.2 shows the mean reading times for high and low frequency words for these measures in adults and children.

Table 4-2: Mean reading times on high and low frequency words for adults and children. Standard deviations in parentheses.

	Adults		Children	
	High frequency	Low frequency	High frequency	Low frequency
First fixation duration	217 (*62*)	217 (*69*)	261 (*108*)	306 (*161*)
Single fixation duration	221 (*63*)	217 (*67*)	270 (*118*)	315 (*169*)
Gaze duration	232 (*73*)	237 (*103*)	316 (*145*)	413 (*282*)
Total time	266 (*117*)	274 (*137*)	463 (*340*)	698 (*687*)

A 2 (group: adults, children) x 2 (word frequency: high, low) mixed design ANOVA showed an effect of word frequency on first fixation duration, reliable by participants but marginal by items, $F1\ (1, 38) = 11.39$, $p < .005$; $F2\ (1, 19) = 3.72$, $p = .07$. There was also a reliable effect of group, $F1\ (1, 38) = 24.5$, $p < .001$; $F2\ (1, 19) = 63.62$, $p < .001$, with children making longer first fixations than adults. Finally there was a significant interaction between group and frequency, $F1\ (1, 38) = 11.30$, $p < .005$; $F2\ (1, 19) = 5.17$, $p < .05$. While adults showed no effect of frequency (indeed they exhibited identical first fixation durations on both high and low frequency words), children made significantly longer first fixations on low as compared to high frequency words, $t1\ (9) = 2.33$, $p < .05$; $t2\ (19) = 2.22$, $p < .05$.

In single fixation durations, there was no effect of frequency: while it was reliable by participants, $F1\ (1, 38) = 5.97$, $p < .05$, it was not reliable by items, $F2\ (1, 19) = 1.37$, $p = .26$. There was also an effect of group $F1\ (1, 38) = 27.16$, $p < .001$; $F2\ (1, 19) = 24.70$, $p < .001$, with children making longer single fixations than adults. There was no interaction as it was not reliable by items, $F1\ (1, 39) = 7.13$, $p < .05$; $F2\ (1, 19) = 2.60$, $p = .12$. However, the numerical trend was the same as in first fixation duration (that is that children exhibited longer single fixations on low frequency than high frequency words but adults did not).

In gaze duration, there was a significant effect of frequency, $F1\ (1, 380 = 17.90$, $p < .001$; $F2\ (1, 19) = 6.31$, $p < .05$, a reliable effect of group, $F1\ (1, 38) = 39.26$, $p < .001$; $F2\ (1, 19) = 45.13$, $p < .001$, with children exhibiting longer gaze durations than adults, and a reliable interaction, $F1\ (1, 38) = 14.72$, $p < .001$; $F2\ (1, 19) = 6.49$, $p < .05$ (see Figure 4-2). When analysed separately, adults showed no difference between the high and low frequency conditions ($ps > .4$) but children looked significantly longer (97ms) at low, as compared to high, frequency words, $t1\ (9) = 2.53$, $p < .05$; $t2\ (19) = 2.62$, $p < .05$. Figure 4-2 shows that across three early measures of processing difficulty, while children exhibited large differences in reading times between high, and low frequency words, adults showed no such differences.

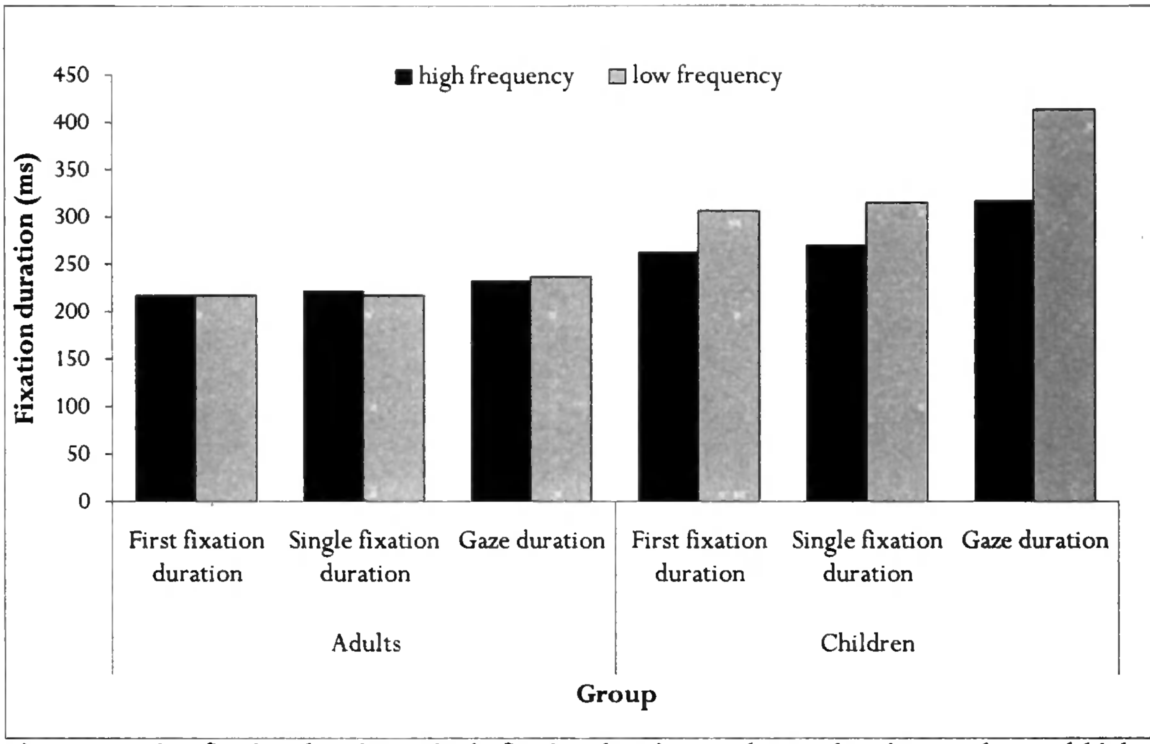

Figure 4-2: First fixation durations, single fixation durations and gaze durations on low and high frequency words for adults and children

Finally, in total time, there was an effect of frequency, $F1\ (1, 38) = 17.90$, $p < .001$; $F2\ (1, 19) = 6.31$, $p < .05$, an effect of group, $F1\ (1, 38) = 39.26$, $p < .001$; $F2\ (1, 19) = 45.13$, $p < .001$, and a reliable interaction, $F1\ (1, 38) = 14.72$, $p < .001$; $F2\ (1, 19) = 6.49$, $p < .05$. When analysed separately, adults showed no difference in their total reading times between high and low frequency words ($ps > .4$), but children showed longer total reading times on low as compared to high frequency words, $t1\ (9) = 2.44$, $p < .05$; $t2\ (19) = 3.26$, $p < .005$.

4.3.3 Discussion

The results from Experiment 3 very clearly show that while children's eye fixations were reliably influenced by word frequency, as indexed by counts from children's reading material, adults' eye fixations were not. These results show that for children, as well as adults, the frequency with which we encounter a word has an immediate and robust influence on how long it takes to lexically identify that word. Indeed, the time course of these effects is the same for adults and children, that is, it is evident in the very first fixation made on a word. This result is striking because it shows that children as young as seven, or in other words only two years into the reading acquisition process, already have their lexicons organised in terms of how often they have encountered a word. Moreover, this effect is independent of the age they acquired the word, showing that word frequency is a critical characteristic of the lexicon over and above AoA.

Furthermore, the results show not only that the frequency with which children encounter words is not the same as the frequency with which adults encounter words, but that frequency counts derived from child corpora do not generate the robust frequency effects in adult readers as frequency counts from adult corpora are known to do. This finding has methodological consequences for future research: how often a word is encountered differs over a lifetime, and while using adult corpora when conducting experiments with adult participants is appropriate, it is not appropriate when doing research with children.

4.4 General Discussion

The results from Experiment 2 and 3 clearly show a differential pattern of effects for adults and children with regard to word frequency. While adults exhibited significantly longer fixation durations on low than high frequency words when frequency was indexed by counts based on adult texts, they did not exhibit this same effect when frequency was indexed by counts based on children's texts. Moreover, children's eye movements were influenced by word frequency when reading these same sentences in which frequency was indexed according to age-appropriate texts. Finally, word frequency had an effect on eye fixations independently of any Age-of-Acquisition effects in child readers.

The results lead to three main conclusions. First, the results from both Experiment 2 and Experiment 3 have shown that linguistic influences are primary in driving eye movements during reading, consistent with cognitive models of eye movement control. Although target words were controlled for word length and plausibility, and the sentences were identical apart from the target word, fixation durations changed relative to a linguistic variable. This conclusion supports cognitive models of eye-movement control during reading such as the E-Z Reader (Reichle et al., 2004) which assume that ongoing linguistic processing influences eye movement behaviour during reading, more so than oculomotor factors. While the finding isn't inconsistent with the idea that eye movements depend on visual as well as linguistic characteristics of the text, it nevertheless shows that the mental processes associated with understanding a word determine how long the eye looks at that word.

The second conclusion which can be drawn from the data reported here is that word frequency is a fundamental characteristic in the organization of the lexicon in English for children as well as adults. While it is extremely well-documented that word frequency enjoys a privileged position in the lexicon of adult readers, and that the frequency with which adults encounter a word affects the speed with which that word is lexically identified, the results from Experiment 3 show that this is also the case in children. Very early in reading development, word frequency is central to the lexical processing system, and has a profound effect on the ease or difficulty with which a word is processed.

The final conclusion to be drawn from these results is that the frequency with which we encounter a word over a lifetime, as well as the age at which we learn a word, is an important factor in how quickly we identify that word. While word recognition processes are affected by the frequency of encounter of a word in both adults and children it cannot be assumed that the indices we have in place to estimate the degree to which adults encounter or experience different features of language can, or should be, applied to children. It seems that there are (at least) three factors which affect the time course of lexical identification. First, as shown by previous research (Juhasz & Rayner, 2006; Juhasz, 2005), if a word was acquired early in life, that word will be identified significantly faster than a word which was acquired later in life. Second, the frequency with which a word has recently been encountered, a kind of snapshot of frequency, affects the time taken to lexically identify it: words which are encountered often are identified more quickly. Finally, the cumulative frequency with which we have encountered a word over a lifetime affects the time needed to identify it and this measure of frequency clearly changes over the course of development (as the language an individual is exposed to changes). Cumulative frequency is of particular importance in relation to child readers as the frequency of occurrence of a word in childhood may be more dynamic, with each encounter having a greater impact on frequency levels as compared to in adulthood when frequency levels may be more static. Overall, the frequency with which a word is encountered is cumulative, and therefore changes over time. This necessarily affects psychological processing in that the nature of the lexicon impacts on reading behaviour. Age-of-Acquisition, "snapshot' frequency and cumulative frequency should all be considered when investigating word recognition processes.

Chapter 5 : Lexical ambiguity

5.1 Introduction

Experiments 2 and 3 in Chapter 4 showed that the frequency with which a word is encountered fundamentally affects the ease with which both adults and children identify that word during reading. That is, word frequency is important in respect of both the organisation and the function of the developing and the mature mental lexicon. Indeed, in order that frequency effects can be observed so immediately and robustly during reading, frequency must play a critical role in word recognition processes. Experiment 4, reported in this chapter, builds on the findings from Experiments 2 and 3 to examine the role of meaning in lexical identification in children and adults by focusing on words which are lexically ambiguous. A word can be lexically ambiguous in terms of meaning, that is, a single lexical item maps onto two (or more) semantic meanings. For example *bark* can refer to the sound a dog makes, or to the material on the outside of a tree trunk. Alternatively, a word can be lexically ambiguous in terms of having more than one syntactic category, as in the case of *produce* which can refer to a noun or a verb. In this chapter, and elsewhere in the thesis, the term "lexically ambiguous word" is used to refer to a word with two (or more) semantic meanings, but only one orthographic and one phonological from (that is, a homophonic homograph: Pacht & Rayner, 1993). Possessing two meanings but a single orthography and phonology makes lexically ambiguous words particularly interesting to experimentalists in the field of language processing, as orthographic and phonological processing, both of which are known to play an important role in lexical access (see Chapter 1, Sections 1.3.1 and 1.3.2), can be held constant while meaning is manipulated.

This chapter reports one experiment in which both the ambiguity of target words, and the alternative meanings of those words was manipulated experimentally in order to investigate whether well-documented adult preferences regarding the processing of ambiguous words also exist in child readers. The chapter will be structured as follows: first, Section 5.1 will review the relevant adult and child literature, although there have not been any studies which have used eye movements to investigate lexical ambiguity effects in child readers; second, Section 5.2 will describe the method used, in particular focusing on the extensive pre-screening procedures employed in order to ensure that ambiguous words were as well-controlled as possible for the participant populations and that the manipulation was maximally effective; third, Section 5.3 will report the results; and finally, Section 5.4 will draw conclusions from the data reported.

5.1.1 Research with adults

While previous research using cross-modal priming has shown that adults initially access both meanings of an ambiguous word (e.g. Swinney, 1979), research investigating lexical ambiguity effects during normal text reading has shown that whether both meanings are equally accessible depends on the nature of the ambiguous word. Some ambiguous words are *balanced* in that they have two salient meanings (and possibly other subordinate meanings) which are approximately equal in terms of usage; while most ambiguous words are *biased* in that they have one *dominant* meaning which is much more frequent than the other *subordinate* meaning(s).

Research has shown that adults initially fixate longer on balanced ambiguous words as compared to biased ambiguous words or control unambiguous words (Duffy et al., 1988; Rayner & Duffy, 1986; Rayner & Frazier, 1989). In addition, reading times in the disambiguating region of the sentence are longer when this region follows a biased ambiguous word than a balanced ambiguous word if the subordinate meaning of the biased word is contextually instantiated. These robust findings show that there is an immediate cost associated with processing a balanced ambiguous word because both meanings are maintained for post-lexical interpretation processes. In contrast, the cost associated with processing a biased ambiguous word is less immediate. Only the dominant meaning of a biased word is maintained available, thereby reducing initial reading times on biased as compared to balanced ambiguous words, as well as on unambiguous control words. It is only later in the sentence, when the meaning of the biased word is disambiguated in favour of the subordinate meaning, that the cost to processing is evident as the reader must re-access the alternative (subordinate) meaning of the ambiguous word.

In the case of the balanced ambiguous words, the reader has both meanings available for post-access processing, and so lexical selection is more efficient once the disambiguating region is encountered. Even if readers did maintain only one meaning of balanced words for post-access interpretation, they would select the correct meaning of the word on about 50% of occasions thereby weakening the ambiguity effect in the disambiguating region. Additionally, when readers need to reinterpret the ambiguous word, the speed with which the alternative meaning is accessed might be greater for balanced as compared to biased words as the second meaning of balanced words would usually be more frequent than the subordinate meaning of biased words.

However, the pattern of effects is very different if the ambiguous word is preceded by a disambiguating context. Duffy et al. (1988) found that when the preceding context supported one meaning of a balanced ambiguous word, the inflated fixation duration on the word disappeared. In contrast, when the preceding context was biased in favour of the subordinate meaning of a biased ambiguous word, fixation durations on the ambiguous word were inflated. This robust finding has been referred to as the *subordinate-bias effect* (Pacht & Rayner, 1993; Rayner, Pacht, & Duffy, 1994). Three types of model have since been offered with respect to the processing of ambiguous words following a biasing context. According to *autonomous models*, prior context has no effect on the access phase of lexical selection and both meanings of an ambiguous word are accessed in the same way as if there were no context. *Exhaustive access models* state that the disambiguating context increases the availability of the appropriate meaning of the word and in this way facilitates lexical access. Finally, *selective access* models propose that the disambiguating context constrains lexical access such that only the appropriate meaning of the ambiguous word is accessed.

Duffy et al. (1988) favoured a version of an exhaustive access model. They argued that meanings were exhaustively accessed following both a neutral and a disambiguating context, but that the context affected the order in which the meanings were accessed. Specifically, they put forward their *reordered access model*, according to which a prior disambiguating context increases the availability of the appropriate meaning without influencing the alternative meaning (in the absence of disambiguating context, alternative meanings simply become available in order of their meaning frequencies). This results in the appropriate meaning of balanced words becoming available earlier than the alternative meaning, and the appropriate meaning (when it is the subordinate meaning) of a biased word becoming available earlier than usual, perhaps simultaneously with the dominant meaning. However, importantly, they did not argue that only the contextually appropriate meaning was activated in the lexicon (as argued by proponents of selective access models).

In contrast to previous studies, Rayner and Frazier (1989) constructed their experimental sentences such that each ambiguous word served as its own control, thereby overcoming the predicament of whether to use the high or low frequency meaning of the ambiguous word as a control, or a combination of the two. Rayner and Frazier used contexts which disambiguated in favour of either the subordinate meaning or the dominant meaning of ambiguous words in order to allow direct comparisons between reading times in these two conditions. Crucially, they also manipulated the position of the disambiguating word, putting it either directly following the ambiguous word, or a few words downstream. As well as replicating results from Rayner & Duffy (1986) and Duffy et al. (1988), they also found that it made little difference if the disambiguation occurred immediately after the ambiguous word or a few words later, showing that lexical selection is immediate (see also Rayner, Cook et al., 2006). Rayner and Frazier also put forward their *integration model* in which they argued that both meanings of an ambiguous word are always automatically generated but if one meaning is successfully integrated into the sentence meaning, further selection processes are discontinued, leading to faster processing times. Although in some ways very similar to Duffy et al.'s reordered access model, the key difference is that Rayner and Frazier's model preserves modularity.

Sereno, Pacht and Rayner (1992) used both high and low frequency controls for biased ambiguous words, thereby controlling for both meanings. They predicted that if the selective access model was correct, there would be no difference between processing of sentences containing the low frequency control words and the subordinate meaning of the biased ambiguous words because the longer fixations observed on words disambiguated in favour of their subordinate meaning could be explained by their low frequency as compared to the dominant meaning. However, Sereno et al. did find differences between the low frequency controls and the ambiguous words and therefore interpreted their results as supportive of the reordered access model.

Dopkins, Morris and Rayner (1992) used positive and negative disambiguating contexts in order to test and discriminate between the reordered access model and the integration model. They found that gaze durations on ambiguous words were longer when those words were preceded by negative context (i.e. providing evidence that the dominant meaning of the ambiguous word should *not* be instantiated, rather than providing evidence that the subordinate meaning should be instantiated), as compared to a neutral context, and that reading times on the disambiguating material were inflated when words were preceded by a neutral context. Dopkins et al. interpreted their pattern of results as providing support for the reordered access model.

More recently, studies have investigated the effect of the strength of the biasing context (Binder & Rayner, 1998; Martin, Vu, Kellas, & Metcalf, 1999), the global discourse bias (Binder & Morris, 1995; Rayner et al., 1994), and there have also been controversies regarding how to control for the alternative meanings of ambiguous words which differ in frequency (Sereno et al., 2006; Sereno et al., 1992). However, for the purposes of this thesis, most of these more complex issues will be sidestepped and only basic ambiguity effects (i.e. in a neutral context) will be investigated. This is because lexical ambiguity effects have not yet been examined in children during normal text reading and it is important to establish whether the basic finding that participants look longer at ambiguous than unambiguous words can be observed in children as well as adults before investigating the effects of additional factors such as context.

Overall, the two main findings with regard to lexical ambiguity resolution in adults are that readers initially take longer to process balanced as compared to biased ambiguous words presented in a neutral context, but look longer at the disambiguating (to the subordinate meaning) information following a biased ambiguous word than a balanced ambiguous word.

Second, when presented following a disambiguating context, balanced words receive equally long fixations as compared to unambiguous words, but if the context disambiguates in favour of the subordinate meaning, biased words will receive longer fixations. The reordered access model provides a good explanation of the data available so-far. These robust findings show that only the dominant meaning of biased words is activated for lexical selection, but both meanings of balanced words are maintained. Experiment 4 will use the preferences that exist in the mature language processor, as observed in a neutral context, to examine whether they are also present in the developing language processor.

5.1.2 Research with children

Although there has been no research to date investigating children's online processing of lexically ambiguous words during natural reading, there has been much interest in the question of when and how children are able to learn words with more than one meaning, or rather to learn that a known word has a second (or third) meaning. It is usually assumed that the mental representation of a lexically ambiguous word consists of a single lexical representation of the word form, and two semantic representations, one for each meaning (Backscheider & Gelman, 1995). If this is the case, then on encountering a second meaning of an already familiar word, a child must create a second semantic representation of this novel referent and develop an association between an existing lexical representation of the known word form and this new semantic representation. This can be contrasted with the more common experience of learning a novel word in which a child must create a lexical representation of the novel word form, and additionally create a semantic representation of the novel referent and develop an association between these two newly created representations (Storkel & Maekawa, 2005).

While pre-school children have the metalinguistic ability to understand that one word may have two distinct meanings (Backscheider & Gelman, 1995; Peters & Zaidel, 1980), it has been shown that children can have difficulty learning words when there is not a one-to-one mapping between a word and its meaning (Markman, 1989), for example, up to age ten, children are relatively unwilling to attach new meanings to known words even when the context demands it (Doherty, 2004; Mazzocco, 1997). Furthermore, even when children are familiar with both meanings of an ambiguous word, there is evidence to show that young children (aged 3-5 years) tend to interpret them in terms of the more frequent meaning, even when this meaning is inappropriate in the context of the texts in which they are presented. Importantly this tendency has been shown to decrease with age (from three to five years, Campbell & Bowe, 1982).

Beveridge and Marsh (1991) found that children aged three to six years preferred the dominant meaning of ambiguous words (which were homophonic but not homographic, e.g. *bear/bare*) before listening to a story which contained those same words in which the subordinate meaning was contextually instantiated. First, children were asked to point to a picture which showed the (homophonic) orally presented word. Children viewed four possibilities, two of which represented alternative meanings of the ambiguous word. They then listened to a story which differed in the richness of contextual information which gave the correct meaning of target words. Following the story, although children preferred the subordinate meaning numerically more than before the story, this difference was not reliable. However, there were several problems with this study: there was no pre-screen procedure to ensure that children knew both meanings of the ambiguous words (and they most likely didn't as many target words were quite difficult such as *buoy* and *quay*); target words were not controlled for length, frequency, or any other possible confounding variable; and no pre-screen procedure was carried out to examine which meaning of the ambiguous word was preferred for children of each age group (the experimenters simply guessed). Given these experimental shortcomings, it is quite possible that very different results may be observed when materials are properly controlled.

In two priming studies particularly relevant to the experiment reported in this chapter, Simpson and Foster (1986) examined children's processing of ambiguous words by asking them to name target words which were primed by ambiguous words. The target words were related to either the subordinate or the dominant meaning of the ambiguous words. In their first experiment, they found that while younger children (ages 7 - 10) exhaustively accessed and maintained both meanings (i.e. target words related to both dominant and subordinate meanings of an ambiguous word were primed), older children (age 12), like adults, maintained the more frequent meaning but not the subordinate meaning (i.e. only target words related to the dominant meaning were primed). In their second experiment, Simpson and Foster varied the amount of time that elapsed

(150ms, 300ms or 750 ms) between the presentation of the prime and that of the target. They found that older children, as well as younger children, initially activated both meanings of the ambiguous words, and the effect of meaning dominance was observed only after 750ms, much like adults have been shown to do in reading studies (Duffy et al., 1988; Rayner & Duffy, 1986).

The research conducted thus far which has investigated children's processing of lexically ambiguous words is somewhat conflicting. While young children do appear to have an understanding that words can have more than one meaning, the way in which children process words with two meanings is not yet clear. There is some evidence to show that younger children use only the dominant meaning of ambiguous words (in a neutral or null context), although the studies which have shown this (Beveridge & Marsh, 1991; Campbell & Bowe, 1982) have been somewhat problematic. In contrast, the most well-controlled study to investigate this issue (Simpson & Foster, 1986) showed that while younger children exhaustively access and maintain both meanings of ambiguous words, older children (age 12), like adults, do not maintain both meanings for post-lexical interpretation for more than 750ms.

For Experiment 4, children were split into two age groups: younger children (ages 6.5 - 8.9 years) and older children (ages 9.2 – 12.0 years) in order to examine whether, as the research outlined above suggests, the mechanism that children have in place for processing words with more than one meaning develops with age. While the older children in the experiment reported in this chapter were not as old as those in the Simpson & Foster study, they were, almost without exception, extremely precocious readers with a mean reading age of 13.3 years (a measure of reading ability was not taken in the Simpson and Foster study). Following Simpson and Foster's study, and given that fixation time is typically much shorter than the naming time response measure used in their study (Pacht & Rayner, 1993), it is possible that monitoring the eye movements of the older group of children in this study might be sensitive enough to reveal their initial processing preferences, whereby they are expected to access and maintain for post-access interpretation both meanings of ambiguous words. However, it may be that a different pattern of effects, for either group, or both groups, of children is observed in text reading rather than a simple word naming task.

It is known that adults look longer at ambiguous than unambiguous words. It is further known that adults look longer at balanced than biased ambiguous words in a neutral preceding context. The stimuli used in Experiment 4 were constructed in order to investigate whether children of different ages exhibited these same robust effects as adults, and in order that direct comparisons could be made between the eye movement behaviour of adults and children as they read the same sentences containing the same linguistic manipulation. In Experiment 4, pairs of experimental sentences contained ambiguous and unambiguous control words (see sentences 1a-1b below). Sentence (1a) contained the subordinate meaning of the ambiguous word (if biased) and sentence (1b) contained an unambiguous control word which was matched to sentence (1a). Sentences (1a) and (1b) were identical apart from the target word (*straw* or *chick*).

(1a) The little boy played with the straw inside the barn yesterday.

(1b) The little boy played with the chick inside the barn yesterday.

5.1.3 Predictions

Adults. It was predicted, in line with previous research, that adults would exhibit longer fixations on ambiguous than unambiguous control words. Furthermore, adults would look longer at balanced ambiguous words than biased or control words during first pass, but would look longer at the disambiguating region following a biased word than a balanced or control word. It was anticipated that a primary result would be the replication of the previous well-documented adult findings.

Children. Following the findings from previous research on children's processing of ambiguous words, it was predicted that both older and younger children would exhibit an effect of ambiguity, showing longer reading times on ambiguous than unambiguous words (as adults do). However, unlike adults, and following Simpson and Foster (1986), it was anticipated that younger and older children would maintain both meanings of all (i.e. biased and balanced) ambiguous words for post-access interpretation. This would mean that children would exhibit no effect of word bias. Furthermore, an interaction would be anticipated, whereby children would exhibit longer fixation durations on biased ambiguous words as compared to adults due to the increased processing demands of maintaining two meanings simultaneously. Furthermore, while adults would be expected to exhibit longer fixation durations on balanced as compared to biased ambiguous words, younger and older children would not be expected to show any difference, as both meanings would be maintained in all cases.

It was also predicted that older children would not maintain the subordinate meaning of biased ambiguous words for more than 750ms, and so inflated second pass reading times might be observed on the disambiguating region following biased, but not balanced, ambiguous words (or control words). This might occur because with greater reading experience older children may have developed this adult-like strategy of dealing efficiently with biased ambiguous words, although the time course in which this strategy was exhibited was predicted to be delayed in older children relative to adults. In contrast, younger children would not be expected to exhibit any reading time differences in the disambiguating region between balanced and biased conditions as they would have maintained both meanings of both biased and balanced ambiguous words and so, unlike the older children and adults, would have the subordinate meaning readily available when the disambiguating information instantiated it as the appropriate meaning.

5.2 Method

5.2.1 Participants

28 adults and 36 children took part in this experiment. Participants were recruited as in Experiment 1 (Chapter 3, Section 3.2.1). Of the 36 children, 18 were categorised as 'younger' and 18 were categorised as 'older'. The mean age in the younger group was 7.9 years (range 6.5 – 8.9 years), and the mean age in the older group was 10.4 years (range = 9.2 – 11.7 years). As in all the experiments in the thesis, children completed the WORD (Rust, Golombok, & Trickey, 1993) test for reading ability. Children in the younger group had a mean reading age of 10.5 years (range = 6.8 – 14.7 years), and children in the older group had a mean reading age of

13.3 (range = 8.6 – 17.0 years). As noted, both groups of children were extremely precocious in their reading ability.

5.2.2 Materials

Pre-screen procedure. Before the main experiment, an extensive pre-screen procedure was undertaken. The pre-screen procedure was carried out as two separate tasks. Part 1 was administered to children only, and was used to ensure that children were familiar with both meanings of the target ambiguous words to be used in the main experiment. Part 2 was administered to both adults and children and was used to assess which ambiguous words were biased and which were balanced for each age group, and which meanings were dominant and subordinate for the biased words.

Part 1. First, 41 ambiguous words, which were expected to be familiar to children as young as six years old, were chosen. For each word, two pictures representing each word meaning plus two pictures of unrelated items were selected. A computer programme was designed using Macromedia Flash MX software so that for each word, a screen appeared with that word printed at the top together with the four pictures (see Figure 5.1 below for an example).

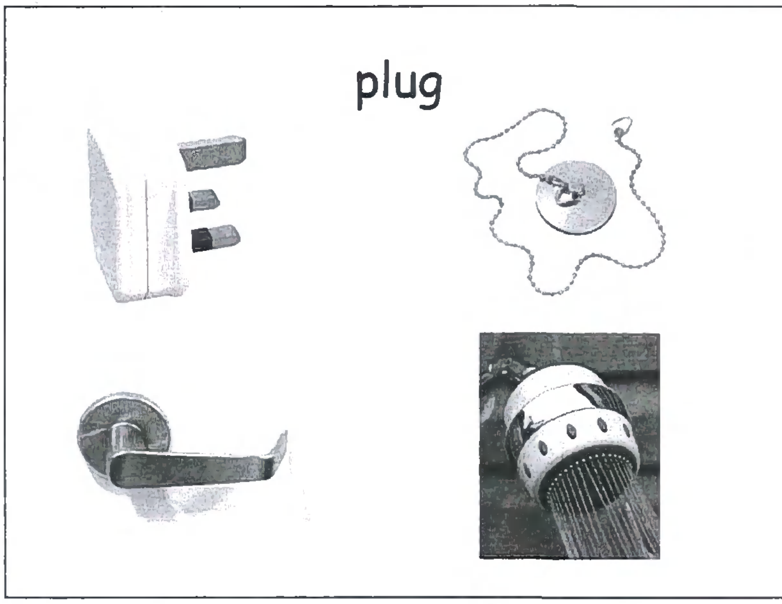

Figure 5-1: Sample item (*plug*) from pre-screen procedure.

Thirty four children in total completed the first pre-screen procedure. Children were recruited from local primary schools but importantly did not take part in the main eye tracking experiment. Of these children, sixteen were aged 7-8 years, and sixteen were aged 10-11 years. The screening procedure took place in schools. Instructions were first given verbally. Children were told that they would see a screen with a word and four pictures. They were told to read the word and then use a computer mouse to click on any pictures that matched the word. They were told to look for two pictures but to click on only those pictures which they thought matched the word. Note that for 10 of the 41 items, there were three meanings represented pictorially for the ambiguous word. In these cases, if a child selected only two of the three pictures the answer was judged to be correct as the child was aware that the word had two meanings. Note also that only five of these were used in the main experiment. Children read written instructions and were given the opportunity to ask any questions. They then completed a practice trial. If they did not answer correctly on the practice trial, the correct answer was given verbally and explained to them. Children then completed the procedure in one session which lasted between five and ten minutes. Children were free to ask questions relating to the words and pictures throughout the session.

Results from this first pre-screen procedure resulted in the exclusion of 13 items for which fewer that 65% of children knew both word meanings. Of the remaining 28 items, for the children aged 10-11, all children knew both meanings of a minimum of 88% of words, with children knowing 96% of both meanings on average. In the 7-8 year old children, performance was slightly lower; all children knew both meanings of at least 69% of words, with children knowing both meanings of 82% of words on average.

Part 2. Once the 28 items had been selected, the second pre-screen procedure was designed to assess whether target ambiguous words were balanced or biased, and if biased, which meaning was dominant for each age group. The same group of children took part in this second pre-screen procedure, as well as 30 adults who were students at Durham University. In total 82 children took part: of these 50 children were aged 7-8 years, and 32 were aged 10-11 years. Testing took place in schools and children were supervised by their teacher who answered any questions they had and helped children with writing if necessary.

Participants were required to complete a written questionnaire in which each target word was printed. Participants were asked to do two things: (1) write the first associated word which occurred to them on reading the target word; and (2) write a sentence containing the target word. They were given the following written instructions: "Below you will see some words. There are two lines after each word. On the first line, write the first word that comes into your head when you read the word. On the second, longer line, write a sentence with the first word in it." Children were given instructions orally in addition to the written instructions. This procedure has been used in both adult (Duffy et al., 1988; Rayner & Frazier, 1989) and child studies (Beveridge & Marsh, 1991) and is known to elicit preferred meanings of ambiguous words (in adults) effectively.

Following Rayner and Duffy (1986) and Duffy et al. (1988), words were categorised as balanced if the dominant meaning had a probability between 0.47 and 0.69 (that is, that meaning was rated as the preferred one in between 47% and 69% of cases). Words were categorised as biased if the dominant meaning had a probability between .7 and 1.0 (i.e. the dominant meaning was preferred at least 70% of the time). For adults, this meant that 2 items were uncategorisable (in both cases because the word had three meanings all of which were preferred in fewer than 47% of cases), 14 words were categorised as biased and 12 were categorised as balanced. The mean dominant probability for the balanced words was .59 (range = .48-.67), and the mean dominant probability for the biased words was .84 (range = .73-1.0). For older children, this meant that 18 items were categorised as biased, and 10 items were categorised as balanced. The mean dominant probability for the balanced words was .59 (range = .50-.69), and the mean dominant probability for the biased words was .83 (range = .71-.97). Finally, for the younger children, 1 item was excluded because it could not be categorised as balanced or biased, 19 items were categorised as biased, and 8 items were categorised as balanced. The mean dominant probability for the balanced words was .57 (range = .48-.65), and the mean dominant probability for the biased words was .89 (range = .70-1.0).

Experimental sentences were then constructed and rated for predictability by 15 adults who did not take part in the main experiment (the same procedure as that outlined in Chapter 3, Section 3.2.2 was used). Sentences were classed as predictable if 70% of respondents completed the sentence using the target word. This criterion resulted in the exclusion of five of the sentences, for which a minimum of 11/15 adults completed the sentences with the target word. This left 23 sentences which were used in the main experiment. For those remaining 23 items, there was no difference in predictability between the ambiguous (mean = 0.12; SD = 0.15) and the control (mean = 0.12; SD = 0.19) words. Children did not complete a cloze test and therefore adult ratings of predictability were used for all age groups. While this was clearly far from ideal, during testing in schools, both children and teachers were unhappy about completing an additional test (as they had already completed several pre-screen procedures for other experiments), and for this reason only adults completed this part of the pre-screen procedure.

For each ambiguous word, two versions of a sentence were constructed (see sentences 1a-1b, Section 5.1.3). As outlined earlier, sentence (1a) contained the subordinate meaning of the ambiguous word (if biased) and sentence (1b) contained an unambiguous control word which was matched for frequency and length to sentence (1a). Sentences (1a) and (1b) were identical apart from the target word (*straw* or *chick*). Ambiguous words were disambiguated in the post-target region (*the barn*). At least one word always followed this disambiguating region. Using both adult and child corpora, there was no difference in frequency between target ambiguous words and control words Using an adult corpus (Baayen et al., 1995), the mean frequency for the ambiguous words was 58 counts in a million (SD = 58), and for the control words, it was 33 counts in a million (SD = 57), $p > .15$. Using a child corpus (Stuart et al., 2002), the mean frequency for the ambiguous words was 104 counts in a million (SD = 137), and for the control words, it was 59 counts in a million (SD = 61), $p > .2$. Target words were identical in length across conditions.

5.2.3 Apparatus

The apparatus used was identical to that used in Experiment 2 (Chapter 4, Section 4.2.1).

5.2.4 Procedure

The procedure for the main eye-tracking experiment was identical to that in Experiment 2 (Chapter 4, Section 4.2.1). Of the 28 experimental sentences (due to human error the predictable sentences were excluded only after the experiment had already been run so participants read all 28 sentences), seven were followed by comprehension questions. In addition, 29 of the 36 children who took part in the main experiment also completed the test which was used for Part 1 of the pre-screen procedure outlined above (Section 5.2.2) to ensure that participants knew both meanings of the ambiguous target words in the experiment (unfortunately the remaining seven children were too tired or else did not have time to complete this test). 97% of items were answered correctly (i.e. both meanings of ambiguous words were correctly identified) with no less than 80% of children knowing both meanings of any one ambiguous word. Note that children completed the task after completing the main experiment so as not to influence their reading of the ambiguous words while their eye movements were monitored.

5.3 Results

All participants scored very highly on the comprehensions questions. Adults answered questions correctly on 99.5 % of occasions, and children answered correctly on 95% of occasions. All participants scored higher than 70%. Fixations longer than 1200ms and shorter than 80ms were systematically excluded from the data set. Trials in which there was tracker loss or many blinks were also excluded. In addition, any trial in which the participant did not fixate either the target region or the post-target region (*straw* or *inside* in sentence 1a below) was deleted. In total, this resulted in an exclusion of 13% of data for adults and 13% of data for children.

	REGION 1	REGION 2	REGION 3	REGION 4	REGION 5	REGION 6
(1a)	The little boy played with/	the /	straw /	inside /	the barn /	yesterday. /

In order to give a clear structure to the somewhat complicated results, the analyses will be reported in four sections. The first section will address whether a basic first pass ambiguity effect exists in the data across all age groups. The second, third and fourth sections will present analyses for each group (adults, older children, and younger children) separately, in order that group meaning preferences (i.e. whether words are balanced or biased, and whether meanings are dominant or subordinate for biased words for each participant group) are taken into account.

5.3.1 Overall ambiguity effect

First all 23 items were included in an analysis to examine whether there was an overall cost associated with processing an ambiguous as compared to an unambiguous control word during first pass reading. Because ambiguous words were categorised differently for each group as to whether they were balanced or biased, this variable was not entered into these initial analyses. Furthermore, due to human error, the design was not a fully-counterbalanced Latin Squares design. For this reason, an additional dummy variable (file) was entered into the analyses in order to control for which version of each experimental sentence each participant read. No effects of this dummy variable were reliable across both participants and items analyses in any measure and therefore they will not be reported in any of the subsequent analyses as they are of no theoretical interest. First fixation durations, single fixation durations, gaze durations and spillover durations (the duration of the first fixation following the fixation(s) on the target word) were all calculated (see Chapter 3, Section 3.3.3 for a description of the other measures).

Table 5-1: Mean first pass reading times (ms) on ambiguous and unambiguous target words for adults, older children and younger children (Standard deviations in parentheses).

	Adults		Older children		Younger children	
	Ambiguous	Control	Ambiguous	Control	Ambiguous	Control
First fixation duration	233 (*77*)	221 (*60*)	289 (*139*)	285 (*122*)	276 (*141*)	296 (*172*)
Single fixation duration	237 (*80*)	222 (*60*)	310 (*167*)	246 (*88*)	278 (*152*)	264 (*144*)
Gaze duration	257 (*105*)	232 (*76*)	392 (*254*)	296 (*182*)	298 (*162*)	280 (*145*)
Spillover duration	232 (*62*)	231 (*95*)	258 (*139*)	246 (*116*)	262 (*175*)	246 (*107*)

Table 5.1 shows first pass reading times for all groups of participants on the ambiguous and unambiguous target words. Because the design was somewhat complicated, it will be described in some detail here. There were three independent variables: age group, ambiguity and (the dummy variable) file. Age group was a between participants, but within items, variable; ambiguity was a within participants, and within items, variable; and file was a between participants, and between items, variable.

A 3 (group: younger children, older children, adults) x 2 (ambiguity: ambiguous, unambiguous control) x 4 (file) mixed design ANOVA showed no effect of ambiguity on first fixation durations ($ps > .7$). There was an effect of group, $F1\ (2, 52) = 6.21$, $p < .005$; $F2\ (2, 36) = 15.24$, $p < .001$. Pairwise analyses showed that adults made reliably longer first fixations than older children, $t1\ (21.8) = 3.78$, $p = .001$; $t2\ (21) = 5.95$, $p < .001$, and reliably longer first fixations than younger children, $t1\ (20.2) = 2.88$, $p < .01$; $t2\ (21) = 5.24$, $p < .001$, but there was not a reliable difference in first fixation durations between older and younger children ($ps > .8$). Note that there was not equal variance in the adult group as compared to the child groups ($ps > .05$ in Levene's test for Equality of Variance) and so the corrected t values are reported here. There were no reliable interactions (all ps > .1). This pattern of differences in fixation durations across age groups was observed across most measures and regions. Because basic oculomotor differences between adults and children independent of the linguistic manipulation employed are not the focus of this (or any other) chapter, and because such differences are already well-documented, further analyses which show the above pattern will not be reported in full.

In single fixation durations, there was no effect of group, although the effect was reliable by items, F1 (2, 50) = 2.37, p = .10; F2 (2, 30) = 4.06, p < .05: adults made numerically longer single fixations than both groups of children. There was no effect of ambiguity, which was reliable by items, but not by participants, F1 (1, 50) = 1.92, p = .17; F2 (1, 15) = 5.71, p < .05. There was no interaction between group and ambiguity (ps > .5).

In gaze durations, there was an effect of ambiguity which was reliable by participants and marginal by items, F1 (1, 52) = 8.45, p < .01; F2 (1, 36) = 3.55, p = .076. Although narrowly failing to reach significance, this important result suggests that there is a cost associated with processing a word which has two (or more) meanings as compared to a word with only one meaning, and that this holds for adults (25ms effect), older children (96ms effect), and younger children (18ms effect). There was also an effect of group, F1 (2, 52) = 10.23, p < .001; F2 (2, 36) = 14.98, p < .001, with adults showing longer gaze durations than both groups of children (all ps < .01), but no reliable difference between the two groups of children (ps > .05). There was no reliable interaction (ps > .2). Finally, there were no reliable effects in spillover durations (all ps > .05).

These results show tentative evidence that for children, as well as adults, there is a cost associated with processing ambiguous words as compared to unambiguous control words. Although the effect was not reliable by items, it may be that a real effect is masked by differences in reading time on balanced and biased words for each group. Because there were differences between age groups in their categorisation of ambiguous words as biased or balanced, it was necessary to conduct analyses on each group separately in order to explore these predicted differences.

Section 5.3.2 therefore examines ambiguity effects in adults, Section 5.3.3 examines effects in older children, and Section 5.3.4 examines those same effects in younger children.

5.3.2 Adults

Following Rayner and Duffy (1986) and Duffy et al. (1988), it was expected that a difference in reading times between ambiguous and control words would be present only on balanced ambiguous words for which both meanings were maintained. For this set of analyses, target words were categorised as balanced or biased (as per the criteria outlined in Section 5.2.2), resulting in eight balanced words, and 14 biased words.

Table 5.2 shows mean first pass reading times on balanced and biased ambiguous words and their controls. As well as the measures which were calculated in the overall analyses, the proportion of first pass regressions, go past time (the duration of all fixations from first fixating a region to leaving it to the right), and total reading time were all calculated for the disambiguating region.

Table 5-2: Reading time measures (ms) and regression probabilities in the target region and disambiguating region for adults for balanced and biased ambiguous words and their controls (Standard deviations in parentheses).

		Balanced		Biased	
		Ambiguous	Control	Ambiguous	Control
Target word (Region 3)	First fixation duration	232 (*90*)	229 (*60*)	237 (*72*)	218 (*61*)
	Single fixation duration	238 (*94*)	227 (*61*)	239 (*74*)	219 (*61*)
	Gaze duration	263 (*117*)	250 (*88*)	257 (*99*)	221 (*64*)
	Spillover duration	232 (*54*)	231 (*77*)	228 (*65*)	237 (*110*)
Disambiguating region (Region 5)	First fixation duration	257 (*103*)	236 (*105*)	252 (*78*)	242 (*75*)
	Single fixation duration	287 (*161*)	225 (*70*)	268 (*91*)	257 (*88*)
	Gaze duration	574 (*353*)	573 (*293*)	403 (*188*)	353 (*152*)
	First pass regressions out	0.19(*0.18*)	0.08 (*0.07*)	0.08 (*0.06*)	0.07 (*0.07*)
	Go past time	719 (*497*)	605 (*319*)	506 (*449*)	409 (*261*)
	Total reading time	585 (*391*)	553 (*304*)	507 (*244*)	442 (*279*)

Separate 2 (ambiguity: ambiguous, control) x 4 (file) mixed design ANOVAs were conducted for balanced and biased words. In the target region, there was no effect of ambiguity in first fixation durations, single fixation durations, gaze durations, or spillover durations (all $ps > .1$) for balanced words. For biased words, there were no reliable effects of ambiguity in first fixations, $F1\ (1, 21) = 8.75$, $p < .01$; $F2 < 1$, $p > .3$, or single fixation durations, $F1\ (1, 20) = 6.79$, $p < .05$; $F2 < 1.5$, $p > .3$. However, there was a reliable effect of ambiguity for biased words in gaze durations, $F1\ (1, 12) = 12.38$, $p < .005$; $F2\ (1, 10) = 5.08$, $p < .05$, with adults showing longer gaze durations on ambiguous than control words. There were no effects of ambiguity in spillover durations ($ps > .3$) These results show that, contrary to predictions, there was a cost associated with processing biased, but not balanced, ambiguous words as compared to unambiguous words for adults.

Fixation durations on the disambiguating region (Region 5) were then examined (see Table 5-2). There were no reliable effects of ambiguity in first fixation durations for balanced or biased words (all $ps > .1$). It was not possible to carry out analyses for single fixation durations as there were too few data points. In gaze durations, there was no effect of ambiguity for balanced words ($ps > .8$), but there was an effect of ambiguity for biased words, $F1\ (1, 17) = 5.06$, $p < .05$; $F2\ (1, 10) = 8.87$, $p < .05$. There were no effects of ambiguity in the probability of making a first pass regression out of the disambiguating region, in go past time, or in total time for biased or balanced words (all $ps > .05$.

Overall, the expected pattern of effects was not observed, and indeed it appeared that adults needed longer to process biased ambiguous words than balanced ambiguous words, in contradiction of previous research. The failure to replicate robust lexical ambiguity effects in adult readers may have been due to the inaccurate classification of target words as balanced or biased, or due to the restrictions inherent in creating experimental sentences suitable for use with both adults and children. These, and alternative explanations for the lack of significant effects will be commented on further in the Discussion section.

5.3.3 Older children

The data from 18 older children (aged 9-12 years) were then analysed. For the older children, using the same criteria as those used for the adults, of the 23 items, eight items were classed as balanced and 15 items were classed as biased. Table 5-3 below shows the reading times and regression probabilities on the target word and the disambiguating region for both balanced and biased, ambiguous and control conditions. Because older children made very few single fixations on the critical regions, it was not possible to conduct statistical analyses for this measure.

Table 5-3: Reading time measures (ms) and regression probabilities in the target region and disambiguating region for older children (Standard deviations in parentheses).

		Balanced		Biased	
		Ambiguous	Control	Ambiguous	Control
Target word	First fixation duration	286 (*136*)	269 (*94*)	291 (*141*)	293 (*135*)
	Gaze duration	435 (*237*)	303 (*116*)	370 (*243*)	292 (*209*)
	Spillover duration	251 (*99*)	225 (*72*)	262 (*154*)	256 (*131*)
Disambiguating region	First fixation duration	267 (*81*)	255 (93)	279 (*112*)	241 (*90*)
	Gaze duration	489 (*246*)	528 (*346*)	493 (*420*)	473 (*273*)
	First pass regressions	0.23 (*0.21*)	0.27 (*0.38*)	0.31 (*0.25*)	0.30 (*0.25*)
	Go past time	862 (*823*)	862 (*870*)	901 (*791*)	731 (*428*)
	Total reading time	862 (*669*)	769 (*439*)	835 (*690*)	765 (*456*)

As with the adults, separate analyses were conducted for balanced and biased words. A 2 x 4 mixed design ANOVA showed no reliable effects of ambiguity in first fixation durations for the older children (ps > .2) for biased or balanced words. Likewise, in gaze durations, there were no reliable effects of ambiguity, (all ps ≥ 0.05). There was, however, an effect of ambiguity in spillover durations for balanced words, , $F1\ (1, 11) = 8.70$, $p < .05$; $F2\ (1, 4) = 7.8$, $p < .05$, but no such effect for biased words (ps > .8). Overall, the data from the target word region suggest that older children require longer processing times for balanced, but not biased, ambiguous words than control words.

In the disambiguating region, there was a reliable effect of ambiguity in first fixation durations on biased words, $F1\ (1,\ 14) = 6.54$, $p < .05$; $F2\ (1,\ 11) = 5.65$, $p < .05$, but not on balanced words ($ps > .5$). However, this trend was not observed in gaze durations, number of first pass regressions made out of a region, go past times, or total times on either balanced or biased wordss ($ps > .05$).

Overall then, the data suggest that older children access and maintain both meanings balanced but not biased ambiguous words for post-lexical selection. That is, they preferentially select the dominant meaning of a biased ambiguous word during reading, and consequently exhibit disruption on encountering the disambiguating region, which requires that they instantiate the subordinate meaning of the ambiguous word. However, these effects are not observed consistenetly across measures, and in the light of opposite pattern of effects in adults, very firm conclusions cannot be drawn from these effects and they should be regarded as merely suggestive.

5.3.4 Younger children

Finally, the data from 18 younger children (aged 6.5-8 years) were analysed. For the younger children, seven of the target ambiguous words were categorised as balanced, one target word could not be categorised as either balanced or biased and 15 ambiguous target words were categorised as biased. Table 5-4 below shows reading times and regression probabilities for the 22 items which could be categorised. As with the older children, single fixation durations are not reported for the younger children because for these analyses there were insufficient data points.

Table 5-4: Reading time measures (ms) and regression probabilities in the target region and disambiguating region for younger children (Standard deviations in parentheses).

		Balanced		Biased	
		Ambiguous	Control	Ambiguous	Control
Target word (Region 3)	First fixation duration	300 (*176*)	306 (*211*)	272 (*127*)	292 (*159*)
	Gaze duration	269 (*94*)	277 (*140*)	307 (*181*)	287 (*150*)
	Spillover duration	285 (*191*)	263(*116*)	250 (*171*)	237 (*105*)
Disambiguating region (Region 5)	First fixation duration	264 (*126*)	286 (*136*)	295 (*132*)	296 (*113*)
	Gaze duration	507 (*292*)	609 (*319*)	570 (*397*)	532 (*387*)
	First pass regressions	0.29 (*0.26*)	0.18 (*0.28*)	0.18 (*0.24*)	0.23 (*0.20*)
	Go past time	774 (*628*)	698 (*349*)	724 (*439*)	768 (*533*)
	Total reading time	700 (*409*)	752 (*379*)	821 (*562*)	825 (*579*)

Younger children showed no effects of ambiguity in their first fixation durations, gaze durations, spillover durations for biased or balanced words (all ps > .1). Furthermore, there were no effects of ambiguity in the disambiguating region (all ps > .2). While it should be noted that there were many missing data in these analyses, in part because so few items were categorised as balanced, overall the analyses for the younger children show no evidence that children require longer to process ambiguous than unambiguous words, either for words which are balanced or biased.

To summarise the results from the lexical ambiguity experiment, when analysed together, there was a marginal overall ambiguity effect, indicating that there was a small initial cost to processing for ambiguous than unambiguous words for all groups. However, when analysed separately, neither adults nor younger children showed any robust effects of ambiguity or bias. In contrast, the data from older children suggest that they access and maintain both meanings of biased and balanced ambiguous words for post-access selection. No disruption to processing was observed in the disambiguating region in which the subordinate meaning of the ambiguous word was instantiated for older children, in accordance with the argument that they had both meanings available for processing at this point in the sentence. Overall, it is likely that when all participants and items are combined, the added power in the analyses produces a (marginal) effect of ambiguity. However, when groups and items are broken down in order to carry out subanalyses, the resulting lack of power means that robust effects are not observed in two of the three age groups.

5.4 Discussion

Experiment 4 investigated adults' and children's processing of lexically ambiguous words. It was predicted that there would be a cost associated with processing a lexically ambiguous word as compared to an unambiguous control word for adults and children. While there was a marginal effect of ambiguity in the overall analyses, the effect did not hold when groups were analysed separately. However, although robust ambiguity effects in the predicted direction were not observed in adults and younger children, older children did exhibit reliably longer gaze durations on balanced ambiguous than unambiguous words. This result provides tentative evidence that, for older children, there is a cost associated with processing a word which has two equally frequent meanings rather than just one, although the effects should be treated with caution in the context of the pattern of effects in adults. Nevertheless, the finding suggests that accessing two meanings from the lexicon, and keeping them available for post-lexical interpretation, places additional processing resources on the developing lexical processor and this results in children aged approximately 9-12 years old, making longer fixation durations as they read such words.

Second, it was predicted that fixations would be longer on balanced than biased ambiguous (and unambiguous control) words for adults but not for children. Unfortunately this prediction was not met and adults exhibited longer first pass reading times on biased, but not balanced, ambiguous words as compared to unambiguous control words. This may be due to an inaccurate classification of target words as balanced or biased (see Section 5.4.1 below), or it may be due to a lack of power in the separate adult analyses as those words categorised as balanced were fewer than those categorised as biased (for all groups). It was predicted that children would not fixate balanced words longer than biased words, but in fact older children did exhibit longer spillover durations on balanced than control words, as was predicted for the adult group This effect suggests that, contrary to predictions, older children maintain both meanings of balanced, but not biased words, as is observed in adult readers (in previous studies) who maintain only the dominant meaning of biased words.

Third, it was predicted that adults, but not children, would exhibit longer reading times on the disambiguating regions of sentences following biased than balanced ambiguous words. While this prediction was met, it failed to have much significance in light of the pattern of effects in the target region. In relation to this point, it was also predicted that older children would eventually discard the subordinate meaning of biased words, while younger children would maintain both meanings resulting in a differential pattern of effects. This prediction was partially met. The older children showed a reliable effect of ambiguity in the post-target region, indicating that they may have maintained both meanings of biased words for a little longer than adults (in previous studies) who generally exhibit ambiguity effects on the target word itself. Younger children, however, exhibited no effects ambiguity, showing that they were not affected by the manipulation employed.

It was predicted that the magnitude of ambiguity effects would decrease with age; that is, younger children would exhibit stronger effects of ambiguity than older children, and older children would exhibit stronger effects than adults. This prediction was not met. As there were no reliable effects in the adults (in the predicted direction) or the younger children, it is of course unwise to make comparisons between a non-significant and a significant effect. It was important to replicate robust ambiguity effects in adult readers in order to legitimise the stimuli used in the current experiment for investigating lexical ambiguity effects in children. The failure to obtain reliable effects may have been due to flaws in the experimental design; for example, the inadequate categorisation of words as balanced or biased. An alternative explanation is that because the manipulation employed typically generates only small effects, the stimuli used, which were relatively easy to read for experienced adults, were not sensitive enough to produce these effects reliably in the adult population. These issues will be discussed further in the following section.

5.4.1 Limitations of experimental manipulation

Despite the tentative conclusions drawn from the data, as mentioned, the data from the child analyses must be treated with extreme caution. The failure to replicate well-documented effects in adults must not be overlooked, and a shortcoming of Experiment 4 is that the manipulation employed was not sufficiently good to generate these robust effects in adult readers. Although many participants were tested, and there were a good number of items, when the groups were analysed separately there was not enough power to produce statistically reliable effects.

However, the failure to produce reliable effects in adults is potentially informative regarding those experimental manipulations which are appropriate in developmental research, and indeed whether any aspect of the experimental manipulation was flawed leading to smaller effects than were possible. While Experiments 2 and 3 examining word frequency (and those in the following chapters examining syntactic ambiguity and thematic plausibility) produced statistically reliable effects, this may be, in part, because the magnitude of these effects is relatively large in the adult population (e.g. there was a 37ms effect of frequency in Rayner & Duffy, 1986). In contrast, lexical ambiguity effects are more subtle (e.g. there was a 17ms effect of ambiguity on balanced words in that same study). Note that adults actually showed a 25ms effect of ambiguity in Experiment 4 but with the increased variability introduced by including children in the analyses, together with the restrictions placed on any linguistic manipulation employed for use with a developmental population (e.g. that words must be acquired relatively early, syntax must be simple, semantic meaning must not be very complex), statistically significant online effects of lexical ambiguity during reading were not observed. Some of the possible experimental limitations, as well as wider issues concerning conducting developmental research in language processing, will be discussed in this section.

The failure to fully replicate robust findings in the adult literature is disappointing but there are several possible explanations for why this may have happened, and why robust effects were not observed in the children either. One possibility is that although target words were controlled for predictability, ratings were taken from adults only, and it may be that children have different perceptions of what is predictable. However, it is the lack of reliable effects in adults rather than children that is of primary concern, especially as ambiguity effects were stronger in (older) children than adults.

Another issue which may have relevance to the lack of effects is that of the categorisation of ambiguous words as balanced or biased. While the pre-screening procedure was extremely thorough, how balanced or biased an ambiguous word is, is almost always a matter of degree (there was only one case in which all adults used only one meaning of an ambiguous word in the pre-screen procedure, no cases in which older children used only one meaning, and only three cases in which younger children used only one meaning). It may be then that a better way to grade meaning bias would be on a scale rather than by category. Importantly, adults showed reliable longer gaze durations on biased ambiguous words than control words, but not on balanced amabiguous words than control words. Given the well-documented pattern of effects in adult readers, these surprising results do indicate that the pre-screen procedure or the categorisation criteria did not result in accurate classification of ambiguous words as biased or balanced. Nevertheless, it should be noted that the screening procedure for the experimental stimuli was extremely rigorous. Although the manipulation did not generate robust effects, great care was taken to ensure that the target words, and overall sentences, were well-controlled, and confounding variables were minimised.

A final, and arguably the most important, factor which may have influenced the magnitude of effects in the adult participants was that of age-appropriateness of the linguistic stimuli. As in all the experiments reported in this thesis (with the exception of Experiment 2; see Chapter 4), experimental sentences were identical for both adults and children in order that direct comparisons could be made between groups. However, this necessarily means that the stimuli were very easy for adults to read. While having identical stimuli is an important premise of the thesis, the relative ease with which adults process the sentences may have attenuated the effects produced by the experimental manipulation employed. For example, lexically ambiguous words in the current experiment were short, relatively high frequency and were acquired relatively early in life. In addition they were embedded in syntactically and semantically simple sentences. These factors would decrease the frequency, and the duration, of fixations on those words (and their controls) meaning that the potential for finding reliable effects of ambiguity would be reduced. This, of course, is a danger with all the experiments reported in the thesis (see also Chapter 3), but it may mean that manipulations which generate only small effects in adult studies are not appropriate for use in developmental studies when stimuli must be modified to be suitable for beginning readers. This is an important methodological consideration which will receive more discussion in the final chapter of the thesis (Chapter 8, Section 8.6).

Chapter 6 : Syntactic ambiguity

6.1 Introduction

Up to this point, the focus of the experiments reported has been exclusively on the processing of individual words. While lexical processing is a fundamental part of the reading process, there is clearly more to sentence interpretation. After word identification has occurred, syntactic processing can commence. Syntactic processing, or parsing, involves the construction of the structural relationships between the words of the sentence. Syntactic ambiguity occurs when a phrase, clause or sentence could have more than one interpretation, and syntactic ambiguities have proven to be a useful means to examine adults' parsing preferences, and to test competing theories of syntactic processing.

In this chapter, two different types of (temporarily) syntactically ambiguous sentences will be used to examine whether children exhibit the same initial parsing preferences as adults. It is important to explain in some detail the syntactic structure of the experimental sentences in terms of the ambiguities they contain, because the way in which adults read such sentences is extremely informative regarding the mechanisms they have in place to parse sentences. While much is known about the way that adults process sentences containing these kinds of syntactic structures, it is not yet known whether children's parsing preferences are the same as those of adults. If children show the same pattern of reading time effects as adults, it seems reasonable to assume that the mechanisms they have in place for building syntactic structure is the same as those found in adult readers.

The chapter will be structured as follows. The Introduction (Section 6.1) will delineate the two types of syntactic structure used in Experiments 5 and 6, and discuss how different models of syntactic processing might account for the eye movement behaviour observed in adults as they read sentences containing these kinds of ambiguity. Section 6.1 will also review research which has been conducted with both adults and children, and make predictions based on these previous findings. Section 6.2 will report Experiment 5 which investigated adults' and children attachment of prepositional phrases during sentence reading. Section 6.3 will report Experiment 6 which examined the time course of adults' and children's closure of a currently-processed phrase during reading. Finally, Section 6.4 will draw some general conclusions from the results from both experiments.

6.1.1 Research with adults

Chapter 1 provided a review of some of the literature pertaining to adults' processing of syntactic ambiguity. Several of the studies outlined in Chapter 1 (e.g. Rayner et al., 1983; Taraban & McClelland, 1988) used sentences with the same syntactic structure as sentences (1a-1b: taken from Experiment 5) in which the attachment of a prepositional phrase (*with the long stick/with the long trunk* in sentences 1a-1b) is ambiguous such that it can be attached high in the syntactic tree (see Figure 6-1) to the verb (*poked*), or low in the syntactic tree to the noun phrase (*the elephant*; see Figure 6-2).

(1a) The boy poked the elephant with the long stick from outside the cage.

(1b) The boy poked the elephant with the long trunk from outside the cage.

Figures 6-1 and 6-2 show that if the prepositional phrase is attached to the noun phrase, an additional level of nodes is needed in the syntactic tree (6 levels of nodes in total; Figure 6-2) as compared to when the prepositional phrase is attached to the verb (5 levels of nodes in total; Figure 6-1). That is, sentence (1b) is syntactically more complex than sentence (1a).

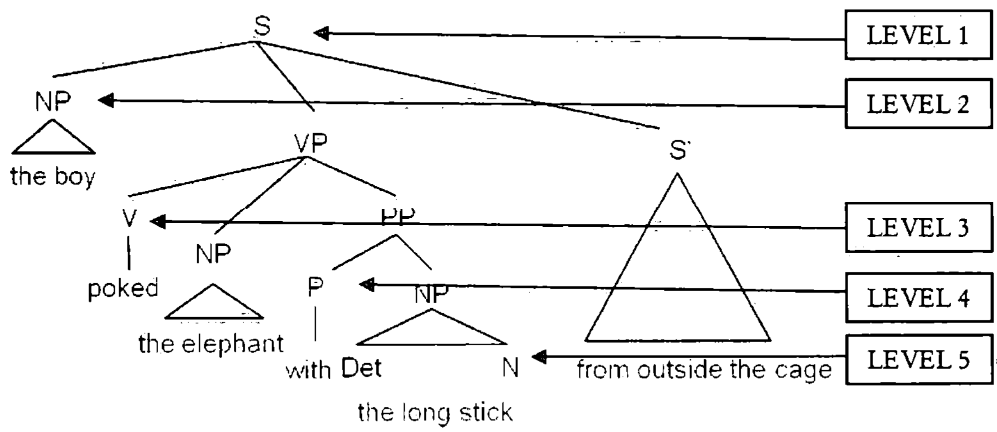

Figure 6-1: Syntax tree for Sentence 1a

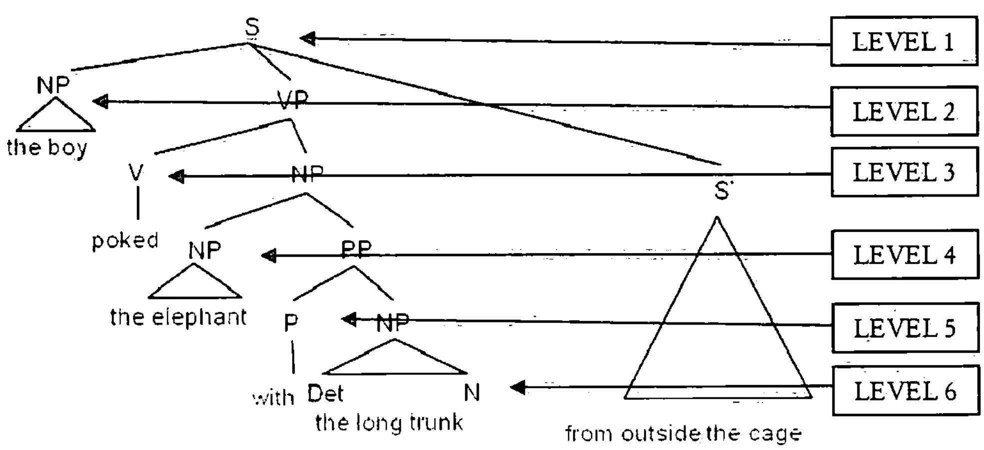

Figure 6-2: Syntax tree for Sentence 1b

As outlined in Chapter 1 (Section 1.4), while some theorists argue that parsing is guided only by syntactic principles (Clifton, Speer, & Abney, 1991; Frazier, 1978; Frazier & Rayner, 1982; Rayner et al., 1983) , some argue that other sources of information (e.g. the semantic content of a sentence, the nature of a preceding referential context, etc.) also influence initial attachment decisions (Altmann & Steedman, 1988; Crain & Steedman, 1985; MacDonald et al., 1994; Marslen-Wilson, 1975; Spivey-Knowlton & Sedivy, 1995; Tanenhaus et al., 1989; Taraban & McClelland, 1988; Tyler & Marslen-Wilson, 1977). While there are alternative accounts of how the syntactic processor operates, the Garden Path model and constraint-based models (see Chapter 1, Section 1.4 for a description of both types of models), will be highlighted here as they have been the most influential, as well as the most contentious, and they stand in direct opposition to one another.

Proponents of the Garden Path theory argue that during initial parsing the syntactic processor has access only to syntactic information, and the construction of the initial analysis is guided exclusively by two principles: Minimal Attachment and Late Closure. Specifically, they argue that, in sentences such as (1a-1b), adults will initially parse the sentence according to Minimal Attachment, and attach the prepositional phrase *with the long stick/long trunk* high to the verb *poked* (i.e. interpret it as an instrument used to *poke* the *elephant*), as this is the syntactically simpler alternative. This analysis is correct in sentence (1a) but not in (1b), leading to the prediction of longer reading times for sentence (1b) as compared to sentence (1a). Rayner et al. (1983) provided compelling evidence for their arguments, showing that adults exhibited inflated reading times on the disambiguating region of sentences like (1b) (*long trunk*) in the non-Minimal Attachment sentences as compared to in the Minimal Attachment versions of the same sentences. Rayner et al. argued that adult readers favoured this syntactic interpretation irrespective of sentential contexts that provided a semantic bias towards an alternative analysis; that is, although a *trunk* cannot (in normal circumstances) be used by a human being to *poke* someone, adults nevertheless exhibit disruption to processing when the prepositional phrase (*with the long trunk*) attaches low to the noun phrase (*the elephant*).

In contrast, proponents of interactive theories, particularly the constraint-based hypothesis, argue that information from all levels of representation are used in the construction of the initial analysis, and thus, parsing is not solely guided by syntactic principles but also by other sources of information such as the semantic content of the preceding sentence. As outlined in Chapter 1 (Section 1.4), Taraban and McClelland (1988) conducted an a self-paced reading task, in addition to sentence completion and rating tasks, in which they showed that adults differed in their attachment preferences contingent on the semantic content of the sentence up to the ambiguity, rather than dependent on whether the prepositional phrase was minimally attached. In short, they showed that readers used thematic role expectations associated with the noun-verb-preposition combinations to guide initial attachment preferences. Specifically, using Rayner et al.'s materials, they replicated Rayner et al.'s results, but using their own (tightly-controlled) materials, they showed that readers exhibited disruption when reading sentences such as (1a) relative to (1b) in direct opposition to the results from Rayner et al. However, importantly the participants in the Taraban and McClelland study did not read sentences normally (as per eye tracking procedures) but rather read sentences in a self paced reading procedure, whereby words were exposed one by one as a button was pressed. This less natural reading procedure might produce very different effects to those observed in normal text reading.

Additionally, proponents of the Referential theory (Altmann & Steedman, 1988; Crain & Steedman, 1985) argued that referential context determines how readers interpreted syntactically ambiguous sentences rather than parsing principles alone. Crain and Steedman (1985) were able to induce and prevent Garden Path effects in their participants by manipulating referential expectations induced by the preceding context. In addition, Altmann and Steedman (1988) showed that whether adults exhibited longer reading times on syntactically ambiguous sentences in which the prepositional phrase could be attached to the verb phrase or the noun phrase, was not dependent on whether the prepositional phrase was minimally attached, but rather on the precise nature of the preceding referential context. Finally, through analyses of text corpora, sentence completion tasks and self-paced reading experiments, Spivey-Knowlton and Sedivy (1995) also found that preceding referential context could induce a preference to attach a prepositional phrase to the noun phrase (as in 1b) rather than to the verb phrase (1a), although only in certain situations. Specifically, they found that the preference to attach the prepositional phrase to the verb phrase is quite strong when action-like verbs are used (e.g. hit), but not when perception verbs are used (e.g. expect), and that this was also modulated by the definiteness of the article (i.e. *a* car versus *the* car). They argued that both verb-specific attachment preferences and referential properties of the definiteness of the noun phrase play important roles in online syntactic ambiguity resolution. However, again, none of these studies monitored eye movements during normal text reading and a different pattern of effects may have been observed than during natural reading.

What is clear from the preceding discussion is that there are conflicting findings regarding whether readers consistently initially attach a prepositional phrase high or low in the phrase structure tree. However, it is also clear that by carefully constructing stimuli and pre-screening them thoroughly, it is possible to develop a set of sentences for which there is a strong and consistent preference for readers to initially attach the prepositional phrase high to the main verb. Exactly why such a preference exists could be a consequence of a number of factors (e.g., thematic verb preferences, semantic expectations, referential factors etc). However, for the purposes of the Experiment 5, the reason for the strong preference is not particularly important. Instead, the important question is whether, for a set of stimuli for which we know adult readers exhibit a strong initial parsing preference, children also exhibit such a preference. Again, the approach is to use evidence of a clear processing preference in adult readers as a diagnostic of whether a similar processing preference exists in children. In Experiment 5, sentences contained verbs which frequently take an instrument (such as *hit*, *poke*, *cut*). These were presented in a null-context and were pre-screened carefully to ensure that they would induce garden path effects whereby adults produce longer reading times on those sentences in which the prepositional phrase is attached low to the noun phrase, than high to the verb phrase. In this way, it will therefore be possible to examine whether children have the same mechanisms in place that produce similar initial syntactic processing preferences during reading as adults.

In a second experiment investigating children's syntactic processing preferences, Experiment 6 examined adults' and children's processing of sentences such as (2a-2b). In these sentences, the adverbial phrase (*yesterday/tomorrow*) can be attached low to the second (i.e. more recent) verb phrase *I bought* (thereby providing information on when the skirt was purchased) or high to the first (earlier) verb phrase *I'll wear* (providing information about when the skirt will be worn).

(2a) I think I'll wear the new skirt I bought tomorrow. It's really nice.

(2b) I think I'll wear the new skirt I bought yesterday. It's really nice.

Proponents of the Garden Path theory argue that adults will exhibit disruption to processing when reading sentences such as (2a) as compared to sentences such as (2b). This is because, according to the principle of Late Closure, readers attach incoming material to the phrase marker that is currently open, so in sentence (2a) adults initially attach the adverb *tomorrow* to the phrase marker associated with the verb phrase *I bought* resulting in disruption to processing due to a temporal mismatch between the adverb and the verb to which it initially attaches. Importantly, advocates of constraint-based theories would also predict the same pattern of effects, but the explanation for this is couched in terms of activation rather than parsing preferences. If it is assumed that the representations corresponding to the argument structures of each verb (*wear* and *buy* in sentences 2a-2b) become activated when the verb is encountered but then decay gradually over time, then it would be expected that the representation associated with the most recent verb will have decayed less than the representation associated with the earlier one (Altmann, 1998). On this basis it would be expected that readers would exhibit disruption to processing for sentences like (2a) but not like (2b).

However, and as mentioned earlier, constraint-based theorists (MacDonald et al., 1994; McRae et al., 1998; Trueswell et al., 1994) argue that semantic content influences parsing decisions in adults, and therefore how readers attach the incoming item (*tomorrow/yesterday*) will depend on the preceding context. In particular, according to Altmann and Steedman's Referential Theory (1988; Altmann, van Nice, Garnham, & Henstra, 1998; see also Crain & Steedman, 1985) referential context influences whether a reader incorporates incoming material into the current, or most recent, clause (see also MacDonald et al., 1994). Altmann et al. (1998) found that readers did exhibit disruption to processing in sentences like (2a) which violate the principle of Late Closure, in accordance with the findings from Rayner et al. However, when embedded in a context which explicitly directed attention towards the predicate associated with high attachment, context overrode this preference to generate longer reading times on late-closure sentences like (2b). Importantly, without a biasing referential context, however, adults did exhibit a preference for sentences such as (2b) rather than (2a), in line with the principle of Late Closure.

To summarise, the sentences used in Experiment 6 were carefully constructed to induce garden path effects in those sentences in which an adverbial phrase was attached high to an early verb, rather than low to a more recently-encountered verb. While previous research has shown that it is possible to override this preference with a biasing referential context, thematic preferences have been demonstrated to have an immediate influence when the stimuli are constructed carefully. It is therefore possible to use demonstrated processing preferences in adult readers to assess whether this same processing preference is observed in children, and moreover whether the magnitude and time course of any effects observed are the same for adults and children.

6.1.2 Research with children

To my knowledge, there have been no studies which have used eye movements to examine syntactic ambiguity effects during reading in children. However, there have been a number of studies that have used alternative methodologies to investigate the integration of context, pragmatic information and real world knowledge into ongoing meaning representations. Although the experiments reported in this chapter do not explicitly address the issue of whether semantic information influences initial parsing decisions, these studies will be outlined briefly as they also concern the issue of children's basic parsing preferences.

Trueswell Sekerina, Hill and Logrip (1999) investigated the influence of visual context on children's processing of syntactically ambiguous sentences, using the visual world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Four and five year-old children listened to temporarily ambiguous (or unambiguous) sentences, such as 'Put the frog (that's) on the napkin in the box,' in which a prepositional phrase (*on the napkin*) could represent either a destination for, or a modifier to, the preceding noun (*frog*) while their eye movements were monitored. They followed instructions by moving objects in a visual array. The visual context was also manipulated to favour one of the two possible interpretations of the prepositional phrase. For example, single referent contexts showed only one frog that was on a napkin, and therefore supported a destination interpretation of *on the napkin*. In contrast, two referent contexts supported a modifier interpretation as two frogs were shown, one on a napkin, and one alone. If children used the referential context to guide their parsing decisions, then their interpretation of the ambiguous sentence would depend on the visual display available to them. If, on the other hand, they did not use context to guide their attachment preferences, they would interpret the sentence in the most syntactically simple way, regardless of referential context.

Trueswell et al. found that while both adults and children showed signs of rapid incremental interpretation, children preferred the destination interpretation (the syntactically less complex alternative) regardless of visual or syntactic context. That is, children were more likely to direct their gaze to an incorrect destination during ambiguous than unambiguous trials despite a biasing referential context. This shows that, unlike adults, children did not make use of contextual cues to disambiguate syntactically ambiguous structures. Furthermore, children were not able (or else were reluctant) to revise initial commitments to the destination interpretation. Children appeared to parse according to a principle like Minimal Attachment, automatically choosing the least complex syntactic structure; but when this interpretation was incorrect, they were not able to come up with an alternative.

These findings suggest some insensitivity to pragmatic information during parsing in children. It may be that children are less adept than adults at integrating non-structural information into their online language processing, although it is worth noting that Hurewitz, Brown-Schmidt, Thorpe and Trueswell (2000) showed that children were able to use contextual cues in a language production task, indicating that the effects found in the Trueswell et al. study may be task-specific (see also Meroni & Crain, 2003; Snedeker et al., 2001 for comparable findings). The findings also indicate that strong parsing preferences may be observed in children in the experiment outlined in this chapter, and that greater disruption might be exhibited in the sentences which do not conform to Minimal Attachment and Late Closure as children may have more difficulty than adults revising their initial parsing commitment.

Traxler (2002) investigated the effect of subcategorisation and plausibility information on syntactic ambiguity resolution in children (aged 8-12 years) using a self-paced reading task, in which children read sentences such as (3a - 3c) below.

(3a) When Sue tripped the girl fell over and the vase was broken.

(3b) When Sue tripped the table fell over and the vase was broken.

(3c) When Sue fell the policeman stopped and helped her up.

He showed that children, like adults, were garden-pathed by sentences like (3a) but also by sentences like (3b) where they did not use plausibility information to help them when faced with a syntactic ambiguity (i.e. it is implausible to trip an inanimate object). He did, in his third experiment, find that while children were garden-pathed by sentences like (3c), there was a small negative correlation between intransitivity preference (how frequently a verb was intransitive as compared to transitive) and disruption to processing (as measured by total reading times). This relationship suggests that children did use subcategory information to guide attachment decisions, although the correlation could also be explained by children using subcategory information to aid recovery from misanalysis, rather than for initial parsing decisions. Nevertheless, subcategory information did not prevent children from computing the incorrect syntactic analysis. Overall, Traxler argued that the results from his experiments showed that children, in contrast to adults, favour the structurally simpler analysis whether it is plausible or not, and these conclusions sit well with those of Trueswell et al.

Finally, Felser, Marinis, and Clahsen (2003) used a self-paced listening task to investigate children's (aged 6-7 years) understanding of ambiguous relative clause sentences such as sentences 4a-4b below.

(4a) The doctor recognised the nurse of the pupils who was feeling very tired.

(4b) The doctor recognised the nurse of the pupils who were feeling very tired.

(4c) The doctor recognised the pupils with the nurse who was feeling very tired.

(4d) The doctor recognised the pupils with the nurse who were feeling very tired.

In sentences such as (4a-4d) above, while adults' preferences were influenced by the semantic properties of the preposition (*of/with*) adjoining the two potential antecedent noun phrases (*nurse* or *pupils*), children exhibited a preference for either the first or the second noun phrase dependent on their working memory capacity, but irrespective of the type of preposition involved. Like Trueswell et al., and Traxler, Felser et al. argued that children primarily rely on structural, rather than semantic, information during processing of modifier attachment ambiguities (see also Clahsen & Felser, 2006, and Sekerina Stromswold & Hestvik, 2004 for further evidence to support this claim).

Together, the results from the three studies outlined above provide strong evidence that children rely predominantly on structural information and disregard additional contextual information such as semantic or pragmatic fit during reading, and it appears that children have strong syntactic preferences regarding which analysis the parser builds. It is likely then that the children tested in Experiments 5 and 6 will exhibit strong parsing preferences, and indeed that the magnitude of these effects will be stronger than those observed in adults, due to their inability (or reluctance) to revise their initial analysis.

6.2 Experiment 5

6.2.1 Predictions

Experiment 5 investigated whether children and adults exhibited the same initial attachment preferences when reading temporarily ambiguous sentences such as (1a-1b). In line with Rayner et al. (1983), it was predicted that both adults and children would exhibit disruption to processing on the disambiguating region when the prepositional phrase (*long stick* or *long trunk* in sentences 1a and 1b) was not attached to the verb (*poked*) phrase but rather to the noun phrase (*elephant*): that is, when the sentence was not minimally attached. While there have been studies to show that adult readers do not always interpret sentences in this way, the materials were designed, like those used by Rayner et al., to induce a thematic preference (i.e. an instrument interpretation of the prepositional phrase) in that they contained action-type verbs which usually take an instrument, and did not follow a referential context. However, importantly, this was not simply assumed, but rather a pre-screen sentence completion task was employed to determine preferences of both adults and children. In the experiment proper, it was predicted that adults and children would make longer fixations, and more regressions out of prepositional phrases which attached low to the noun phrase rather than high to the verb phrase.

The wide range of ages in the child participants in this experiment provided the opportunity to examine within a single experiment whether there are age-related changes that take place in relation to children's online parsing commitments. It was tentatively predicted that there might be a difference in the time course of the effects observed in younger children, older children and adults, specifically that while adults were expected to show increased fixation durations on, and regressions from, the target region, it was anticipated that in children such effects might be spatially localised to words downstream of the target region, or else observed during second pass measures of reading time.

Finally, following the results from Experiments 1, 2 and 4, in which children showed stronger effects of word length, word frequency and lexical ambiguity than adults and the results from the Trueswell et al. study, it was further anticipated that the magnitude of the effects observed would be greater in children than in adults, and indeed greater in younger children than older children; that is, there would be a developmental decrease in the strength of the effects observed.

6.2.2 Method

Participants. The same 30 adults, who took part in Experiments 2 and 3 (see Chapter 4, Section 4.2.1) took part in the experiment, as part of the same experimental session. Twenty four children took part in the experiment. The children were split into two age groups. In the younger group, the mean age was 7.7 years (range = 6.5 - 8.9 years), and in the older group the mean age was 10.3 years (range = 9.1 - 11.7 years). The mean reading age was 10.3 years in the younger group, and 12.8 years in the older group (see Chapter 3, Section 3.2.1 for details of the reading measure used).

Materials. A pre-screen procedure was undertaken before the main experimental stimuli were prepared to examine which of the two possible syntactic analyses (i.e. Sentence 1a or Sentence 1b) adults, older children and younger children preferred, and whether the groups differed in their preferences. Eighteen adults, 16 older children (aged 10-11 years) and 14 younger children (aged 7-8 years) took part in a sentence completion task.

The sentence completion task was computerised using Macromedia Flash MX software. All participants read written instructions and children received additional verbal instructions. Participants viewed a screen which showed the beginning portion of the proposed experimental sentences (e.g. *The boy poked the elephant with the...*). Participants were instructed to complete the sentences by typing one or more words following the sentence fragment. They could finish the sentence in any way they chose but the sentences had to be grammatically correct. Once the participant had typed a response, s/he clicked a button to view the following sentence. There were sixteen sentences in total.

Responses were then categorised as high-attached (HA), such as '*The boy poked the elephant with the long stick*' or as low-attached (LA), such as '*The boy poked the elephant with the long trunk*'. In the HA condition, the prepositional phrase (*with the stick*) is an instrument of the verb (*poked*), while in the LA condition, the prepositional phrase modifies the noun phrase (*the elephant*). Results showed that adults completed the sentences with an instrument (i.e. the completions were HA) on 78.5% of occasions. Older children completed the sentences with an instrument on 86.3% of occasions, and younger children on 86.2% of occasions. These results show (1) that all groups of participants had the same (offline) attachment preferences; (2) that this preference is stronger in children than adults; and (3) that there is no difference between older and younger children in terms of these particular parsing preferences.

Two versions of the sixteen experimental sentences were then constructed, as shown in sentences (1a: HA) and (1b: LA) repeated here.

REGION1 REGION2 REGION3 REGION4 REGION5

(1a) The boy poked the elephant with/ the long/ stick/ from/ outside the cage.

(1b) The boy poked the elephant with/ the long/ trunk/ from/ outside the cage.

Sentences were divided into five regions. Region 3 was the target region. All target words were preceded by the word *the* or *a* and an adjective. All but two of the sentence pairs were identical apart from the target word (*stick* or *trunk* in sentences 1a and 1b). For these two sentence pairs, the adjective was different for each condition. However, there was no difference in length or frequency (using both adult and child corpora) between conditions for either the adjective preceding the target word, or the target word itself (all ps > .2). At least one long word, or two short words, always followed the target word.

Apparatus. The apparatus used was identical to that used in Experiment 2 (Chapter 4, Section 4.2.1).

Procedure. The procedure for the main eye-tracking experiment was identical to that in Experiment 2 (Chapter 4, Section 4.2.1). Of the 16 experimental sentences, four were followed by comprehension questions.

6.2.3 Results and Discussion

Fixations longer than 1200ms and shorter than 80ms were systematically excluded from the data set. Trials in which there was tracker loss or excessive blinks were also excluded. In addition, any trials in which the participant did not fixate either the target region (*stick/trunk* in sentences 1a and 1b) or the post-target region (*from* in sentences 1a and 1b) were deleted. Outliers (2.5 Standard Deviations above or below the mean) were also excluded. This resulted in the exclusion of 12% of the data in total. All participants performed very well on the comprehension questions with children answering 93% of questions correctly, and adults answering 97% of questions correctly. All participants answered a minimum of 75% of questions correctly.

First fixation durations, gaze durations, the probability of making a first pass regression out of a critical region, go past time and total reading time were all calculated for eye movements in the target and the post-target regions (see Chapter 3, Section 3.3.3 for a description of eye movement measures). Mean reading times and regression probabilities in the target region are shown in Table 6-1.

Table 6-1: Mean reading times and regression probabilities for adults, older children and younger children in the target region in high-attached (HA) and low-attached (LA) conditions. Standard deviations in parentheses.

		First fixation duration	Gaze duration	First pass regressions out	Go past time	Total time
Adults	HA	229 (*64*)	250 (*82*)	0.05 (*0.09*)	266 (*103*)	303 (*133*)

	LA	229 (*54*)	259 (*84*)	0.10 (*0.11*)	283 (*114*)	321 (*153*)
Older Children	HA	275 (*101*)	341 (*155*)	0.13 (*0.16*)	402 (*226*)	429 (*242*)
	LA	271 (*92*)	337 (*125*)	0.11 (*0.15*)	391 (*198*)	468 (*260*)
Younger children	HA	293 (*122*)	363 (*188*)	0.10 (*0.15*)	424 (236)	483 (*302*)
	LA	314 (*134*)	394 (*191*)	0.08 (*0.14*)	439 (*229*)	517 (*303*)

A 3 (group: adult, older children, younger children) x 2 (attachment: high-attached, low-attached) mixed design ANOVA showed no reliable effects of attachment in any measure (all $ps > .15$). There were effects of group in every reading time measure, with both groups of children making longer first fixations, gaze durations, go past times, and total word reading times than adults ($ps < .05$), but there was no difference between groups in the number of first pass regressions made ($ps > .1$). Furthermore, there were no differences in reading times or refixation probabilities between older and younger children ($ps > .1$).

There were no reliable interactions (all $ps > .15$). However, adults made marginally more regressions out of the target region in the LA condition than the HA condition, reliable by items and marginal by participants, $t1\ (29) = 1.89$, $p = .069$, $t2\ (15) = 2.69$, $p < .05$ (see Figure 6-3 below). While these pairwise analyses did not follow a reliable interaction, the graph is presented here as it shows a numerical trend, perhaps reflecting an effect in the adults. In contrast to the adults, both groups of children showed the opposite pattern of effects as those predicted although this difference was small (and not reliable).

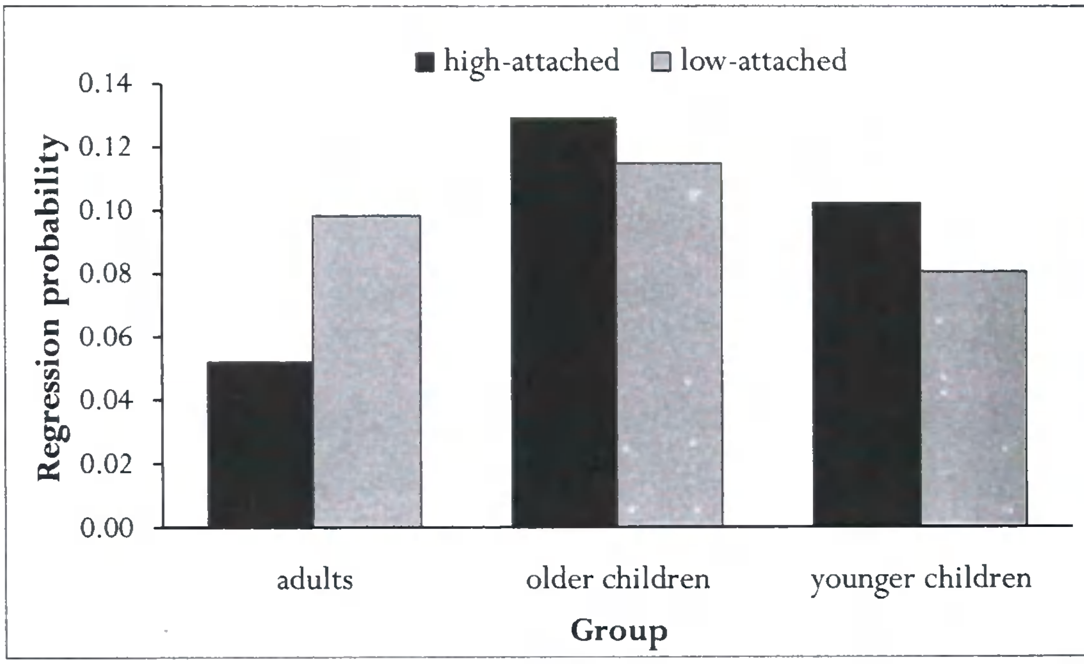

Figure 6-3: Regression probabilities for adults, older children and younger children in the target region for high-attached (HA) and low-attached (LA) conditions.

Analyses were then conducted on the post-target region (see Table 6-2). Note that the post-target region was not the final region in the sentence, and hence effects observed in the post-target region should not be due to clause wrap-up (Rayner, Kambe, & Duffy, 2000).

Table 6-2: Mean reading times (ms) and regression probabilities for adults, older children and younger children in the post-target region in high-attached (HA) and low-attached (LA) conditions. Standard deviations in parentheses.

		First fixation durations	Gaze durations	First pass regressions out	Go past times	Total times
Adults	HA	251 *(90)*	322 *(155)*	0.11 *(0.16)*	402 *(255)*	380 *(198)*
	LA	255 *(92)*	305 *(132)*	0.16 *(0.17)*	405 *(248)*	400 *(203)*
Older children	HA	281 *(131)*	428 *(252)*	0.17 *(0.18)*	517 *(374)*	515 *(293)*
	LA	294 *(132)*	420 *(232)*	0.25 *(0.25)*	645 *(523)*	609 *(327)*
Younger children	HA	309 *(129)*	416 *(240)*	0.13 *(0.15)*	513 *(409)*	518 *(373)*
	LA	313 *(145)*	451 *(258)*	0.25 *(0.18)*	791 *(606)*	575 *(341)*

In the post-target region, there were no effects of attachment in first fixation durations or gaze durations (ps > .2). There was no effect of attachment in the number of regressions made out of the post-target region, which was marginal by participants, F1 (1, 51) = 3.86, p = .055, but not reliable by items (p > .9). However, there was a reliable effect of attachment in go past time, F1 (1, 50) = 10.78, p < .005; F2 (1, 15) = 9.37, p < .01. Furthermore, there was an interaction in this same measure, which was reliable by participants and almost reliable by items, F1 (2, 50) = 5.13, p < .01; F2 (2, 30) = 3.01, p = .06 (see Figure 6-4 below). Pairwise analyses showed no difference in go past times for adults (ps > .5), no difference for older children (although the effect was reliable by items), t1 (11) = 1.05, p = .32; t2 (15) = 2.74, p < .05, and a reliable difference for younger children, t1 (11) = 2.78, p < .05; t2 (15) = 2.23, p < .05. This was the earliest point at which (younger) children exhibited an effect of attachment and shows that children were delayed in their processing of syntactic ambiguity as compared to adults (assuming the effect of attachment in the target region in adults was real).

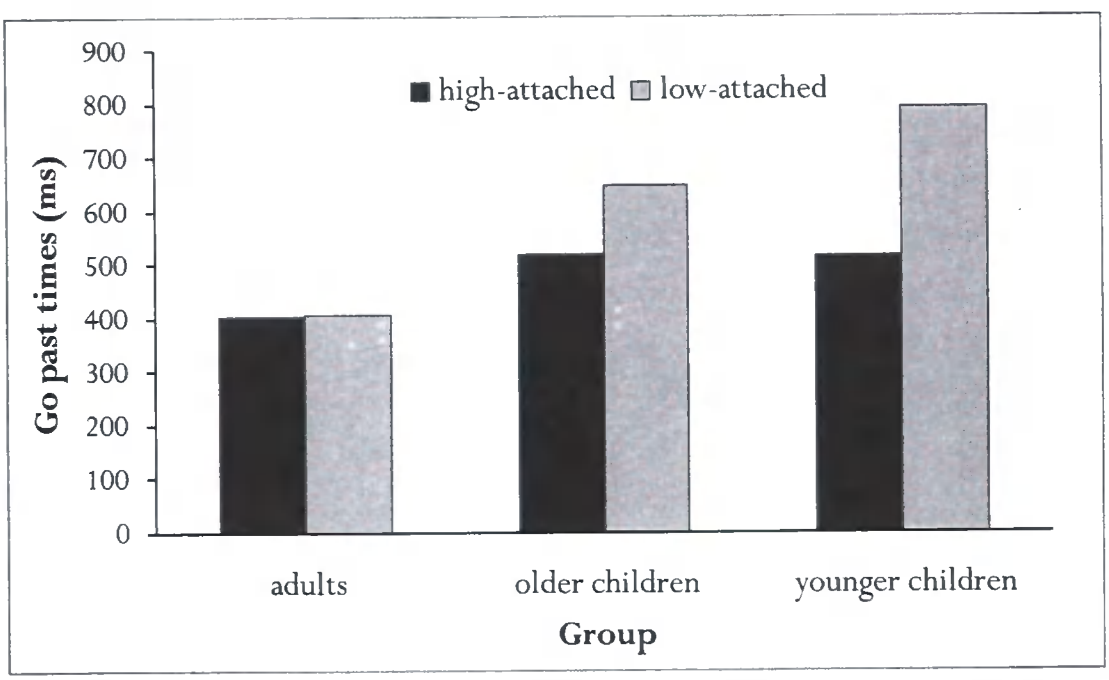

Figure 6-4: Go past times (ms) for adults, older children and younger children in the post-target region for high-attached (HA) and low-attached (LA) conditions.

There was an effect of attachment in total word reading time in the post-target region, reliable by participants but marginal by items, F1 (1, 50) = 5.93, p < .05; F2 (1, 15) = 3.79, p = .07.

There was also an effect of group in all reading time measures (all ps < .05), showing the same pattern of effects as reported in the target region. There were no interactions (ps > .2) apart from that reported above in go past time. Overall, the results from the post-target region show that younger children and older children (marginally) were disrupted by the LA sentences as compared to the HA sentences. Adults, however, showed no effect of attachment preference in the post-target region, having already showed a (marginal) effect in the target region.

Overall, it was predicted that all age groups would exhibit a preference for attaching the prepositional phrase high to the verb, rather than low to the noun phrase, and that children would show delayed effects of attachment as compared to adults. This prediction was (marginally) met. While adults made more regressions out of the target region, thereby showing an immediate effect of attachment, children exhibited this same pattern of effects only in the post-target region.

Before discussing the implications of this result, it should be noted that some of these conclusions are based on marginal results. Older children showed longer go past times in the post-target region, but this effect was not reliable by participants. Adults made more regressions out of the target region in the LA than the HA condition, but this difference was marginal by participants, and it did not follow a reliable interaction. The lack of robust effects may be due to a number of factors. First, there were only 16 items in the analyses which resulted in a lack of power. The reason that so few items were included was that it was thought that 32 items in total (i.e. from Experiments 5 and 6 together), in addition to a further 24 sentences from Experiment 2, would be the most that younger children could be expected to read in a single experimental session. Another factor which may have influenced the lack of significant effects was the same one outlined in Chapter 5 (Section 5.4.1): the sentences were age-appropriate for young children making them very easy for the adults to read and this may have accounted for the relatively weak effects in the adult participants. Finally, as remarked on in Section 6.2.1, the increased variability introduced by the child data made finding significant effects in the adults (and the children) difficult, Indeed, as mentioned, it was for this reason that pairwise analyses were conducted on the adult data separately which revealed an earlier (marginal) effect in the adult group.

Overall, Experiment 5 shows that both adults and children preferentially attached a prepositional phrase high in the syntactic tree, resulting in fewer levels of nodes. This finding is supportive of the Rayner et al. findings. Furthermore, the data strongly suggest that both older and younger children were delayed relative to adults in exhibiting this preference during normal reading. However, no differences between older and younger children were observed in terms of the time course of effects, although younger children exhibited stronger effects than older children. It appears, then, that early in the process of reading development, the mechanism that children have in place to attach a prepositional phrase high to a verb phrase, rather than low to a noun phrase, is qualitatively the same as that of adults, but the time course in which this takes place is not the same.

6.3 Experiment 6

6.3.1 Introduction and predictions

Experiment 5 showed that, like adults, children prefer to attach a prepositional phrase high to a verb rather than low to a noun during reading (although this process is delayed in children relative to adults), and that this impacts on their eye movement behaviour. Experiment 6 investigated children's and adults' processing of a different ambiguous syntactic structure, in which whether an incoming adverb was attached to the currently-processed clause, or to an earlier noun phrase, was manipulated.

The principle of Late Closure in the Garden Path theory stipulates that participants should preferentially attach new items to the phrase or clause currently being developed rather than starting a new phrase or clause. According to the Garden Path model, in sentences (2a) and (2b), readers should initially attach the adverb (*tomorrow/yesterday*) to the phrase currently being processed (*I bought*) and so should exhibit disruption on target words in the early-closure (EC) condition (*tomorrow*) as compared to in the late-closure (LC) condition (*yesterday*). While this pattern of effects can change when preceded by a biasing context (see Altmann et al., 1998), unlike the sentences in Experiment 5, the finding that adult exhibit less disruption to processing when reading sentences like (2b) than sentences like (2a) is relatively uncontroversial. For this reason, it was not deemed necessary to carry out a pre-screen procedure as was done for Experiment 5.

It was predicted that adults, older children, and younger children would exhibit disruption to processing in the EC than the LC condition which would be observed in longer reading times and increased regressions on the EC sentence. More specifically, and in line with previous research (Altmann et al., 1998), it was predicted that adults would exhibit longer gaze durations on the target word (*tomorrow*) and would make increased regressions out of this same word to refixate the previous portion of the sentence (*I bought*) to which they had (wrongly) attached the adverb. While adults would show these effects on the target word (*tomorrow/yesterday*), it was anticipated that children would exhibit later effects, in line with the results from Experiment 5. Furthermore, it was expected that younger children would be more delayed in these effects as compared to older children.

6.3.2 Method

Participants. The same 30 adults and 24 children from Experiment 5, plus an additional four children (four children were excluded from Experiment 5 due to tracker loss or blinking) took part in Experiment 6, totalling 28 children in all. The children were split by age as in Experiment 5. The younger age group had a mean age of 7.9 years (range = 6.5 - 9.0 years) and the older group had a mean age of 10.4 (range = 9.5 - 11.7 years). The mean reading age in the younger group was 10.4 years and the mean reading age in the older group was 13.3 years.

Materials. Experimental sentences were constructed in which the attachment of an adverbial phrase was manipulated. In sentences (2a) and (2b) below (repeated from above), the adverbial phrase was either attached high to the verb phrase *I'll wear* (*tomorrow*), resulting in early closure (EC) of the currently processed phrase, or low to the verb phrase *I bought* (*yesterday*), resulting in late closure (LC) of the currently processed phrase in accordance with the Garden Path theory.

The target region was Region 5. Region 6 was included in the experimental sentences in order to avoid clause wrap-up effects (Rayner et al., 2000) on the target word.

REGION1	REGION 2	REGION3	REGION4	REGION5	REGION 6

(2a) I think/ I'll wear/ the new skirt/ I bought/ tomorrow./ It's really nice./

(2b) I think/ I'll wear/ the new skirt/ I bought/ yesterday./ It's really nice./

Apparatus. The apparatus used was identical to that used in Experiment 2 (Chapter 4, Section 4.2.1).

Procedure. The procedure for the main eye-tracking experiment was identical to that in Experiment 2 (Chapter 4, Section 4.2.1). Of the 16 experimental sentences, four were followed by comprehension questions.

6.3.3 Results and Discussion

Fixations and trials were excluded according to the same criteria as in Experiment 5 (see Section 6.2.3). In addition one trial was excluded completely from the analyses due to a typographical error leaving just 15 items. This resulted in the exclusion of 9% of the data in total. All participants performed well on the comprehension questions with children answering 98.9% of questions correctly, and adults answering 99.5% of questions correctly. All participants answered a minimum of 75% of questions correctly.

First fixation durations, gaze durations, the probability of making a first pass regression out of a region, total word reading time, and second pass reading time (total fixation durations in a region after having left that region to the right) were all calculated in the target region (Region 5). The probability of making a regression into the region and second pass reading time were also calculated for Regions 2 and 4 which were the regions that contained the two possible verb phrases to which the adverbial phrase in Region 5 could be attached (see sentences 2a-2b). Mean reading times and regression probabilities in the target region are shown in Table 6-3 below (see Chapter 3, Section 3.3.3 for a description of eye movement measures).

Table 6-3: Mean reading times and regression probabilities for adults, older children and younger children in the target region for early closure (EC) and late closure (LC) conditions. Standard deviations in parentheses.

	Adults		Older children		Younger children	
	Early closure	Late closure	Early closure	Late closure	Early closure	Late closure
First fixation duration	230 *(64)*	219 *(58)*	277 *(107)*	271 *(124)*	268 *(92)*	279 *(100)*
Gaze duration	362 *(191)*	331 *(171)*	506 *(300)*	506 *(275)*	543 *(342)*	508 (269)
No. first pass regressions out	0.21 *(0.21)*	0.15 *(0.13)*	0.27 *(0.18)*	0.15 *(0.16)*	0.30 *(0.22)*	0.20 *(0.15)*
Total time	457 *(249)*	411 *(214)*	875 *(509)*	719 *(407)*	917 *(587)*	746 *(519)*
2nd pass reading time	363 *(240)*	229 *(101)*	471 *(306)*	687 *(622)*	681 *(475)*	614 *(501)*

In the target region, a 3 (group: adults, older children, younger children) x 2 (closure: EC, LC) mixed design ANOVA showed no effect of closure in first fixation durations (ps $> .5$) or gaze durations (ps $> .1$). There was an effect of group in both measures (all ps $< .001$) with both groups of children making reliably longer first fixations and gaze durations than adults (ps $\leq .01$), but no differences between the older and younger children ($p > .5$). As in Experiment 5, the pattern of effects of participant group was the same across measures and regions, and these effects will not be reported in full unless they differ from the above pattern. There were no reliable interactions between group and closure (all ps $> .3$). Note that pairwise analyses with the adult group alone revealed reliably longer gaze durations in the EC than LC sentences on the target region, $t1\ (29) = 2.32$, $p < .05$; $t2\ (14) = 2.12$, $p = .05$ for adult readers. As in Experiment 5, while this result did not follow a reliable interaction, this effect of closure in the target region suggests that this was the earliest point at which adults exhibited disruption to processing the EC structure.

There was a reliable effect of closure, however, on the number of first pass regressions made out of the target region, $F1\ (1, 55) = 8.79$, $p < .005$; $F2\ (1, 14) = 6.80$, $p < .05$. Moreover, there was no effect of group ($ps > .3$) and no interaction ($ps > .6$) showing that all age groups exhibited disruption to processing in the EC compared to the LC condition (see Figure 6-5).

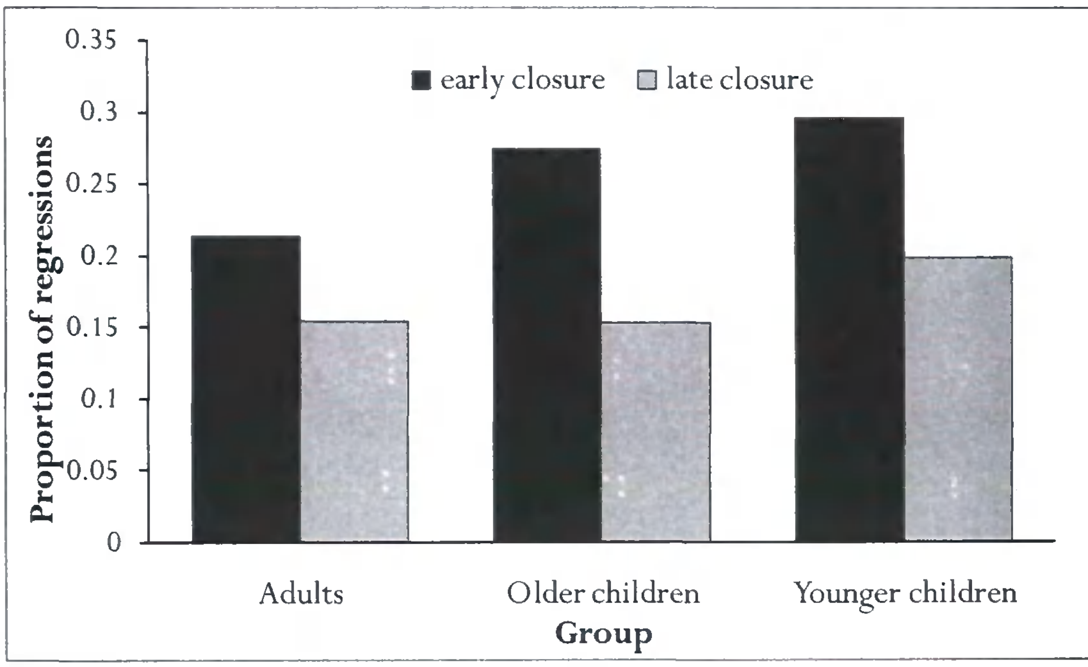

Figure 6-5: Proportion of regressions made out of the target region by adults, older children and younger children in early closure and late closure conditions.

In total reading time, there was a reliable effect of closure, $F1\ (1, 55) = 24.47$, $p < .001$; $F2\ (1, 14) = 12.81$, $p < .005$, and an effect of group ($ps < .001$). There was also a marginal interaction, $F1\ (2, 55) = 2.22$, $p = .1$; $F2\ (2, 28) = 3.58$, $p < .05$. Pairwise analyses showed that while older children ($t1\ (13) = 2.31$, $p < .05$; $t2\ (14) = 3.28$, $p < .01$) and younger children ($t1\ (13) = 2.72$, $p < .05$; $t2\ (14) = 2.66$, $p < .05$) showed reliably longer total reading times in EC than LC conditions, adults showed marginal differences, $t1\ (29) = 2.95$, $p < .01$; $t2\ (140 = 1.77$, $p = .098$. Children therefore showed stronger effects of closure in total reading time, perhaps suggesting that the disruption to processing they exhibited was longer lasting than for the adults (see Figure 6-6).

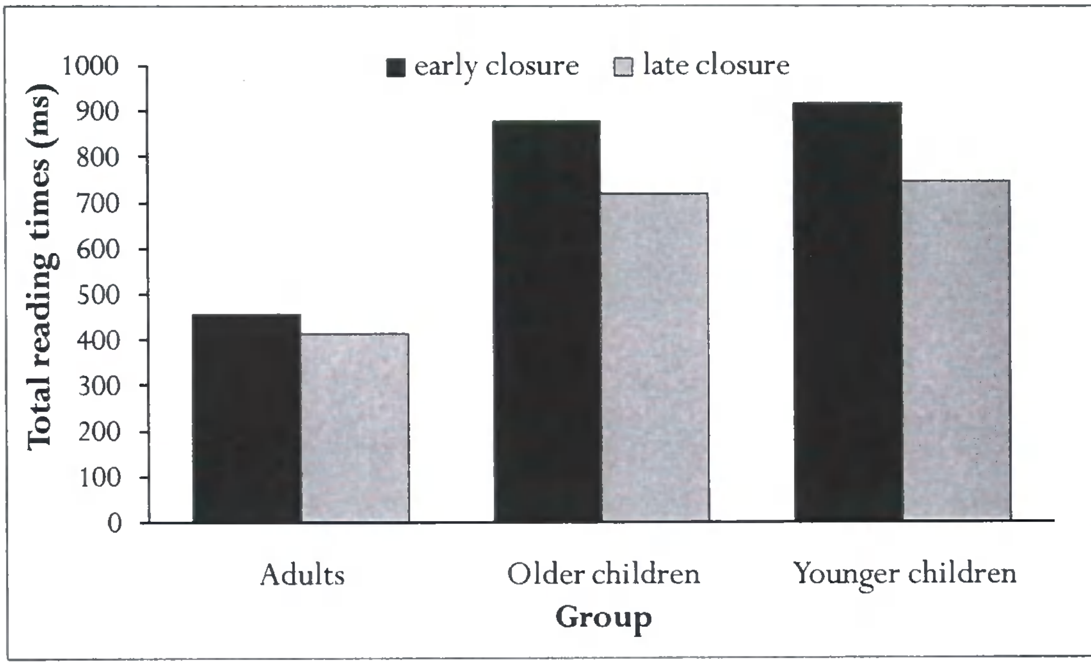

Figure 6-6: Total word reading times in the target region for adults, older children and younger children in early closure and late closure conditions.

Finally, in the target region, there was no effect of closure ($ps > .5$), an effect of group, $F1\ (2, 32) = 8.20$, $p < .005$; $F2\ (2, 12) = 5.70$, $p < .05$, but no reliable interaction, $F1\ (2, 32) = 1.91$, $p = .16$; $F2\ (1.7, 7) = 6.23$, $p < .05$ in second pass reading time.

There were two other important regions in the experimental sentences in which eye movement records were examined: Region 2 (*I'll wear* in sentences 2a and 2b) and Region 4 (*I bought* in sentences 2a and 2b). Table 6-3 below shows the proportion of regressions made into Regions 2 and 4, and second pass reading times in those same regions.

Table 6-4: Probability of regressing into, and second pass reading times (ms) in, Regions 2 and 4, for adults, older children and younger children for early closure (EC) and late closure (LC) conditions. Standard deviations in parentheses.

		Region 2		Region 4	
		Regressions into region	Second pass reading time	Regressions into region	Second pass reading time
Adults	EC	0.20 (*0.15*)	336 (*167*)	0.29 (*0.21*)	288 (*154*)
	LC	0.20 (*0.15*)	296 (*175*)	0.24 (*0.16*)	284 (*156*)
Older children	EC	0.36 (*0.22*)	486 (*251*)	0.40 (*0.22*)	569 (*410*)
	LC	0.32 (*0.23*)	513 (*366*)	0.29 (*0.16*)	472 (*281*)
Younger children	EC	0.29 (*0.23*)	584 (*363*)	0.40 (*0.25*)	577 (*357*)
	LC	0.28 (*0.18*)	517 (*324*)	0.25 (*0.18*)	463 (*373*)

In Region 2, there was no effect of closure (ps > .4) in either measure. There was an effect of group (ps < .05) in both measures, and no interactions (ps > .1). However, in Region 4, there was a reliable effect of closure, with all groups making more regressions into Region 4 in the EC than LC condition, F1 (1, 55) = 8.21, p < .01; F2 (1, 14) = 5.79, p < .05, in line with predictions. There was no effect of group (ps > .08) and no interaction (ps > .4) in this measure.

Overall, the results show that both adults and children (older and younger) exhibited disruption to processing in the EC condition on the target word. All groups of participants made more regressions out of the target word in the EC condition to refixate the previous region which contained the verb phrase to which participants had attempted to attach the adverbial phrase. They did not, however, make more regressions into the region which contained the verb phrase which the adverbial phrase was in fact attached to in the EC condition (i.e. *I'll wear* in sentences 6a-6b).

It was predicted that children would exhibit delayed effects of attachment compared to adults and that these effects would be more delayed in younger than older children. These predictions were partially met. Adults exhibited longer gaze durations on the disambiguating target word while older and younger children did not. Although these pairwise comparisons did not follow a reliable interaction between group and closure, it is argued that the time course of closure effects was indeed different for adults and children (i.e. that adults were slightly quicker than children to detect that their initial syntactic interpretation was not appropriate). However, the only very firm conclusion that can be drawn from the data is that all groups of participants showed effects of the manipulation on the target word itself. This result shows that very early in reading development, children learn to attach incoming items as a constituent of the phrase being currently parsed, and such an observation suggests that this principle is crucial to successful and efficient written language processing.

6.4 General Discussion

Two experiments investigating adults' and children's parsing preferences were reported in this chapter. It was predicted that children, like adults, would exhibit disruption to processing sentences in which (1) a prepositional phrase was attached low to a noun phrase rather high than to a verb phrase, and (2) an adverbial phrase did not attach to the clause directly before it but rather to an earlier verb phrase. It was further predicted that there would be a difference in the time course of these effects: the effects observed would be delayed in older children relative to adults, and delayed in younger children relative to older children. Finally, it was predicted that the magnitude of the effects would be greater in younger children than older children, and greater in older children than adults.

While the effects observed were not quite as robust as predicted, the data reported from the two experiments suggest that the first prediction was correct. All groups exhibited reliable effects of attachment and closure in the predicted direction. That is, all groups exhibited disruption to processing when reading sentences which were syntactically more complex. This result suggests that adults and children (as young as seven years old) have the same mechanisms in place with which they parse sentences: the developing and the mature syntactic processing systems appear to have the same attachment preferences which are used to guide reading. Although the experiments reported in this chapter did not seek to discriminate between competing models, it should be noted that the preferences exhibited by all groups were in line with those predicted by the Garden Path model.

Furthermore, in relation to the second prediction, adults exhibited first pass effects of syntactic ambiguity on the target word (that is immediately) in both experiments while children exhibited effects in the post-target region in Experiment 5, and on the target region in a later measure in Experiment 6. It can, therefore, be argued that while all groups were garden-pathed, children took a little longer than adults to detect their misanalysis. It can be said then that there do not appear to be any qualitative differences in the syntactic processing mechanisms of adults and children, only small quantitative differences which affect the time course with which garden-path effects are observed.

Finally, in relation to the third prediction, children exhibited stronger effects of syntactic ambiguity than adults in both experiments, and younger children exhibited stronger effects of attachment than older children in Experiment 5. The effect of syntactic ambiguity was also longer lasting in children than in adults. It seems then that while children need a little longer to process that they have parsed a sentence incorrectly, once they have detected their misparse, they are more disrupted, and take longer than adults to recover from it. This interpretation goes well with findings from Trueswell et al. (1999) who found that children were less able than adults to revise initial parsing commitments, and tended to perseverate with their original (syntactically simpler) analysis.

In conjunction with the results from Experiments 1, 2 and 3 in which adults and children exhibited word length and word frequency effects, the results from Experiments 5 and 6 suggest that adults and children are similar in how they process language at a lexical, and post-lexical level, but that children have not yet developed effective strategies to revise parsing commitments as adults have. This means that during reading of simple unproblematic text, children behave very much like adults, but when children encounter a structure which is not preferred (i.e. is not the syntactically simplest analysis), behavioural differences can be observed in the eye movement records as they struggle to reanalyse the sentence relative to adults.

There are several issues which merit further discussion in relation to the results obtained. First, it is worth noting that the effects observed in Experiment 6, which investigated the closure of a currently-processed phrase, were more immediate in children than the effects in Experiment 5 investigating the attachment of a prepositional phrase. It could be that there is something about the closure of a phrase which means it is particularly important in the developing syntactic processor, and that this criterion by which sentences are parsed actually emerges earlier in children than that of the attachment of a prepositional phrase. However, it is perhaps more likely that the experimental sentences chosen simply contained a more robust manipulation. That is, while participants were disrupted by encountering a prepositional phrase which modified the noun rather than the verb in Experiment 5; in Experiment 6, the disruption caused when participants attempted to attach the adverbial phrase to a verb phrase which denoted a completely different tense, resulted in a garden path of greater magnitude and immediacy. It may be that readers simply struggled to comprehend this syntactic misanalysis to a greater degree.

Finally, some of the same methodological issues which were raised in Chapter 5 in relation to the lexical ambiguity manipulation are also relevant to this chapter. The increased variability in eye movement measures brought into the analyses by the child data may have clouded some of the effects. Indeed, this was assumed to be the case and was the justification used to conduct pairwise analyses on the adult group separately which resulted in the claims made regarding differential effects in the adult and child groups. Another methodological difficulty in both of the experiments reported here was that there were comparatively few experimental sentences. This meant that for many of the analyses, effects were reliable by participants but not by items, making the arguments put forward weaker. It is for this reason, together with the ever-present problem of increased variability, that despite relatively weak effects, it has been argued that these do show real differences between the adult and child participants in terms of the time course of effects.

In summary, the experiments reported in this chapter showed that both adults and children exhibited disruption when processing sentences which were syntactically more complex. While adults showed more immediate effects of syntactic ambiguity than children, children, in particular younger children, exhibited stronger and longer-lasting effects. In conclusion, children appear to have the same sentence-parsing mechanism in place as adults, but the time course in which it operates is slightly delayed relative to adult readers.

Chapter 7 : Thematic processing

7.1 Introduction

Six experiments have now been described which have shown that while children and adults are remarkably similar in terms of their word-level, online processing, children are slightly delayed relative to adults in their syntactic processing. The final experiment conducted as part of this thesis investigates children's and adults' processing of thematic implausibility and anomaly, the highest level of linguistic processing investigated in the thesis.

7.1.1 Thematic roles

As outlined in Chapter 1, when a reader encounters a verb, as well as the verb's subcategorisation frame, the verb's thematic grid also become available. A thematic role of a verb refers to the part played by an entity in an event. Verb arguments are an intrinsic part of the verb's meaning and are central to understanding the sentence as a whole, as well as the verb. There are a finite number of entities which serve as plausible thematic roles to any one verb, and while not all the thematic roles for a particular verb must necessarily be explicitly filled (Mauner, Tannenhaus, & Carlson, 1995), they must all be plausible if the reader is not to experience disruption to processing. The experiment reported in this chapter involved a manipulation of the plausibility of the thematic roles associated with a verb in order to investigate whether children, like adults, exhibited disruption to processing when faced with an implausible or anomalous verb argument.

Two competing theories of thematic roles and the place they occupy within the language processor have been proposed which are relevant to the current work. One theory proposes that a verb's thematic role is a concept which is gradually formed from everyday experiences of who does what to whom and when (Ferretti, Gagne, & McRae, 2003; Ferretti, McRae, & Hatherell, 2001; McRae, Ferretti, & Amyote, 1997), while a competing theory suggests that thematic roles are part of a verb's lexical structure, specific to that verb, and are only activated when that particular verb is encountered (Conklin, Koenig, & Mauner, 2004; Koenig, Mauner, & Bienvenue, 2003; Koenig, Mauner, Bienvenue, & Conklin, 2008; Mauner et al., 1995). If the former theory is correct, then assignment of a verb's thematic roles requires access to real world knowledge. If the latter is true, thematic role assignment does not call upon external sources of information and can be viewed as a purely linguistic process.

McRae, Ferretti and colleagues' theory is especially interesting in terms of the objectives of the thesis because if thematic role concepts develop over time and with experience, then it might be expected that a systematic difference in the way in which adults and children process sentences containing implausible and anomalous thematic roles would be observed. While adults would be expected to have formed strong, stable expectations regarding the co-occurrence of particular verbs and arguments, children's expectations might be much less rigid and their processing of sentences containing implausible or anomalous thematic relations may therefore be less disrupted in comparison with their processing of control plausible sentences, as compared to adult readers.

7.1.2 Research with adult readers

Much of the research investigating adults' processing of plausibility information has concentrated on its role in the guidance of parsing in sentences containing a syntactic ambiguity (Blodgett & Boland, 2004; Henderson & Ferreira, 1990; Pickering & Traxler, 1998; Thornton & MacDonald, 2003; Traxler, Morris, & Seely, 2002). However, as outlined briefly in Chapter 1 (Section 1.5), Marslen-Wilson, Brown and Tyler (1988) conducted some early research into semantic as compared to pragmatic anomalies, independently of syntactic structure. In their experiment, participants listened to sentences such as (1a-1d) below and were asked to detect a monitoring target word (*guitar* in all sentences) by pressing a response button.

1a. John carried the guitar. (control)

1b. John buried the guitar. (pragmatic anomaly)

1c. John drank the guitar. (semantic anomaly)

1d. John slept the guitar. (categorical anomaly)

Sentences of types (1b) and (1c) are of particular relevance to the experiment reported in this chapter. Marslen-Wilson et al. differentiated between these two conditions in terms of linguistic and non-linguistic knowledge. Sentence (1b) is anomalous in a pragmatic sense and inferences about the real world are necessary in order to recognise the anomaly. While it is certainly unusual for people to bury guitars, it is quite possible to find a real world context in which such an action would be perfectly acceptable. In contrast, sentence (1c) does not require knowledge about the real word, only a concept of 'something drinkable' and of a guitar as 'something solid' (and therefore non-drinkable) to detect the anomaly. Results showed significantly longer monitoring latencies in response to both pragmatic, and in particular semantic, anomalies, as compared to control sentences, implying that both types of anomaly have immediate effects on processing, and that non-linguistic domains of interpretation and inference are quickly integrated with thematic properties of verb argument frames.

However, it is important to note that Marslen-Wilson et al. examined spoken (rather than written) language comprehension, and used a relatively crude measure of processing time that was supplementary to the listening task: response latency to diction of a target word. It is entirely possible that using a sensitive and naturalistic measure such as monitoring eye movements, as well as conducting the study using the modality of reading rather than listening, will yield different results, perhaps highlighting a small but significant difference in the time course of detection of semantic as compared to pragmatic anomalies.

The most relevant adult study to the present experiment was an investigation by Rayner, Warren, Juhasz and Liversedge (2004) on the effects of thematic plausibility on adults' reading behaviour (although see also Braze et al., 2002; Ni et al., 1998; Warren & McConnell, 2007; Warren et al., 2008). As mentioned in Chapter 1 Rayner et al. used sentences that described events in which an individual performed an action with an instrument. In each case, the verb had three thematic roles (see sentences 2a-c): an agent (*John*), an instrument (*knife*, *axe* or *pump*) and a patient/theme (*carrots*).

2a. John used a knife to chop the large carrots for dinner.

2b. John used an axe to chop the large carrots for dinner.

2c. John used a pump to inflate the large carrots for dinner.

In all three sentences, the instrument (*knife*, *axe* or *pump*) could be plausibly used in conjunction with the main verb (*chop* or *inflate*). However, in the implausible condition (2b) the patient (*carrots*) was incongruous as the object of the verb (*chop*) given the particular instrument used (*axe*). That is, although axes are often used to chop things, and carrots are often chopped, an axe is not often used to chop carrots. To this extent, the sentence is implausible (though it does not describe a situation that is virtually impossible). By contrast, in the anomalous condition (2c), the patient (*carrots*) could not be used in conjunction with the verb (in this case *inflate*) since under ordinary circumstances, carrots do not have the attribute of inflatability. However, the attribute of inflatability in carrots may be considered possible in certain circumstances such as cartoons, and in this context anomaly effects can be eliminated (Filik, 2008), although some research shows that the elimination is not immediate (Warren et al., 2008).

A crucial difference between the implausible and anomalous conditions was that in the implausible condition, the incongruity occurred due to the incompatibility of two objects involved in the event denoted by the main verb. By contrast, in the anomalous condition, the incongruity arose due to the sentential object being an inappropriate argument of the verb. Rayner et al. characterised the former manipulation as describing an implausible situation, whereas the latter manipulation describes an impossible, or extremely unlikely event, and was categorised as anomalous. This distinction is again comparable to that made by Marslen-Wilson et al. The anomalous condition equates to a semantic anomaly, while the implausible condition equates to a pragmatic anomaly, and requires knowledge of the real world in order to detect the implausibility.

In the Rayner et al. study, disruption to processing occurred earlier when the sentences were anomalous rather than implausible. Specifically, for anomalous sentences, disruption was apparent in first pass reading time as soon as the critical word *carrots* was fixated; whereas in the implausible condition, disruption was less immediate and only apparent for fixations made after a regressive saccade from the critical word. Rayner et al. suggested that the differential effects may be due to either the severity of the violation (i.e. how implausible it is perceived to be), or because, in most cases, anomalous violations can be detected on the basis of lexical information alone (a verb argument violation - Mauner et al., 1995). Implausible violations, on the other hand, can only be detected at a later stage of processing after the semantic evaluation of the combination of a verb and the objects involved in the event it denotes.

7.1.3 Parafoveal-on-foveal effects

A second area of eye movement behaviour that was investigated by Rayner et al. was parafoveal-on-foveal processing. The extent and depth to which parafoveal words are processed prior to direct fixation has received considerable attention in the field of eye movement research in recent years, largely because the question of whether parafoveal words are lexically identified prior to fixation has direct implications for whether words are identified serially (e.g., the E-Z Reader model, Pollatsek et al., 2006; Reichle et al., 1998; Reichle et al., 2004); or in parallel (e.g., the SWIFT model, Engbert et al., 2002; Engbert et al., 2005), and this has provoked fierce debate (e.g. Kennedy & Pynte, 2008; Rayner, Pollatsek, Liversedge, & Reichle, 2009). Although models such as SWIFT, in which words are identified in parallel according to an attentional gradient, can explain parafoveal-on-foveal effects quite easily (although see Reichle et al., 2009), these findings are less readily accommodated by serial attention shift models. As discussed in Chapter 2, the E-Z Reader model specifies that $word_{n+1}$ can be identified while the reader is still fixating $word_n$, and under such circumstances $word_{n+1}$ would then be skipped. However, when full identification of $word_{n+1}$ does not occur while $word_n$ is still fixated, the linguistic properties of a non-fixated word (i.e., $word_{n+1}$) should not influence fixation durations on $word_n$. In this way, the E-Z Reader model specifically rules out the possibility that words are lexically identified in a non-sequential (i.e. parallel) fashion. Thus, any evidence of parafoveal-on-foveal effects that cannot be attributed to oculomotor error would provide a significant challenge to the E-Z Reader model.

Previous research investigating lexical-level parafoveal-on-foveal effects has been somewhat inconsistent (see Carpenter & Just, 1983; Henderson & Ferreira, 1993; Hyönä & Bertram, 2004; Inhoff, Radach et al., 2000; Inhoff & Rayner, 1986; Inhoff, Starr et al., 2000; Kennedy, 2000;

Kennedy et al., 2004; Murray & Rowan, 1998; Murray, 1998; Pynte et al., 2004; Rayner, White et al., 2003). Rayner et al. (2003) observed an influence of $word_{n+1}$ on fixation durations on $word_n$ only for those fixations on $word_n$ that were very close to the end of that word (i.e., within three characters of $word_{n+1}$). Consistent with the serial attention shift position, Rayner et al. interpreted these effects as resulting from oculomotor error, arguing that the inflated fixations on $word_n$ were a consequence of mislocated fixations due to saccadic undershooting (Nuthmann, Engbert, & Kliegl, 2005) rather than genuine parafoveal-on-foveal effects. That is, they argued that readers intended to make a saccade to $word_{n+1}$, but the saccade fell short of the intended target and therefore fixation durations reflected processing of this word. Experiment 7 will provide a further opportunity to investigate the possibility that semantic processing of a parafoveal word can take place and whether it affects the current fixation duration.

7.1.4 Research with child readers

In common with other aspects of linguistic processing, there has been very little research examining thematic processing in children. However, there is one study (Nation et al., 2003) which is particularly relevant (see also Brock, Norbury, Einav, & Nation, 2008) as it has used eye movements to investigate children's assignment of thematic roles, though it was not a reading experiment. Other studies have investigated children's processing of plausibility information but in terms of how it affects their ongoing syntactic processing (e.g. Hurewitz et al., 2000; Tanenhaus et al., 1995; Trueswell et al., 1999), and will not be discussed here (see Chapter 6, Section 6.1.2 for a description of these studies).

Nation, Marshall and Altmann (2003) conducted an experiment in which they monitored children's eye movements to objects in a visual scene as they listened to spoken sentences such as (3a) and (3b):

(3a) Jane watched her mother eat the cake.

(3b) Jane watched her mother choose the cake.

On hearing the target verb (*eat*) of sentence (3a), children made fast anticipatory eye movements towards the target object on the screen (*cake*) when that object was the only edible entity. In contrast, on hearing sentence (3b), children did not move their eyes until they heard the word (*cake*) when all entities in the visual display were 'choosable' (i.e. when the verb's selectional restrictions did not confine the possible object of the verb to a single entity in the display). Nation et al. argued that these findings clearly indicate that children, like adults, are sensitive to verb selection restrictions and are able to integrate this with information extracted from the visual context with the same accuracy and speed as adults. That is, although Nation et al. did not make direct comparisons between adults and children, a previous study (Altmann & Kamide, 1999) showed that adults also made eye movements towards the target object on hearing a verb which allowed only one entity as its direct object in the display. Nation et al.'s results therefore show that thematic role information is quickly activated and assigned on the basis of plausibility or thematic fit in children in the visual world paradigm. Whether or not these effects generalise to text reading is the focus of the experiment reported in this chapter.

7.1.5 Predictions

To summarise the literature, in Rayner et al.'s adult study, readers exhibited earlier disruption to processing of sentences containing anomalous as compared to implausible thematic relations. On this basis, it was anticipated that adult readers in the experiment reported here would spend longer fixating the critical region of the anomalous than the implausible or control sentences during first pass reading. Also in line with Rayner et al.'s findings, it was predicted that adults would exhibit delayed disruption to processing for the implausible sentences as compared to the anomalous sentences, such that disruption may first occur within second rather than first pass fixations. Additionally, direct comparisons between the adult data from Rayner et al.'s and the adult data from the experiment reported here were made in order to establish that the Rayner et al. results replicate. For this reason, the adult data were analysed separately, as well as in conjunction with, the child data. Finally for the adults, data were analysed in order to examine further parafoveal-on-foveal effects of anomaly and implausibility.

In terms of the children, the predictions were a little more tentative. The vast majority of studies that have investigated whether children show a sensitivity to semantic information have manipulated semantic (visual or linguistic) context to examine initial syntactic processing preferences (e.g., Trueswell et al., 1999). Semantic influences on processing are inferred on the basis of initial syntactic commitments. The only study that has examined children's online processing of plausibility information per se is that of Nation et al. (2003), and they found no differences in processing between children and adults (in a separate experiment). However, given that Nation et al.'s experiment was not a reading experiment, with a highly constrained visual context that made verb arguments relatively predictable, it may be unwise to assume that children and adults read sentences containing implausibilities and anomalies in the same way. This is particularly the case given that it is known that children are slower to process written language generally as compared to adults (Rayner, 1986), and given that Experiments 5 and 6 showed that children were delayed relative to adults in their detection of syntactic misanalysis.

Thus, on the assumption that linguistic influences will occur with less immediacy in the eye movement records of children compared with adults, and following the results from Experiments 5 and 6, it was anticipated that children might exhibit delayed disruption to processing when reading sentences containing anomalous and implausible thematic relations. Additionally, while adults were expected to show first pass effects of anomaly, it was predicted that similar effects may be delayed for children. Specifically, it was anticipated that effects for children would be spatially localised to words downstream from the critical word, or fixations made during second pass reading (after a regressive saccade and later in the eye movement record than for adults). Similarly, it was also anticipated that implausibility effects would be delayed for children relative to adults. It may even be that children exhibit no disruption to processing in response to the implausible sentences if, as suggested by McRae, Ferretti and colleagues (Ferretti et al., 2003; Ferretti, Kutas, & McRae, 2007; Ferretti et al., 2001; McRae et al., 1997), they do not yet have sufficient experience of the co-occurrence of verbs and their arguments to have stable expectancies regarding what constitutes a plausible or implausible thematic role. As with the adult readers, the child data were analysed in order to examine further parafoveal-on-foveal effects of anomaly and implausibility.

7.2 Method

7.2.1 Participants

Twenty-four adults and twenty-four children took part in the experiment. Adults and children were drawn from the same participant pools as in Experiment 1 (Chapter 3, Section 3.2.1). The mean age of the child participants was 9 years and 6 months, with ages ranging from 7 years 0 months to 12 years 0 months. As in all other experiments, child participants completed the Wechsler Objective Reading Dimension (WORD; Rust, Golombok & Trickey, 1992) as a measure of their reading ability. The average reading age was 12.1 years.

7.2.2 Apparatus

The same apparatus was used as in Experiment 1 (Chapter 1, Section 3.2.3), but in this experiment the viewing distance was 80cm rather than 100cm.

7.2.3 Materials

Thirty-six experimental items were constructed, and for each of these, there were three versions of each sentence (see sentences 4a-c below for control, implausible and anomalous versions respectively).

(4a) Robert used a trap to catch the horrible mouse that was very scared.

(4b) Robert used a hook to catch the horrible mouse that was very scared.

(4c) Robert used a radio to play the horrible mouse that was very scared.

As in Rayner et al.'s study, all items were designed so that the plausibility violation always occurred at the noun of the adjectival noun phrase (the critical target word) following the infinitival verb. The critical word (*mouse* in sentences 4a-c) was identical across conditions (as were the two immediately preceding words), which ensured that any effects observed at this point were unlikely to be due to factors other than the plausibility/anomaly manipulation. On average, there were no significant differences in length (M = 4.42, SD = 1.05 (anomalous condition); M = 4.56, SD = 1.00 (implausible and control conditions)) or frequency (M = 95.8, SD = 82.6 (anomalous condition; M = 153.1, SD = 185.4 (implausible and control conditions)) of the infinitival verb across conditions (*catch* or *play*), ps > .09. Similarly, there was no significant difference in the frequency of the noun (*trap*, *hook* or *radio*) denoting the instrument across conditions (Anomalous: M = 27.1, SD = 51.2; Implausible: M = 15.4, SD = 18.1; Control: M = 30.9, SD = 53.0; p > .5). There was a significant difference in length (F (1, 35) = 5.2, p < .01) of this same noun. In the anomalous and control conditions, words were respectively 5.7 (SD = 2.4) and 5.9 (SD = 2.5) characters long on average whereas in the implausible condition, they were 6.5 (SD = 2.6) characters. However, this difference is in a part of the sentence prior to the regions of interest.

Questionnaires containing all 108 sentences (i.e. all three versions of each sentence) were given to 19 undergraduate students (who did not take part in the experiment proper). Eighty-two children also completed a 36-item questionnaire (the same as that given to adult participants but a third of the length to make the task easier for the children). There were two age groups of children: younger children (age 7-8 years) and older children (age 10-11 years). All children were recruited from three local primary schools and did not take part in the main experiment.

Analyses of adult and child ratings were carried out separately. Participants were asked to rate the sentences on a 5-point scale where 1 was 'very strange', 3 was 'a bit strange' and 5 was 'not at all strange / normal'. The mean ratings for each condition, together with the standard deviations, are shown in Table 7.1.

Table 7-1: Adults' and children's mean ratings on a 1-5 scale for sentences in the Anomalous, Implausible and Control conditions in an offline questionnaire. Standard deviations in parentheses.

	Adults	Younger children	Older children
Anomalous	1.55 (*0.51*)	1.98 (*0.74*)	1.71 (*0.51*)
Implausible	2.44 (*0.68*)	2.04 (*0.98*)	2.02 (*0.72*)
Control	4.83 (*0.16*)	4.44 (*0.43*)	4.41 (*0.39*)

For the adults, there was a significant difference between ratings for the three groups of sentences, $F(2, 70) = 458.3$, $p < .001$; adults rated the anomalous sentences as being significantly stranger than the implausible sentences, $t(1, 35) = 6.84$, $p < .001$; the anomalous sentences as being significantly stranger than the control sentences, $t(1, 35) = 38.44$, $p < .001$; and the implausible sentences as being significantly stranger than the control sentences, $t(1, 35) = 20.72$, $p < .001$. For the children, there was also a significant difference between ratings for the three conditions, $F(2, 70) = 212.5$, $p < .001$; children reliably rated the anomalous sentences as stranger than the implausible sentences, $t(1, 35) = 2.7$, $p < .05$, the anomalous sentences as stranger than the control sentences, $t(1, 35) = 19.5$, $p < .001$, and the implausible sentences as stranger than the control sentences, $t(1, 35) = 16.9$, $p < .001$. There was no significant difference between the ratings of the younger as compared to the older children, $p > .6$. These results indicate that in an offline task, children were able to perceive a difference between all three conditions and this served to validate the plausibility manipulation employed in Experiment 7 across both participant populations.

7.2.4 Procedure

The procedure was identical to that in Experiment 1 (Chapter 3, Section 3.2.4). In addition, participants were warned that some of the sentences were somewhat strange but that they should read the sentences as normally as possible.

7.3 Results

All sentences were divided into five regions as shown in Example 4c (repeated here), three of which were of particular interest.

REGION 1 REGION 2 REGION 3 REGION 4 REGION 5

(4c) /Robert used a radio to play/ the horrible/ mouse/ that was/ very shy. /

The three regions of particular interest were the pre-target region (Region 2) which comprised the determiner and adjective (e.g., *the horrible*); the target word region (Region 3) comprising the noun (e.g., *mouse*); and the post-target region (Region 4) comprising the two short words (or one long word), following the target noun (e.g., *that was*).

All participants performed at a minimum of 75% accuracy on the comprehension measure. Adult had a mean accuracy of 89% and children scored 97% of questions correctly on average. Trials were excluded if (i) the participant did not fixate three or more regions in total, (ii) the participant did not fixate the first two regions or (iii) the participant did not fixate two of the three regions of interest (regions 2, 3 and 4). Furthermore, a proportion of the child participants did not complete all the trials due to fatigue or tracker loss. On this basis, 37 trials were eliminated (4.4% of trials) from the adult data and 129 trials (14.9% of trials) were eliminated from the children's data (exclusions were equally distributed across conditions). As in all experiments in this thesis, fixations less than 80 ms were combined with fixations on adjacent letters or else eliminated if they were not within 3 characters of another fixation. Fixations longer than 1200 ms were also excluded.

The following eye movement measures were computed: first fixation duration, gaze duration, go past reading time, and total reading time. Two (group: adults vs. children) x 3 (plausibility: anomalous vs. implausible vs. control) mixed design ANOVAs were carried out for analyses of the pre-target region, the target region, and the post-target region.

7.3.1 Pre-target region

Table 7-2: Mean fixation times (ms) in the pre-target region for anomalous, implausible and control conditions. Standard deviations in parentheses.

	Adults			Children		
	Anomalous	Implausible	Control	Anomalous	Implausible	Control
First fixation duration	245 (*65*)	241 (*75*)	238(*59*)	285 (*133*)	287 (*109*)	292(*107*)
Gaze duration	371 (*161*)	366 (*153*)	338 (*139*)	551 (*360*)	539(*320*)	518 (*318*)
Go past reading time	433 (*224*)	424 (*210*)	405 (*281*)	734 (*817*)	642 (*466*)	624 (*461*)
Fixation duration prior to Region 3	252 (*69*)	244 (*71*)	242 (*61*)	309 (*141*)	297 (*128*)	291 (*122*)
Last 3 characters	251 (*65*)	251 (*70*)	240 (*64*)	327 (*158*)	304 (*111*)	302 (*122*)
Total reading time	572 (*295*)	512 (*285*)	451 (*254*)	873 (*580*)	677 (*429*)	678 (*466*)

Table 7.2 shows the reading time measures for the pre-target region. There were no effects of the plausibility manipulation on the duration of the first fixation (ps > .8), but there was a reliable effect of group, F1 (1, 46) = 13.05, p < .005; F2 (1, 33) = 131, p < .001: children made longer first fixation durations as compared to adults. Again, note children's fixation durations were consistently longer than those of adults across measures and regions in all instances in which there was an effect of group. There was no reliable interaction between group and plausibility (ps > .3). In gaze durations, there was a suggestion of an effect of the plausibility manipulation (a difference of 33ms), but the effect was not reliable, F1 (2, 92) = 2.94, p = .058; F2 (2, 66) = 1.55, p = .2, and there was no reliable interaction (ps > .8). Once again, there was an effect of group F1 (1, 46) = 14.95, p < .001; F2 (1, 33) = 129, p < .001. These results replicate those of Rayner et al. who found no first pass effects of the plausibility manipulation, although they did find a numerical difference between the anomalous, implausible, and control conditions.

Although the effect of the plausibility manipulation for gaze duration was not reliable by items, there was a 33 ms difference in the predicted direction between the anomalous as compared to the control condition for both adults and children. Following Rayner et al., some fine grain analyses to investigate parafoveal-on-foveal effects were conducted. First, all fixations prior to the first fixation made in the target region were examined. For these fixations, there was no reliable effect of plausibility, $F1\ (2, 92) = 1.52, p = .23$; $F2\ (2, 70) = 2.15, p = .12$, although there was a numerical trend in the predicted direction: the difference in prior fixations in the anomalous compared to the control (18 ms) was larger than the difference between the implausible and control conditions (7 ms). There was a reliable effect of group, $F1\ (1, 46) = 11.00, p < .005$; $F2\ (1, 35) = 93.9, p < .001$, but no interaction ($Fs < 1$).

To examine this tendency in more detail, exclusively those fixations made within three characters of the target region were then considered, on the assumption that such fixations would allow the reader a good preview of the anomalous or implausible $word_{n+1}$. A marginal effect of plausibility was obtained, $F1\ (2, 84) = 2.34, p = .10$; $F2\ (1.6, 52) = 2.39, p = .10$; participants made marginally longer fixations in the anomalous as compared to the control condition marginal by items, $t2\ (27) = 2.00, p = .055$, although this was not significant by participants ($p > .19$). There was no reliable difference between fixation durations in the implausible and control conditions ($ps > .5$). In summary, there were consistent numerical differences in fixation durations on the fixation preceding the target word and when that fixation was within three characters of the target; differences between the anomalous and control conditions were greater than those between the implausible and control conditions (though these differences missed statistical significance). The data from these analyses are entirely consistent with those reported by Rayner et al., and also with their explanation of a saccadic overshoot.

Finally in the pre-target region, total reading time showed a highly significant effect of plausibility, $F1\ (2, 92) = 21.87, p < .001$; $F2\ (2, 70) = 16.4, p < .001$. Participants spent longer overall reading the anomalous as compared to the control sentences, $t1\ (47) = 5.59, p < .001$; $t2\ (35) = 4.96, p < .001$, but there was no difference in reading time between the implausible and control sentences ($ts < 1.4, p > .1$). Once again, there was also a significant effect of group, $F1\ (1, 46) = 12.96, p < .005$; $F2\ (1, 35) = 87.08, p < .001$ with longer total reading times for children than for adults. Finally, there was an interaction between plausibility and group that was statistically significant by items and marginal by participants, $F1\ (2, 92) = 2.79, p = .07$, $F2\ (2, 70) = 4.88, p < .05$.

To explore this interaction, the adult and child data were analysed separately. Adults showed effects of anomaly and implausibility in total reading time in the pre-target region, $F1\ (2, 46) = 19.09$, $p < .001$; $F2\ (2, 70) = 13.44$, $p < .001$; they spent longer reading the pre-target region in the anomalous as compared to the control condition, $t1\ (23) = 6.82$, $p < .001$; $t2\ (35) = 5.74$, $p < .001$, and the implausible as compared to the control condition, $t1\ (23) = 2.98$, $p < .01$; $t2\ (35) = 2.84$, $p < .01$. Given that this measure includes second pass fixations, it indicates that adults ultimately detected both the anomaly and the implausibility. The results were somewhat different for the children. Like adults, they showed a reliable effect of the plausibility manipulation in total reading time, $F1\ (2, 46) = 11.22$, $p < .001$; $F2\ (2, 70) = 11.41$, $p < .001$, but only the difference between the anomalous and control conditions was reliable, $t1\ (23) = 3.75$, $p < .005$; $t2\ (35) = 3.73$, $p < .005$. Total reading times were no different in the implausible as compared to the control condition ($ps > .7$). Thus, while adults showed effects of both anomaly and implausibility, children were influenced only by the anomaly manipulation. Rayner et al. found an anomaly effect, but not an implausibility effect in total reading time with their adult participants. It is not clear why this effect occurred in the experiment reported here, but not in that of Rayner et al. However, since the effect occurred in the relatively late measure of total reading time, it should not be taken as an indication of initial detection of anomaly or implausibility.

To summarise the results from the pre-target region, although first pass effects (in gaze durations) in the pre-target region were not reliable, a 33 ms difference in gaze durations between the anomalous and control conditions was observed (16ms longer than in the Rayner et al study). Marginal parafoveal-on-foveal effects of plausibility were also found, suggesting that readers were experiencing some disruption to processing at this early stage, although not reliably so. These findings are consistent with the Rayner at al. data. In total reading time reliable effects of anomaly and implausibility were obtained for adults; however, for children the only reliable effects obtained were for anomalies. As indexed by longer total reading times, children showed no sensitivity to the implausible sentences. Finally, there were consistent effects of group in every measure, showing that children took longer to read sentences than adults.

7.3.2 Target region

Table 7-3: Mean fixation times (ms) in the target region for anomalous, implausible and control conditions. Standard deviations in parentheses.

	Adults			Children		
	Anomalous	Implausible	Control	Anomalous	Implausible	Control
First fixation duration	273 (*92*)	263 (*75*)	258 (*81*)	317 (*126*)	311 (*122*)	306 (*129*)
Gaze duration	324 (*141*)	299 (*111*)	300 (*110*)	428 (*264*)	406 (*204*)	417 (*254*)
Go past time	395 (*226*)	369 (*211*)	353 (*185*)	604 (*602*)	541 (*512*)	526 (*457*)
Total time	507 (*311*)	428 (*250*)	394 (*244*)	645 (*473*)	522 (*364*)	552 (*440*)

Table 7.3 shows the reading time measures for the target region. As per Rayner et al., for first fixation durations, despite a numerical trend in the predicted direction (anomaly effect = 13 ms; implausibility effect = 5 ms), there was not a reliable effect of plausibility (ps > .1). There was a significant effect of group, F1 (1, 46) = 11.61, p < .005; F2 (1, 35) = 51.47, p < .001, with longer first fixations for children than for adults, but no interaction (Fs < 1). In gaze durations, there was an effect of plausibility that was marginal by participants but not reliable by items, F1 (2, 92) = 2.62, p = .07; F2 (2, 70) = 1.58, p > .2. There was also an effect of group F1 (1, 46) = 15.6, p < .001; F2 (1, 35) = 117, p < .001, but no reliable interaction (Fs < 1). When the adult data were analysed separately in order to replicate the Rayner et al. data, there was a reliable effect of the plausibility manipulation in gaze durations, F1 (2, 46) = 4.57, p < .05; F2 (2, 70) = 4.31, p < .05. Adults exhibited longer gaze durations on anomalous than control words, t1 (23) = 2.31, p < .05; t2 (35) = 2.21, p < .05, but showed no difference between implausible and control words (ps > .7).These effects are consistent with the results from the Rayner et al study.

In go past time, there was an effect of the plausibility manipulation which was reliable by items but not by participants, F1 (2, 92) = 1.46, p > .2; F2 (2, 70) 3.61, p < .05. There was an effect of group F1 (1, 46) = 14.4, p < .001; F2 (1, 35) = 64.5, p < .001, with children showing longer go past reading times than adults. There was no reliable interaction (ps > .5). Again, this pattern of data replicates that of Rayner et al. who also obtained no significant effect of implausibility in the go past measure in the target region (although like in the present experiment they did obtain a numerical trend).

As per Rayner et al., total reading time was influenced by the plausibility manipulation, F1 $(2, 92) = 13.28$, $p < .001$; F2 $(2, 70) = 15.05$, $p < .001$; readers spent longer reading the anomalous as compared to the control sentences, t1 $(47) = 4.34$, $p < .001$; t2 $(35) = 3.73$, $p < .005$, but showed no difference in total reading time between the implausible and the control sentences ($ps > .3$). There was a main effect of group, F1 $(1, 46) = 8.37$, $p < .01$; F2 $(1, 35) = 41.57$, $p < .001$, with longer total reading times for children than adults. There was no interaction between group and the plausibility manipulation ($Fs < 1.2$, $ps > .3$).

In summary, in the target region, there were first pass effects (in gaze durations) of the plausibility manipulation with adults looking 24 ms longer at the target word in the anomalous as compared to the control condition, and children looking 11 ms longer, and both adults (42 ms) and children (78 ms) taking longer to go past the target word in the anomalous as compared to the control condition. There were also reliable anomaly effects in total reading time. There were no effects of implausibility in the target region. These findings fully replicate those of Rayner et al.

7.3.3 Post-target region

Table 7-4: Mean fixation times (ms) in the post-target region for anomalous, implausible and control conditions. Standard deviations in parentheses.

	Adults			Children		
	Anomalous	Implausible	Control	Anomalous	Implausible	Control
First fixation duration	273 (*100*)	273 (*115*)	263 (*110*)	292 (*115*)	297 (*119*)	297 (*147*)
Gaze duration	376 (*232*)	388 (*251*)	373 (*237*)	481 (*316*)	494 (*332*)	466 (*314*)
Go past time	697 (*736*)	595 (*565*)	491 (*399*)	968 (*1040*)	693 (*550*)	705 (*735*)
Total time	593 (*377*)	566 (*429*)	500 (*300*)	751 (*500*)	669 (*424*)	611 (*461*)

Table 7.4 shows the reading time measures for the post-target region. In first fixation durations, once again there was no effect of plausibility ($ps > .4$) and no interaction ($ps > .5$). There was an effect of group, reliable by items but not by participants, $F1\ (1, 46) = 2.15$, $p = .15$; $F2\ (1, 34) = 16.09$, $p < .001$. In gaze durations, there was no effect of plausibility ($ps > .1$). There was a reliable effect of group, $F1\ (1, 46) = 7.61$, $p < .01$; $F2\ (1, 34) = 61.25$, $p < .001$, but no interactive effect ($ps > .8$). These effects differ from those reported by Rayner et al. who found a significant effect of anomaly in both first fixation durations, and gaze durations, in the post-target region. The failure to obtain reliable effects in this region for these measures may be a consequence of the additional variability within the present data set introduced by the children.

In go past time, there was a significant effect of plausibility, $F1\ (2, 92) = 11.23$, $p < .001$; $F2\ (1.6, 68) = 9.50$, $p < .005$ (see Figure 7-1); go past times were longer in the anomalous than the control condition, $t1\ (47) = 3.97$, $p < .001$; $t2\ (34) = 3.42$, $p < .005$, but there were no differences between the implausible and control conditions ($ps > .2$). There was an effect of group, $F1\ (1, 46) = 8.84$, $p < .01$; $F2\ (1, 34) = 56.9$, $p < .001$, and an interaction between group and plausibility, marginal by items but not reliable by participants, $F1\ (2, 92) = 1.72$, $p = .19$; $F2\ (2, 68) = 2.35$, $p = .10$. As noted, to directly compare the adult data with those from the Rayner et al. study, the data from the adults were analysed separately. Go past times were reliably longer for adults in the anomalous as compared to the control condition, $t1\ (23) = 3.89$, $p < .005$; $t2\ (34) = 3.44$, $p < .005$, and importantly in the implausible as compared to the control condition (although marginal by participants), $t1\ (23) = 1.84$, $p = .079$; $t2\ (34) = 2.24$, $p < .05$. This is the earliest point at which adults exhibited effects of implausibility, and these data replicate Rayner et al's findings.

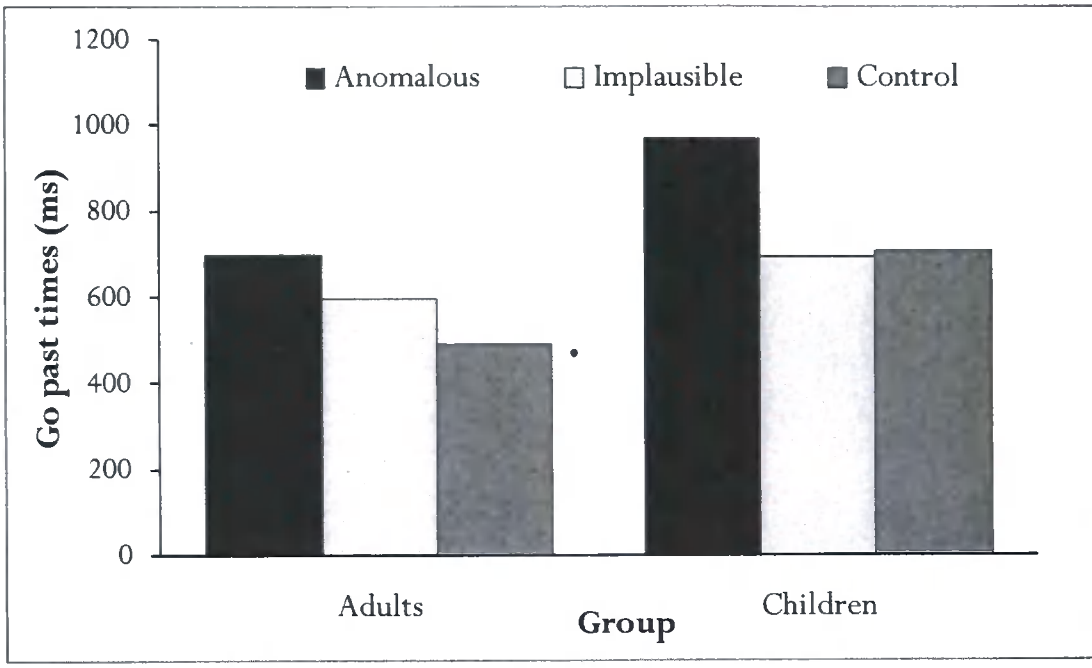

Figure 7-1: Go past reading times for adults and children in the post-target region in the anomalous, implausible and control conditions

Finally, in total reading times there was an effect of the plausibility manipulation, F1 (2, 92) = 10.83, p < .001; F2 (2, 70) = 7.26, p < .005; participants spent longer reading the anomalous as compared to the control sentences, t1 (47) = 5.03, p < .001; t2 (35) = 3.30, p < .005, and the implausible compared to the control sentences, t1 (47) = 2.84, p < .01; t2 (35) = 2.74, p < .05. There was also an effect of group, F1 (1, 46) = 5.59, p < .05; F2 (1, 35) = 47.55, p < .001, but no interaction (ps > .3). These findings indicate that while earlier reading time measures showed differential effects of plausibility, ultimately both adults and children exhibited disruption to processing in the implausible condition. This is the first indication of an implausibility effect in children and shows that children, as well as adults, were able to discriminate between both anomalous and implausible thematic violations during reading. Again, the adult data replicate those of Rayner et al. who also obtained reliable anomaly and implausibility effects in total times in the post-target region.

To summarise the results in the post-target region, go past times revealed disruption to processing as a result of the plausibility manipulation. While adults spent longer reading the anomalous and the implausible as compared to the control sentences (as per Rayner et al.), children showed a difference between the anomalous and control conditions only. This differential pattern of effects for adults and children is the same pattern observed in the total reading time measure in the pre-target region. In contrast, total reading times in the post-target region revealed both anomaly and implausibility effects for both adults and children, indicating that both participant groups had detected the implausible thematic violation in the latter stages of reading the sentences.

7.3.4 Summary of Results

Overall, the earliest point at which reliable disruption to processing was observed in response to the anomalous thematic violation was in the target region in gaze duration. Although the effect failed to reach statistical significance in the overall analyses, adults exhibited a reliable effect of anomaly, and children showed a numerical trend in the predicted direction. Reliable implausibility effects were observed only in the much later measure of total reading time in the post-target region (although adults showed implausibility effects in go past times in the post-target region as per Rayner et al.). As predicted, increased immediacy in the detection of anomalous as compared to implausible thematic violations was therefore observed. Interestingly, reliable differences in the time course of anomaly detection between adults and children were not observed. Although the magnitude of the effects was stronger for adults, both adults and children exhibited disruption to processing during first pass in the anomalous condition in the target region. There were, however, differences in time course of the implausibility detection: in go past times in the post-target region, adults' processing was disrupted by both the implausibility and anomaly, while children showed anomaly effects only. This same pattern of effects was observed in total times in the pre-target region, and an implausibility effect in children was observed only in total times in the post-target region. It appears that although children were delayed in their processing of implausibility as compared to adults, both adults and children showed increased immediacy in their anomaly as compared to their implausibility detection, and exhibited disruption to processing in response to both kinds of manipulation.

7.4 Discussion

There were both commonalities and differences in the time course of processing of the different kinds of thematic relations in adults and children. It was predicted that adults would exhibit earlier disruption to processing of sentences containing anomalous as compared to implausible thematic relations. This prediction was met as adults exhibited disruption to processing during first pass in the anomalous condition in the target region, but showed only later effects of implausibility. In line with Rayner et al., this finding shows that anomalous thematic relations are detected more immediately than implausible ones (see also Warren & McConnell, 2007). It appears that the increased severity of an anomalous (as compared to an implausible) thematic violation induces more immediate and substantive disruption in the eye movement record.

It was also predicted that disruption to processing would be delayed in children when reading sentences containing anomalous and implausible thematic relations, as compared to adults. This prediction was only partially met. Although the adults showed reliable first pass effects of anomaly and those observed for the children failed to reach statistical significance, it seems clear that both groups did show consistent and substantial reading time differences between the anomalous and control sentences during first pass. It is argued then, that the time course of processing associated with thematic anomaly detection is similar in adults and children. This finding seems to fit well with the Nation et al. study (2003) which found that the time course for identifying a plausible (as compared to anomalous distractors) verb argument was immediate, and did not differ for child and adult participants.

Importantly, however, while children did not differ from adults in their anomaly detection, they did appear to be delayed relative to adults in their implausibility detection. While disruption to processing in adults was observed in the implausible condition compared with the control condition for go past reading times in the post-target region (and in total time in the pre-target region), children showed no similar difference in reading times between these two conditions in these measures. Only in the total reading times in the post-target region did children first show a reliable effect of implausibility. These results indicate that although children do detect implausible thematic violations during reading, they are delayed in doing so relative to adults.

Presumably, an anomalous thematic violation may be detected through the recognition of an illegal combination of linguistic constituents. By contrast, the detection of a thematic implausibility relies both on thematic processing, as well as the integration of real world knowledge and pragmatic information with respect to the discourse representation that is being developed on an ongoing basis by the reader. If this is the case, then it appears that children are similar to adults in terms of basic thematic processing during comprehension, but are less efficient than adults in the integration of real world knowledge into the discourse representation. It may be for this reason that children are delayed in their detection of thematic implausibilities relative to adults.

Such an interpretation sits well with previous research into children's online language processing using different methodologies that have found that children are less able to use contextual information to guide parsing (e.g. Trueswell et al., 1999). It also sits well with the distinction made by Marslen-Wilson et al (1988) between semantic and pragmatic anomalies. If violations in the anomalous condition could be detected on the basis of purely linguistic information while the violations in the implausible condition relied on inferences about the real world, then anomalous violations could be detected immediately within the language processor, whereas in the implausible condition theta roles may have been evaluated for their plausibility by referring to real world knowledge following initial assignment, thereby leading to delayed detection. By this argument, the assignment of implausible, but not anomalous, thematic roles would fit well with McCrae and colleagues'(e.g. Ferretti et al., 2001) characterisation of how knowledge of thematic roles is developed. If this is correct, it can be argued that the linguistic mechanisms in place to assign thematic roles to verbs are the same in adult and child readers. However, when outside knowledge about the plausibility of certain events taking place in the real world need to be integrated into an ongoing representation of the sentence meaning, children are less efficient at this aspect of processing and therefore exhibit delayed effects in the implausible condition only.

A limitation of the experiment is that real world plausibility could account for differences in time course of the detection of the anomalous as compared to implausible thematic relations. As in the example in the Marslen-Wilson et al. (1988) study, although it is unusual to bury a guitar (see sentence 1b), it is far more unusual to drink a guitar (see sentence 1c). Therefore, sentences in the anomalous condition in the present experiment were not only linguistically anomalous, but also more pragmatically anomalous, than those in the implausible condition. A graded effect of implausibility therefore exists in the materials used in this experiment in addition to a categorical distinction between semantic and pragmatic anomaly. If pragmatic plausibility were held constant across both the anomalous and implausible conditions, then the present manipulation would be able to discriminate between the time course of accessibility of semantic versus pragmatic anomaly. However, as the experiment stands, pragmatic anomaly, as well as semantic anomaly varies between conditions, and therefore the two types of anomaly cannot be disentangled from one another. A recent experiment (Warren & McConnell, 2007) has addressed this issue and suggests that the magnitude and latency of disruption to linguistic (semantic) violations is greater than that to pragmatic violations (and furthermore these effects are not eliminated by a biasing context, Warren et al., 2008), in keeping with the conclusions made from the current data.

In summary, the key findings from Experiment 7 are as follows. First, to a significant extent, the findings of Rayner et al. were replicated for adult participants. Clear anomaly and implausibility effects were obtained with a similar time course and there was some indication that the effect of anomaly could be detected when the reader fixated just to the left of the target word. These small effects are cautiously interpreted as being primarily due to saccadic undershoots and mislocated fixations. A sensitivity to thematic implausiblities and anomalies in both adults and children was also demonstrated in both an offline rating task and online in eye movements during reading. Additionally, the eye movement records from the children showed thematic anomaly effects that were as immediate as those for adults, whereas, thematic implausibility effects were delayed relative to adults. These data suggest that while children and adults are similar in terms of basic thematic assignment processes that occur during reading, they differ in the efficiency with which they are able to integrate pragmatic and real world knowledge into the discourse representation.

Chapter 8 : Conclusions

This thesis set out to systematically examine children's eye movement behaviour as they read sentences containing specific visual and linguistic manipulations in order to investigate the online cognitive processes that occur as adults and children process written language. The experimental manipulations employed were those already known to generate robust processing preferences in adult readers and could therefore be used diagnostically to examine whether children had the same processing mechanisms in place to deal with different aspects of written language. The final chapter of the thesis will discuss the conclusions that can be drawn from the results of the seven experiments reported. Section 8.1 will discuss the conclusions which can be drawn with regard to children's and adults' visual processing during reading, in particular referring to the findings from Experiment 1 (Chapter 3). Section 8.2 will discuss conclusions which can be made about children's and adults' lexical processing, drawing from Experiments 1, 2, 3 and 4 (Chapters 3, 4 and 5). Section 8.3 will summarise the theoretical implications from the two experiments that investigated syntactic processing (Experiments 5 and 6; Chapter 6), and Section 8.4 will draw conclusions concerning children's and adults' thematic processing (Experiment 7; Chapter 7). Section 8.5 will discuss, more generally, the theoretical ramifications of the thesis as a whole, and Section 8.6 will consider the methodological difficulties that have been encountered in the course of this work and consider their implications for future work. Finally, Section 8.7 will summarise the final conclusions of the thesis.

8.1 Visual processing and oculomotor behaviour during reading

While the majority of experiments in the thesis manipulated some aspect of linguistic processing, Experiment 1 investigated visual processing during reading. This was important because text must be processed visually before linguistic processing can begin and so it was important to establish whether any qualitative differences between adults and children existed at this initial stage of word recognition. Word length was manipulated in order to examine how this visual characteristic of a word affected both temporal and spatial aspects of children's and adults' oculomotor behaviour. Results showed that adults and children were very similar in where they targeted their initial saccade to a word, locating their fixation close to the word centre. Landing position distributions differed slightly with word length, in line with previous (adult) research (McConkie et al., 1988), with both adults and children fixating further into long that short words. Both children and adults used parafoveal information to skip shorter words, and they also made early decisions regarding intra-word refixations, either on the basis of initial fixation location, or perhaps even earlier while programming their initial saccade to a word. In terms of early decisions regarding saccade targeting then, adults and children were strikingly similar. Although it is known that there are some age-related changes that take place in oculomotor behaviour generally (e.g. children have increased saccade latencies in non-reading tasks: Cohen & Ross, 1978), saccadic accuracy, peak saccadic velocity, and saccadic overshoot have all been observed to be the same in children as in adults (e.g. Fukushima et al., 2000; Salman et al., 2006). It appears then that during normal text reading as well, children and adults generally do not differ in their inter-word saccadic targeting.

However, children and adults differed in their refixation behaviour, both in response to changes in word length (children made relatively more refixations on long than short words) and as a consequence of their initial fixation location on a word. Specifically, while both adults and children made more refixations following an initial fixation away from the word centre, children made both progressive and regressive saccades to refixate the word, while adults made mostly progressive saccades. Furthermore, there were some differences that were observed between children and adults in temporal aspects of eye movement behaviour as a result of the word length manipulation. Specifically, children's fixation durations were more inflated by an increase in word length than adults', and they exhibited relatively longer first pass fixations on long than short words.

Overall, the findings showed that adults and children are strikingly similar in their early oculomotor decisions regarding where they direct their initial saccade to a word (or whether they choose to skip a word). However, differences emerged in subsequent fixations on a word: adults made fewer refixations than children, and directed them rightwards, presumably because reading is more efficient when regressive saccades are avoided, while children made more refixations, and these were made in both directions. Such a conclusion supports the argument that children require more and/or longer visual samples of words in order to successfully achieve lexical identification. It is also consistent with previous findings showing that children have less information about word length and letter identity available to them during a fixation than adults (Häikiö et al., 2008; Rayner, 1986).

8.2 Lexical processing

Four experiments investigated lexical-level processing in the thesis: Experiments 1, 2, 3, and 4. First, Experiments 1 - 3 examined word frequency effects in children and adults. Experiment 1 showed that using adult indices of frequency and directly applying them to research with a developmental population does not necessarily generate reliable frequency effects. Although it may be possible to obtain reliable frequency effects in children using adult corpora (see Blythe, Liversedge, Joseph, White, & Rayner, 2009), as outlined in Chapter 4 (Section 4), there are many words which are high frequency as indexed by adult counts but low frequency as indexed by child counts (and vice-versa). The materials used in Experiment 1 did not sufficiently account for this.

Experiment 2 used age-appropriate materials to generate reliable frequency effects in adult readers. Experiment 3 then tested these same adults, as well as a group of children, using a set of sentences in which the frequency manipulation was indexed by frequency counts drawn from age-appropriate texts for children. Together, Experiments 2 and 3 convincingly showed that differential frequency effects are observed for adults and children. That is, while children showed reliable frequency effects when the counts were take from age-appropriate reading material, adults did not exhibit those same effects when reading the same sentences. In sum, while the frequency with which a word is encountered is a fundamental characteristic in the organisation and function of the lexicon in both children and adults, the data across these three experiments clearly showed that the extent to which fixation durations are reduced in children and adults is directly related to the frequency of encounter of a word, and that the frequency of encounter, and the effects that this causes, change systematically with age. Furthermore, indices of adult frequency which have traditionally been used in reading experiments, may not be appropriate for use with developing populations. These findings are important not only for how we understand the structure of the developing lexicon, but also for how we conduct future research with children.

Finally, Experiment 3 showed that word frequency effects exist in children independent of the age at which those same words were acquired in life. That is, although the number of months or years since a word was first learned will clearly have an impact on the cumulative frequency with which it is encountered, and this impact will presumably be quite large for children who have relatively fewer words known to them than adults, and comparatively less time to encounter and re-encounter words, the frequency of encounter of a word nevertheless directly influences the ease with which that word is processed over and above the age at which the word was acquired. In this sense, adults and children are alike in that word frequency affects word recognition processes independent of AoA.

Experiment 4 investigated lexical ambiguity effects, that is, whether words with two or more meanings require greater processing resources than words with just one meaning. While the manipulation employed in this experiment did not generate robust effects, the data were suggestive that, for older children (aged 9-12), there may be a cost associated with processing a lexically ambiguous word than an unambiguous word. If this effect is trustworthy, then it can be taken as a formative indication that older children access and maintain both meanings of an ambiguous word for post-lexical interpretation. While such a strategy may not be optimal in terms of reading efficiency, it would prevent disruption to processing that adults have been shown to exhibit (in previous research, e.g. Duffy et al., 1988) when the less frequent meaning of a biased ambiguous word is contextually instantiated. Importantly, the manipulation employed in Experiment 4 did not produce robust, well-documented ambiguity effects in the adult group (or in the younger children) and so all the data from this experiment, and the resulting interpretations, must be treated with caution. Arguably, the most important conclusions to be drawn from Experiment 4 are methodological, in that the failure to replicate a robust effect is potentially informative regarding the kinds of manipulations that are appropriate to use with developmental populations. These will be discussed in some detail in Section 8.6 of this chapter.

Overall, the findings in relation to lexical processing show that while indices of word frequency (and most likely of other lexical variables such as familiarity) differ for adults and children, both the developing and the mature mental lexicons are functionally organised in terms of the frequency with which words are encountered, and this is independent of the age at which words are acquired. Further research is needed to clarify whether children and adults process words with two meanings in the same way, as well as to investigate further lexical-level effects that have not been examined in this thesis.

8.3 Syntactic processing

Two experiments (Experiments 5 and 6 in Chapter 6) investigated children's and adults' online processing of temporarily syntactically ambiguous sentences. Specifically, these experiments set out to investigate whether children, like adults, preferentially attached a prepositional phrase high to a verb phrase rather than low to a noun phrase (Experiment 5), and whether they preferentially attached an adverbial phrase to the currently-processed verb phrase rather than a verb phrase encountered earlier in the sentence (Experiment 6). If children and adults exhibited disruption when reading sentences containing the same syntactic structure, then it could be postulated that they had the same processing mechanism with the same parsing preferences in place. Results showed that while (younger and older) children were slightly delayed relative to adults in their detection of both types of syntactic ambiguity, they did exhibit the same parsing preferences as adults both online and offline: that is they preferred the syntactically simpler interpretation of the experimental sentences.

Results further showed that once children had detected that their preferred analysis was not the one instantiated in the sentence, the magnitude of the disruption was greater and the effects were also longer-lasting than for adults. This suggests that children, unlike adults, did not revise their initially incorrect interpretation of a syntactically ambiguous sentence. Overall, it appears that children as young as seven years old have the same sentence-parsing mechanism in place as adults, with the same parsing preferences. Importantly, it is at the syntactic level that a difference in the time course of effects was observed. That is, although adults and children are alike in how they initially parse sentences, they are different in the time frame in which they detect an initial syntactic misanalysis.

8.4 Semantic (thematic) processing

The final experiment in the thesis was an investigation of how children and adults process anomalous and implausible thematic relations during reading. Previous research has shown that adults exhibit earlier disruption to processing of sentences containing anomalous, as compared to implausible, thematic relations (Rayner et al., 2004). The results from Experiment 7 show that this is also true of children. Indeed there is not a reliable difference in the time course of anomaly effects in adults and children, showing that thematic anomaly has an immediate disruptive effect on ongoing processing in both age groups. However, children were delayed in their processing of implausible thematic relations as compared to adults, suggesting that while children and adults do not differ in terms of basic thematic assignment processes that occur during reading, they do differ in the efficiency with which they are able to integrate pragmatic and real world knowledge into the discourse representation. Clearly, further research is needed in order to examine the numerous other aspects of semantic processing in children, and whether a difference can be observed between adults and children in the time course and magnitude of these effects as well.

8.5 Overall theoretical conclusions

Brought together, the findings from the seven experiments reported in this thesis indicate that there are both commonalities and differences in how adults and children process written language. Overall, the data suggest that while there is little difference between adults and children in visual and lexical level processing, differences in the time course of effects begin to emerge at higher levels of processing. Children are delayed relative to adults in their detection of syntactic misanalysis, and the efficiency with which they are able to integrate real world knowledge into their discourse representation. It may be then, that at the word level children are as rapid in their processing as adults, but when relationships between words, and knowledge of the world, need to be incorporated into their ongoing representation of the text meaning, children require increasingly more processing resources as compared to adults, and are therefore increasingly delayed in building a coherent representation of sentential meaning.

In Chapter 2 (Section 1.6), some of the more prominent models of eye movements during reading were briefly outlined, more specifically cognitive models such as the E-Z Reader model (Pollatsek et al., 2003; Pollatsek et al., 2006; Reichle et al., 1998; Reichle et al., 2006; Reichle et al., 2004), the SWIFT model (Engbert et al., 2002; Engbert et al., 2005; Kliegl & Engbert, 2003), and oculomotor models (e.g. O'Regan, 1992; O'Regan, 1990; O'Regan & Lévy-Schoen, 1987; Yang & McConkie, 2004; Yang & McConkie, 2001). While these models have successfully accounted for numerous empirical effects observed in adult readers, as yet none has attempted to model data from experiments with children to account for the pattern of eye movements during reading observed in developing populations. The data presented in this thesis may go some way towards beginning this process. While further research is undoubtedly needed, it can now been confidently argued that children, like adults, exhibit strong effects of word length and word frequency and the data from these experiments in particular could be entered into simulations and eventually incorporated into existing models of eye movements during reading. Clearly such an endeavour is far beyond the scope of this thesis, but it is hoped that future collaborative work will address this issue.

8.6 Methodological conclusions

There are several methodological issues which have come to the forefront during the course of this project, and which have implications for future developmental research in the field of eye movements during reading. The first is the issue of the increased variability in a variety of different factors (reading ability, eye movement measures etc.) for children as compared to adults. This is important because greater heterogeneity in children in these factors will impact on their ability to process sentences, and therefore on the likelihood of obtaining reliable effects. As has been noted in several of the experimental chapters, in all measures of eye movements, standard deviations were significantly higher for children as compared to adults. This is true in all of the experiments reported in the thesis, and it is a difficulty which would appear to be inherent in developmental research. Presumably the reason that there is increased variability in children than in adults is that children are in the process of developing their reading skills and there is a lot of variability within this developmental trajectory. In contrast, the adults in the experiments conducted as part of this thesis had already become extremely competent readers (as would be expected of undergraduate students), and therefore there was less variability between participants than in the child groups.

An important question in respect to this issue is whether steps can be taken to address the problem. One potential solution which could be applied to future research concerns the design of the experiments. In the seven experiments carried out for this thesis, and indeed in previous research using eye movements during reading to investigate children's written language comprehension, all were of a cross-sectional design. It is well-documented that there are limitations associated with such studies in relation to developmental research, and given that the increased heterogeneity in developing participant groups is a problem inherent in developmental work, ideally, in the future, studies which adopt longitudinal and cross-sectional approaches in parallel will be carried out. This would allow change to be documented over time within, as well as across, individuals. Such work would strengthen the conclusions that could be formed, and while such a project is beyond the scope of a PhD thesis, it would be an extremely worthwhile and informative way to continue the current line of research.

Second, a greater number of participants (and items) could be tested in order to increase the power of the analyses. It may be that in developmental research, having more participants than would be the norm in an equivalent adult study is simply an added prerequisite. Clearly, controlling for reading ability and having tightly defined age groups would also help to decrease the variability in groups. Third, it may be that only linguistic manipulations which generate very strong effects in adult readers are suitable for use with developing populations. For example, in Experiment 4, which investigated lexical ambiguity effects, reliable effects were not observed (in adults or younger children). It is striking that it was this manipulation, known to produce only relatively small effects in adult readers, for which no effects were obtained. It is, therefore, perhaps unsurprising that this study only produced small or unreliable effects in children. In contrast, those experiments which employed manipulations known to produce large, robust, effects in adult readers (e.g. word frequency and word length) did generate reliable effects. In sum, increased variability in a variety of factors, and the resulting heterogeneity in child participant groups, is an important consideration in any developmental research of this kind, and it contributes to the difficulty of obtaining reliable effects in children. Tightly-controlled, large participant groups, and using manipulations which generate large effects, as well as using a longitudinal design alongside cross-sectional studies, may go some way towards combating the problem.

A second, related methodological issue which merits further discussion is that of using age-appropriate reading material with participants of different age groups. An important premise of this thesis was the use of identical stimuli with all groups of participants (younger children, older children and adults) in order that direct comparisons between eye movement records could be made, and this seems an entirely appropriate experimental approach. However, adults will clearly find sentences which are age-appropriate for young children easier to process than children (and older children easier than younger children), and because eye movements are known to reflect the processing difficulty experienced by the reader, we would expect to see differences in fixation durations and fixation probabilities between different age groups reading identical text (e.g. see McConkie et al., 1991).

It may be that this inevitable difference in processing difficulty for adults and children in the experiments reported has concealed or modulated some of the effects observed, or perhaps more importantly, not observed. Because sentences were necessarily easy for the adults, it may be that a word which was categorised according to the experimental manipulation, as long, low in frequency, or lexically ambiguous, generated only a small increase in processing time. All target words were acquired early in life, were relatively short (target words were never more than eight letters but were considerably shorter in most experiments) and often relatively high in frequency. Therefore they were all processed relatively easily (as compared to, for example, many target words in adult experiments). Furthermore, all sentences were relatively syntactically and semantically simple and so were likely to be read quickly with few regressions. Indeed, in Experiment 2 in which adults read age-appropriate sentences containing a frequency manipulation (for which there none of the usual constraints on AoA, word length etc.), there was a highly significant effect of word frequency. In contrast, in Experiment 1 in which the same corpus (CELEX: Baayen et al., 1995) was used to index word frequency, but in which target words, and sentences in general, were constructed to be suitable for children as young as seven years old, adults failed to exhibit reliable frequency effects. However, despite this difference in processing difficulty between adults and children, in six of the seven experiments reliable effects were obtained, indicating that the manipulations employed were strong enough to overcome the discrepancy.

A possible solution to this problem is for all age groups to read multiple sets of sentences, thereby excluding the confounding variable of processing difficulty across groups (see Häikiö et al., 2008; Rayner, 1986 for examples of this). That is, in the experiments in this thesis, two (or three) sets of sentences could have been developed, one for adults, and one for children (or one for adults, one for older children, and one for younger children). Both adults and children would read both sets of sentences and thus the differences in eye movement behaviour which were due to processing difficulty per se could be teased apart from differences observed which were due specifically to the linguistic manipulation employed. However, this would mean children as young as seven years old would be obliged to read sentences written for adult readers and this raises an ethical, as well as an experimental, difficulty. Clearly these young children would find the sentences very difficult, or even impossible, to read. As a result, they would not understand the text and so their eye-movement behaviour could be extremely difficult to interpret, and in addition, they might very soon become demoralised and unhappy with the testing situation.

It is argued then, that although not ideal, the method employed in the experiments reported in this thesis is the best. While there is no methodological or ethical problem with asking adults to read sentences which are age-appropriate for young children, and indeed, as in Experiment 3 (Chapter 4), asking adults to read a separate set of sentences which are age-appropriate, to ask young children to read sentences which are too difficult for them is not something to be recommended. It is important to be alert to this issue when constructing experimental stimuli, and to try to employ manipulations which will generate strong effects in adult readers as well as children, despite words being short and acquired early in development, and sentences being syntactically and semantically simple (Häikiö et al., 2008). Future research using eye movements during reading with developmental populations should be careful to implement these requirements.

8.7 Final conclusions

Overall, there are four main conclusions to be drawn from the empirical data reported in this thesis. First, children and adults are very similar in how they use visual information to make early oculomotor decisions regarding saccade targeting. Second, word frequency is critical in the organisation and function of the developing as well as the mature mental lexicon. Third, children are delayed in their detection of a syntactic misanalysis as compared with adults. Fourth, children take longer to form a coherent representation of the meaning of a sentence as compared to adults. These four conclusions strongly suggest that children have the same psychological mechanisms in place to process written language online as adults, but that these mechanisms are slightly delayed in the time course in which they operate at a post-lexical level. This series of experiments has demonstrated, for the first time, that eye movements during reading can be used successfully to investigate children's online written language processing in relation to a variety of linguistic phenomena, and it is hoped that the theoretical and methodological implications drawn from this work are helpful to future research in this field.

Bibliography

Adams, M. J. (1990). *Beginning to read*: MIT press.

Aghababian, V., & Nazir, T. (2000). Developing normal reading skills: Aspects of visual processes underlying word recognition. *Journal of Experimental Child Psychology, 76*, 123-150.

Altmann, G. T. M. (1998). Ambiguity in sentence processing. *Trends in Cognitive Sciences, 2*(4), 146-152.

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238.

Altmann, G. T. M., van Nice, K. Y., Garnham, A., & Henstra, J. (1998). Late closure in context. *Journal of Memory and Language, 38*, 459-484.

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

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

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1-2), 285-311.

Arnold, J. E. (2008). THE BACON not the bacon: How children and adults understand accented and unaccented noun phrases. *Cognition, 108*(1), 69-99.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes, 22*(4), 527 - 565.

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM): University of Pennsylvania, Philadelphia: Linguistic Data Consortium.

Backscheider, A. G., & Gelman, S. A. (1995). Children's understanding of homonyms. *Journal of Child Language, 22*, 107-127.

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

Balota, D. A. (1994). Visual word recognition. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 303-358). London: Academic Press.

Balota, D. A., Paul, S., & Spieler, D. H. (1999). Attentional control of lexical processing pathways during word recognition and reading. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 15-57). London: Psychology Press.

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

Baron, J., & Treiman, R. (1980). Use of orthography in reading and learning to read. In J. F. Kavanagh & R.L.Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 171-189). Baltimore, MD: University Park Press.

Beauvillain, C., Dukic, T., & Vergilino, D. (1999). The planning of successive saccades in letter strings. In W. Becker & H. Deubel & T. Mergner (Eds.), *Current oculomotor research: physiological and psychological aspects* (pp. 333-340). New York: Kluwer Academic Plenum Publishers.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*: Wiley.

Beveridge, M., & Marsh, L. (1991). The influence of linguistic context on young children's understanding of homophonic words. *Journal of Child Language, 18*(2), 459-467.

Bijeljac-Babic, R., Millogo, V., Farioli, F., & Grainger, J. (2004). A Developmental investigation of word length effects in reading using a new on-line word identification paradigm. *Reading and Writing, 17*(4), 411-431.

Binder, K. S., & Morris, R. (1995). Eye movements and lexical ambiguity resolution: effects of prior encounter and discourse topic. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(5), 1186-1196.

Binder, K. S., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review, 5*(2), 271-276.

Blodgett, A., & Boland, J. E. (2004). Differences in the Timing of Implausibility Detection for Recipient and Instrument Prepositional Phrases. *Journal of Psycholinguistic Research 33*(1), 1-24.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular co-ordination of eye movements during reading in children and adults. *Vision Research, 46*(22), 3898-3908.

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

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read - a causal connection. *Nature, 301*, 419 - 421

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' Eye Movements Distinguish Anomalies of Form and Content. *Journal of Psycholinguistic Research, 31*(1), 25-44.

Brock, J., Norbury, C., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition, 108*(3), 896-904.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago: University of Chicago Press.

Campbell, R. N., & Bowe, T. (1982). Text and context in early language comprehension. In M. C. Donaldson & R. Grieve & C. Pratt (Eds.), *Early child development and education*. Oxford: Blackwell.

Carpenter, P. A., & Just, M. A. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning & Verbal Behavior, 20*, 137-160.

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

Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Challenges and controversies about word recognition and reading*. London: Psychology Press.

Chall, J. S. (1967). *Learning to read: The great debate*. New York: McGraw-Hill.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Clay, M. M. (1970). An Increasing Effect of Disorientation on the Discrimination of Print: A Developmental Study *Journal of Experimental Child Psychology, 9*(3), 297-306.

Clifton, C., Kennison, S. M., & Albrecht, J. E. (1997). Reading the words Her, His, Him: Implications for parsing principles based on frequency and on structure. *Journal of Memory and Language, 36*(2), 276-292.

Clifton, C., Speer, S., & Abney, S. P. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language, 30*(2), 251-271.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel & M. H. Fischer & W. S. Murray & R. L. Hill (Eds.), *Eye Movements: A Window on Mind and Brain* (pp. 341-371). Amsterdam: Elsevier.

Clifton, C., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language, 49*(3), 317-334.

Cohen, M. E., & Ross, L. E. (1978). Latency and Accuracy Characteristics of Saccades and Corrective Saccades in Children and Adults. *Journal of Experimental Child Psychology, 26*(3), 517-527.

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, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555). London: Academic Press.

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.

Conklin, K., Koenig, J.-P., & Mauner, G. (2004). The role of specificity in the lexical encoding of participants. *Brain and Language, 90*(1-3), 221-230.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. . *Cognitive Psychology, 6*(1), 84-107.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty & L. Karttunen & A. M. Zwicky (Eds.), *Natural language parsing* (pp. 320-354). Cambridge: Cambridge University Press.

Crocker, M. (1999). Mechanisms for Sentence Processing. In S. Garrod & M. J. Pickering (Eds.), *Language Processing*. London: Psychology Press,.

Daneman, M., & Reingold, E. (1993). What eye fixations tell us about phonological recoding during reading. *Canadian Journal of Experimental Psychology, 47*, 153-178.

Daneman, M., & Reingold, E. (1995). What eye fixations tell us about phonological recoding during reading. In J. M. Henderson & M. Singer & F. Ferreira (Eds.), *Reading and language processing* (pp. 25-50). Hove: Erlbaum.

Daneman, M., Reingold, E. M., & Davidson, M. (1995). Time course of phonological activation during reading: Evidence from eye fixations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 618-632.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 17*, 618-632.

Davis, C., & Bowers, J. (2006). Contrasting Five Different Theories of Letter Position Coding: Evidence From Orthographic Similarity Effects. *Journal of Experimental Psychology: Human Perception & Performance, 32*(3), 535-557.

Davis, C. J. (1999). *The Self-Organising Lexical Acquisition and Recognition (SOLAR) model of visual word recognition*. University of New South Wales, Australia (Unpublished doctoral dissertation).

Doherty, M. J. (2004). Children's difficulty in learning homonyms. *Journal of Child Language, 31*, 203-214.

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language, 31*(4), 461-476.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology-Human Perception and Performance, 31*(5), 954-969.

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

Ehri, L. C. (1975). Word consciousness in readers and prereaders. *Journal of Educational Psychology, 67*(2), 204-212.

Ehri, L. C. (1987). Movement into word reading and spelling: How spelling contributes to reading. In J. Mason (Ed.), *Reading and Writing Connections*. Newton, MA: Allyn and Bacon).

Ehri, L. C. (2002). Phases of acquisition in learning to read words and implications for teaching *BJEP Monograph Series II, Number 1 - Learning and Teaching Reading, 1*(1), 7-28.

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135-154). Oxford: Blackwell.

Ehri, L. C., & Wilce, L. S. (1987). Does Learning to Spell Help Beginners Learn to Read Words? *Reading Research Quarterly, 22*(1), 47-65.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior, 20*(6), 641-655.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*(5), 621-636.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777-813.

Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science, 16*, 913 - 920.

Felser, C., Marinis, T., & Clahsen, H. (2003). Children's Processing of Ambiguous Sentences: A Study of Relative Clause Attachment. *Language Acquisition, 11*(3), 127-163.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*(3), 348-368.

Ferretti, T. R., Gagne, C. L., & McRae, K. (2003). Thematic role focusing by participle inflections: Evidence from conceptual combination. *Journal of Experimental Psychology-Learning Memory and Cognition, 29* (1), 118-127.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 182-196.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*(4), 516-547.

Field, A. (2005). *Discovering statistics using SPSS*. London: Sage.

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition, 106*(2), 1038-1046.

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

Folk, J., & Morris, R. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition, 31*(1), 87-99.

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

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition., 22*(3), 696-713.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Unpublished doctoral dissertation. University of Connecticut, Storrs.

Frazier, L. (1987). Sentence processing : a tutorial view. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 559-586). Hove: Erlbaum.

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

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson & J. Marshall & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301-330). Hove: Erlbaum.

Fukushima, J., Hatta, T., & Fukushima, K. (2000). Development of voluntary control of saccadic eye movements: I. Age-related changes in normal children. *Brain and Development, 22*(3), 173-180.

Garrod, S., & Pickering, M. J. (1999). Issues in language processing. In S. Garrod & M. J. Pickering (Eds.), *Language processing* (pp. 1-11). Hove: Psychology Press.

Gentner, D. (1983). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Vol. 2. Language, cognition, and culture.* Hillsdale, NJ: Erlbaum.

Gleitman, L. R. (1981). Maturational determinants of language growth. *Cognition, 10*(1-3), 103-114.

Gleitman, L. R. (1990). The structural sources of verb meaning. *Language Acquisition*(1), 3-55.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42*(1), 73-83.

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-Out Model. *Psychological Review, 103*(3), 518-565.

Groll, S. L., & Ross, L. E. (1982). Saccadic eye movements of children and adults to double-step stimuli. *Developmental Psychology, 18*, 108-123.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2008). Development of letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, (in press).

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

Henderson, J. M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*(2), 201-221.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One Frog, Two Frog, Red Frog, Blue Frog: Factors Affecting Children's Syntactic Choices in Production and Comprehension. *Journal of Psycholinguistic Research, 29*(6), 597-626.

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of nonfixated words influence the processing of fixated words during reading? . *European Journal of Cognitive Psychology, 16*(1-2), 104-127.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word-length and word-frequency. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*(6), 1430-1440.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1612-1627.

Inhoff, A. W. (1984). Two Stages of Word Processing during Eye Fixations in the Reading of Prose. *Journal Of Verbal Learning And Verbal Behavior, 23*, 612-624.

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A. Kennedy & R. Radach & D. Heller & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: effects of word frequency. *Perception and Psychophysics, 40*(6), 431-439.

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

Jared, D., Levy, B. A., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: evidence from proofreading and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3*, 219-264.

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-Letter Effects in Reading: Evidence From Eye Movements and Parafoveal Preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 209-229.

Jordan, T. R., Patching, G. R., & Milner, A. D. (1998). Central fixations are inadequately controlled by instructions alone: Implications for studying cerebral asymmetry. *Quarterly Journal of Experimental Psychology, 51A*(2), 371-391.

Joseph, H. S. S. L., Blythe, H. I., & Liversedge, S. P. (2009). Differential effects of word frequency in adults and children during reading. *Psychonomic Bulletin and Review*, in preparation.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 61*(5), 708-723.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, (under submission).

Judd, C. M., McClelland, G. H., & Culhane, S. E. (1995). Data analysis: Continuing Issues in the Everyday Analysis of Psychological Data. *Annual Review of Psychology, 46*, 433-465.

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

Juhasz, B., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations *Visual Cognition, 13*(7-8), 846-863.

Juhasz, B. J. (2005). Age-of-Acquisition Effects in Word and Picture Identification *Psychological Bulletin, 131*(5), 684-712.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology-Learning Memory and Cognition, 29*(6), 1312-1318.

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

Just, M., & Carpenter, P. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 149-179). Amsterdam: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A*, 429 - 456.

Kennedy, A., Murray, W. S., & Boissiere, C. (2004). Parafoveal pragmatics revisited. *European Journal of Cognitive Psychology, 16*, 128-153.

Kennedy, A., & Pynte, J. (2008). The consequences of violations to reading order: An eye movement analysis. *Vision Research, 48*(21), 2309-2320.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition, 2*, 15-47.

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J. Hyönä & R. Raddach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 391-411). Amsterdam: Elsevier.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 262 - 284.

Koenig, J.-P., Mauner, G., & Bienvenue, B. (2003). Arguments for adjuncts. *Cognition, 89*(2), 67-103.

Koenig, J. P., Mauner, G., Bienvenue, B., & Conklin, K. (2008). What with? The Anatomy of a (Proto)-Role. *Journal of Semantics 25*(2), 175-220.

Kučera, H., & Francis, W. H. (1967). Computational analysis of present-day American English: Providence, RI: Brown University Press.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

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

Lima, S. D., & Inhoff, A. W. (1985). Lexical access during eye fixations in reading: Effects of word initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 272-285.

Liversedge, S. P. (2003). Eye movements and thematic processing. In J. Hyönä & R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement* (pp. 273-290). Amsterdam: Elsevier.

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

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*(4), 676-703.

Markman, E. M. (1989). *Categorization and naming in children*. Cambridge, MA: MIT Press.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic vs thematic relations. *Cognitive Psychology, 16*, 1-27.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121-157.

Marsh, G., Friedman, M. P., Welch, V., & Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 3, pp. 199–221). San Diego, CA: Academic Press.

Marslen-Wilson, W., Brown, C., & Tyler, L. (1988). Lexical representations in language comprehension. *Language and Cognitive Processes, 3*(1-21).

Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (2002). Morphology and meaning in the English mental lexicon. In G. T. M. Altmann (Ed.), *Psycholinguistics: Critical concepts in Psychology* (Vol. II, pp. 17-83). London: Routledge.

Marslen-Wilson, W. D. (1975). Sentence Perception as an Interactive Parallel Process. *Science, 189*(4198), 226-228.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of Discourse Context as a Determinant of the Subordinate Bias Effect. *Quarterly Journal of Experimental Psychology, 52A*(4), 813-839.

Mason, J. (1980). When do children begin to read: An exploration of four-year-old children's letter and word reading competencies. *Reading Research Quarterly, 15*, 203-227.

Mauner, G., Tannenhaus, M. K., & Carlson, G. N. (1995). Implicit Arguments in Sentence Processing. *Journal of memory and language, 34*(3), 357-382.

Mazzocco, M. M. M. (1997). Children's interpretations of homonyms: a developmental study. *Journal of Child Language, 24*, 441-467.

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

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye-movement control during reading: 1. The location of initial eye fixations on words. *Vision Research, 28*(10), 1107-1118.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics, 46*(3), 245-253.

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

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*, 365-368.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. . In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). Boston: CRC Press.

McDonald, S. A., & Vergilino-Perez, D. (2006). Estimating the proportion of preplanned refixations in natural reading using Monte-Carlo simulations. *Visual Cognition, 13*(1), 109-118.

McRae, K., Ferretti, T. R., & Amyote, L. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes: Special Issue on Lexical Representations in Sentence Processing, 12*, 137-176.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the Influence of Thematic Fit (and Other Constraints) in On-line Sentence Comprehension. *Journal of Memory and Language, 38*(3), 283-312.

Meroni, L., & Crain, S. (2003). *On Not Being Led Down the Kindergarten Path.* Paper presented at the Proceedings of the 27th Annual Boston University Conference on Language Development.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition,* 7(4), 323-331.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of Acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A*(3), 528-559.

Murray, W., & Rowan, M. (1998). Early, mandatory, pragmatic processing. *Journal of Psycholinguistic Research, 27*, 1-22.

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

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665-681.

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition, 58*, 221-251.

Nation, K. (in press). Decoding, orthographic learning and the development of visual word recognition. In M. Spivey & K. McRae & M. Joanisse (Eds.), *Cambridge Handbook of Psycholinguistics.*

Nation, K., & Cocksey, J. (2008). Beginning readers activate semantics from orthography. (Unpublished manuscript).

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*, 314-329.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of visual word recognition. *Journal of Research in Reading, 27*, 342-356.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly Detection: Eye Movement Patterns. *Journal of Psycholinguistic Research 27*(5), 515-539.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the effect of inverted optimal viewing position. *Vision Research, 45*(17), 2201-2217.

O'Regan, J. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333-354). New York: Springer.

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

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363-383). Hillsdale, NJ: Erlbaum.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*(4), 443-468.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.

Paap, K., Newsome, S., McDonald, J., & Schvaneveldt, R. (1982). An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89*(5), 573-594.

Pacht, J. M., & Rayner, K. (1993). The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research, 22*(2), 251-271.

Paterson, K. B., Liversedge, S. P., & Davis, C. J. (2008). Inhibitory neighbour priming effects in eye movements during reading. *Psychonomic Bulletin and Review.*

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition, 31*(6), 829-841.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 767-779.

Peters, A. M., & Zaidel, E. (1980). The acquisition of homonymy. *Cognition, 8*, 187-207.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths : An eye-tracking study. *Journal of Experimental Psychology. Learning, Memory, and Cognition 24* (4), 940-961.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: MIT Press.

Pinker, S. (1994). How could a child use verb syntax to learn about verb semantics? *Lingua, 92*, 377-410.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The Role of Morphological Constituents in Reading Finnish Compound Words. *Journal of Experimental Psychology: Human Perception & Performance, 26*(2), 820-833.

Pollatsek, A., Lesch, M., Morris, R., & Rayner, K. (1992). Phonological Codes Are Used in Integrating Information Across Saccades in Word Identification and Reading. *Journal of Experimental Psychology: Human Perception & Performance 18*(1), 148-162.

Pollatsek, A., Reichle, E., & Rayner, K. (2003). Modeling eye movements in reading: Extensions of the EZ Reader model. In J. Hyona & R. Radach & H. Deubel (Eds.), *The mind's eyes: Cognitive and applied aspects of eye movement research*. Amsterdam: Elsevier.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1-56.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1), 178-202.

Rayner, K. (1976). Developmental changes in word recognition strategies. *Journal of Educational Psychology, 68*(3), 323-329.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*(4), 443-448.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*, 21-30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211-236.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*, 468-469.

Rayner, K., Carlson, M., & Frazier, L. (1983). The Interaction of Syntax and Semantics During Sentence Processing - Eye-Movements in the Analysis of Semantically Biased Sentences. *Journal of Verbal Learning and Verbal Behavior, 22*(3), 358-374.

Rayner, K., Cook, A. E., Juhasz, B. J., & Frazier, L. (2006). Immediate disambiguation of lexically ambiguous words during reading: Evidence from eye movements *British Journal of Psychology, 97*(4), 467-482.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*, 1129-1144.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 779-790.

Rayner, K., & Hagelberg, E. M. (1975). Word recognition cues for beginning and skilled readers. *Journal of Experimental Child Psychology, 20*, 444-455.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 167-179.

Rayner, K., & Kaiser, J. S. (1975). Reading Mutilated Text. *Journal of Educational Psychology, 67*(2), 301-306.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 53*(4), 1061-1080.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14*, 385-388.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16*(8), 829-837.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12*, 206-226.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: evidence from eye fixations. *Journal of Memory and Language 33*(4), 527-544.

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: evidence for direct control. *Quarterly Journal of Experimental Psychology A, 33*(4), 351-373.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Hillsdale: Erlbaum.

Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 24*(2), 476-497.

Rayner, K., Pollatsek, A., Liversedge, S. P., & Reichle, E. D. (2009). Eye movements and non-canonical reading: Comments on Kennedy and Pynte. *Vision Research, (in press).*

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3*(2), 245-248.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology-Human Perception and Performance, 22*(5), 1188-1200.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*(6), 1290-1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*(4), 504-509.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception and Psychophysics, 27*, 537-544.

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled letetrs: there is a cost. *Psychological Science, 17*(3), 192-193.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä & R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 213-234). Amsterdam: Elsevier.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology: General, 81*(2), 275-280.

Reichle, E. D., & Laurent, P. A. (2006). Using Reinforcement Learning to Understand the Emergence of "Intelligent" Eye-Movement Behavior During Reading. *Psychological Review, 113*(2), 390-408.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Models of eye movement control in reading: Why the serial allocation of attention is plausible and attention gradients are not. *Trends in Cognitive Sciences, (in press).*

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125-157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*(1), 4-22.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2004). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445-476.

Rumelhart, D., & McClelland, J. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*(1), 60-94.

Rumelhart, D. E., & McClelland, J. L. (1986). On Learning the Past Tenses of English Verbs. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition* (Vol. 2, pp. 216 - 271). Cambridge, MA: MIT Press.

Rust, J., Golombok, S., & Trickey, G. (1993). Wechsler Objective Reading Dimensions Test. Sidcup, UK: Psychological Corporation.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science, 274*(5294), 1926-1928.

Salman, M. S., Sharpe, J. A., Eizenman, M., Lillakas, L., Westall, C., To, T., Dennis, M., & Steinbach, M. J. (2006). Saccades in children. *Vision Research, 46*(8-9), 1432-1439.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory and Cognition, 26*, 1270-1281.

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* (pp. 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.

Sereno, S., O'Donnell, P., & Rayner, K. (2006). Eye Movements and Lexical Ambiguity Resolution: Investigating the Subordinate-Bias Effect. *Journal of Experimental Psychology: Human Perception & Performance, 32*(2), 335-350.

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science, 3*(5), 296-300.

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition, 55*(2), 151-218.

Simpson, G. B., & Foster, M. R. (1986). Lexical ambiguity and children's word recognition. *Developmental Psychology, 22*(2), 147-154.

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge: Cambridge University Press.

Snedeker, J., Thorpe, K., & Trueswell, J. C. (2001). On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. In J. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 964–969). Hillsdale: Erlbaum.

Snowling, M., & Frith, U. (1981). The role of sound, shape and orthographic cues in early reading. *British Journal of Psychology, 72*, 83-87.

Snowling, M. J. (2000). *Dyslexia*. Oxford: Blackwell.

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55*(3), 227-267.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic Performance Profile of Children with Reading Disabilities: A Regression-Based Tests of the Phonological-Core Variable-Difference Model. *Journal of Educational Psychology, 86*(1), 24-53.

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: the role of phonotactic probability and word frequency. *Journal of Child Language, 32*(4), 827-853.

Stuart, M., Masterson, J., Dixon, M., & Quinlan, P. (2002). Children's Printed Word Database: University of Essex).

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-660.

Tanenhaus, M. K., Carlson, G., & Trueswell, J. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes: Special Issue on Parsing and Interpretation, 4*, 211-234.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language, 27*(6), 597-632.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2*, 187-202.

Taylor, S. E., Frackenpohl, H., & Pettee, J. L. (1960). Grade level norms for the components of fundamental reading skill, *EDL Research and Information Bulletin* (Vol. 3). Huntington, NY: Educational Development Laboratories.

Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language, 48*, 740-759.

Traxler, M., Morris, R., & Seely, R. (2002). Processing subject and object-relative clauses: Evidence from eye-movements. *Journal of Memory and Language, 33*, 127-154.

Traxler, M. J. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: evidence from self-paced reading. *The Quarterly Journal of Experimental Psychology: Section A, 55A*(1), 75-96.

Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: Eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology, 50A*, 481-497.

Treiman, R., Cohen, J., Mulqueeny, K., Kessler, B., & Schechtman, S. (2007). Young Children's Knowledge About Printed Names. *Child Development 78*(5), 1458-1471.

Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*(89-134).

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic Influences On Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language, 33*(3), 285-318.

Tyler, L. K., & Marslen-Wilson, W. D. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior, 16*(6), 683-692.

Vergilino-Perez, D., Collins, T., & Dore-Mazars, K. (2004). Decision and metrics of refixations in reading isolated words. *Vision Research, 44*(17), 2009-2017.

Vergilino, D., & Beauvillain, C. (2000). The planning of refixation saccades in reading. *Vision Research, 40*(25), 3527-3538.

Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology, 16*(1-2), 154-177.

Vitu, F., Kapoula, Z., Lancelin, D., & Lavigne, F. (2004). Eye movements in reading isolated words: evidence for strong biases towards the center of the screen. *Vision Research, 44*(3), 321-338.

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy & R. Radach & D. Heller & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301-326). Oxford: Elsevier.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research, 41*(25-26), 3513-3533.

Vitu, F., & O'Regan, J. K. (1995). A challenge to current theories of eye movements in reading. In J. M. Findlay & R. Walker & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes, and applications* (pp. 381-393). Amsterdam: North Holland.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless Reading - Eye-Movement Characteristics Are Similar in Scanning Letter Strings and Reading Texts. *Perception & Psychophysics, 57*(3), 352-364.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal Landing Position in Reading Isolated Words and Continuous Text. *Perception & Psychophysics, 47*(6), 583-600.

Wagner, R., & Torgesen, J. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192-212.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading *Psychonomic Bulletin & Review, 14*(4), 770-775.

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(4), 1001-1010.

Weinberg, A. (1993). Parameters in the theory of sentence processing: Minimal Commitment theory goes east *Journal Journal of Psycholinguistic Research 22*(3), 339-364.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7(1), 49-63.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a New Synthesis. *Annual Review of Psychology, 50*(1), 509-535.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 205-223.

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception & Performance, (in press).*

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8*, 221-243.

Williams, C. C., Perea, M., Pollatsek, A., & Rayner, K. (2006). Previewing the neighborhood: The role of orthographic neighbors as parafoveal previews in reading. *Journal of Experimental Psychology-Human Perception and Performance, 32*(4), 1072-1082.

Williams, R. S., & Morris, R. K. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology, 16*(1-2), 312-339.

Yang, S., & McConkie, G. (2004). Saccade generation during reading: Are words necessary? *European Journal of Cognitive Psychology, 16*(1-2), 226-261.

Yang, S. N., & McConkie, G. W. (2001). Eye movements during reading: a theory of saccade initiation times. *Vision Research, 41*(25-26), 3567-3585.

Appendices

10.1 Appendix 1: Materials for Experiment 1

10.1.1 Length manipulation

Target words are underlined. Sentence (a) contains the long word and sentence (b) contains the short word. All long words were eight letters, and all short words were four letters.

1a. They were all playing in the old dirty <u>carriage</u> yesterday.

1b. They were all playing in the old dirty <u>barn</u> yesterday.

2a. The noisy <u>laughter</u> could be heard across the entire garden.

2b. The noisy <u>band</u> could be heard across the entire garden.

3a. The lovely <u>fountain</u> was in the garden right outside my door.

3b. The lovely <u>arch</u> was in the garden right outside my door.

4a. Her brother was really mean and put some <u>medicine</u> in my tea.

4b. Her brother was really mean and put some <u>salt</u> in my tea.

5a. They went to London to see the famous <u>painting</u> this morning.

5b. They went to London to see the famous <u>park</u> this morning.

6a. The young <u>champion</u> was really pleased with all his success.

6b. The young <u>lord</u> was really pleased with all his success.

7a. The great <u>distance</u> involved made the task really dangerous.

7b. The great <u>risk</u> involved made the task really dangerous.

8a. The local <u>hospital</u> was always busy and full of nice people.

8b. The local <u>shop</u> was always busy and full of nice people.

9a. The clever <u>prisoner</u> owned the book that was for sale today.

9b. The clever <u>hero</u> owned the book that was for sale today.

10a. Your fancy <u>knitting</u> was the best that we saw by a long way.

10b. Your fancy <u>dive</u> was the best that we saw by a long way.

11a. There is a lady standing on the wooden <u>platform</u> and talking.

11b. There is a lady standing on the wooden <u>deck</u> and talking.

12a. He spent a long time talking with your <u>customer</u> in the shop.

12b. He spent a long time talking with your <u>aunt</u> in the shop.

13a. The little <u>creature</u> seemed tame and came quite close to us.

13b. The little <u>bird</u> seemed tame and came quite close to us.

14a. I ignored the argument to avoid any more <u>distress</u> and upset.

14b. I ignored the argument to avoid any more <u>harm</u> and upset.

15a. I saw the baby <u>elephant</u> playing with its mum and dad today.

15b. I saw the baby <u>lamb</u> playing with its mum and dad today.

16a. I talked to your sister in the school <u>corridor</u> this morning.

16b. I talked to your sister in the school <u>pool</u> this morning.

17a. I spilt some tea by mistake on your <u>magazine</u> this morning.

17b. I spilt some tea by mistake on your <u>coat</u> this morning.

18a. I realised that the man had a secret <u>ambition</u> to beat you.

18b. I realised that the man had a secret <u>plot</u> to beat you.

19a. The dirty <u>stranger</u> was just outside the shop doorway today.

19b. The dirty <u>soil</u> was just outside the shop doorway today.

20a. The nasty <u>shepherd</u> got us into lots of trouble with mother.

20b. The nasty <u>liar</u> got us into lots of trouble with mother.

21a. As they were being so careful, their <u>accident</u> was a shock.

21b. As they were being so careful, their <u>bill</u> was a shock.

22a. For our dinner we decided to order some <u>macaroni</u> with peas.

22b. For our dinner we decided to order some <u>stew</u> with peas.

23a. I was looking for all of you in the huge <u>building</u> yesterday.

23b. I was looking for all of you in the huge <u>hall</u> yesterday.

24a. The pretty <u>material</u> had lots of purple flowers all over it.

24b. The pretty <u>hill</u> had lots of purple flowers all over it.

25a. The very rare <u>ornament</u> was always hidden in a secret place.

25b. The very rare <u>ruby</u> was always hidden in a secret place.

26a. There were lots of people near the busy <u>entrance</u> yesterday.

26b. There were lots of people near the busy <u>lane</u> yesterday.

27a. The massive <u>mountain</u> was so impressive that we stared at it.

27b. The massive ship was so impressive that we stared at it.

28a. He got upset and pointed the broken umbrella at me crossly.

28b. He got upset and pointed the broken fork at me crossly.

10.1.2 Frequency manipulation

Target words are underlined. Sentence (a) always contains the high frequency word, and sentence (b) always contains the low frequency word. Frequencies in parentheses (counts per million)

1a. The bitter coffee that you gave me tasted really unpleasant. (92)

1b. The bitter cherry that you gave me tasted really unpleasant. (7)

2a. The sudden danger made us all really scared and we ran away. (94)

2b. The sudden scream made us all really scared and we ran away. (11)

3a. We couldn't stop looking at the lovely garden in the sun. 140)

3b. We couldn't stop looking at the lovely bubble in the sun. (9)

4a. My mum bought me a jumper with a yellow button on the front. (26)

4b. My mum bought me a jumper with a yellow stripe on the front. (7)

5a. They were all a bit scared of the nasty fellow at the park. (44)

5b. They were all a bit scared of the nasty keeper at the park. (9)

6a. The lovely lawyer has really beautiful long, curly red hair. (51)

6b. The lovely maiden has really beautiful long, curly red hair. (6)

7a. The sudden threat was a shock and I forgot what I was doing. (77)

7b. The sudden bother was a shock and I forgot what I was doing. (3)

8a. That strong cheese changed the taste of the pizza entirely. (31)

8b. That strong pepper changed the taste of the pizza entirely. (9)

9a. I think the colour of her hair is quite silver without dye. (51)

9b. I think the colour of her hair is quite normal without dye. (2)

10a. The nice kind worker went swimming with his children today. (240)

10b. The nice kind banker went swimming with his children today. (14)

11a. The gloomy leader walked slowly along the street by himself. (143)

11b. The gloomy beggar walked slowly along the street by himself. (4)

12a. The old broken window nearly hurt someone yesterday morning. (200)

12b. The old broken hammer nearly hurt someone yesterday morning. (11

13a. Yesterday morning I found an old broken record in our shed. (93)

13b. Yesterday morning I found an old broken shield in our shed. (8)

14a. The noisy cattle got plenty of attention from the big crowd. (32)

14b. The noisy pigeon got plenty of attention from the big crowd. (11)

15a. The baby played all day with the pretty marble you gave her. (25)

15b. The baby played all day with the pretty rattle you gave her. (3)

16a. The dusty mirror was by the door where you said it would be. (49)

16b. The dusty armour was by the door where you said it would be. (10)

17a. She spends lot of time talking to the nice people next door. (1480)

17b. She spends lot of time talking to the nice tailor next door. (3)

18a. She laid down and put her head on the soft pillow to sleep.(19)

18b. She laid down and put her head on the soft fleece to sleep. (1)

19a. During the night we heard a really loud shriek in the room. (24)

19b. During the night we heard a really loud insect in the room. (5)

20a. Her lovely speech was ruined by the unexpected rain shower. (93)

20b. Her lovely outfit was ruined by the unexpected rain shower. (13)

21a. The fluffy animal was very cute but it made me sneeze a lot. (260)

21b. The fluffy kitten was very cute but it made me sneeze a lot. 7)

22a. We were all working hard on the special design from France. (81)

22b. We were all working hard on the special puzzle from France. (9)

23a. The special lesson helped him to ride really well yesterday. (57)

23b. The special saddle helped him to ride really well yesterday. 10)

24a. Your torn letter was used to light a fire when we were cold. (206)

24b. Your torn napkin was used to light a fire when we were cold. (7)

25a. I tripped and had an accident in the dark forest last night. (95)

25b. I tripped and had an accident in the dark cavern last night. (3)

26a. He really wanted to wait and see the grand result yesterday. (221)

26b. He really wanted to wait and see the grand finish yesterday. (3)

27a. The mean man had left his dog in the cold church all night. (183)

27b. The mean man had left his dog in the cold kennel all night. (1)

28a. He shouted at the maid and told her to answer the door now. (133)

28b. He shouted at the maid and told her to polish the door now. (7)

10.2 Appendix 2: Materials for Experiments 2 and 3

10.2.1 Experiment 2: Adult frequency

Target words are underlined. Sentence (a) always contains the high frequency word, and sentence (b) always contains the low frequency word. Frequencies in parentheses (counts per million)

1a. They couldn't stop staring at the beautiful <u>animal</u> lying in the sun. (260)

1b. They couldn't stop staring at the beautiful <u>iguana</u> lying in the sun. (1)

2a. She looked out the window and saw the <u>doctor</u> running frantically along the street. (184)

2b. She looked out the window and saw the <u>poodle</u> running frantically along the street. (1)

3a. In her opinion, the <u>people</u> seemed unfriendly and a little frightening. (1480)

3b. In her opinion, the <u>vandal</u> seemed unfriendly and a little frightening. (1)

4a. Although he had a gruelling job, the <u>worker</u> always made time for his family. (204)

4b. Although he had a gruelling job, the <u>joiner</u> always made time for his family. (1)

5a. The girl fancied a <u>change</u> so she left the party and went to her room. (211)

5b. The girl fancied a <u>siesta</u> so she left the party and went to her room. (1)

6a. The breathtaking old <u>church</u> was set in the most beautiful grounds. (183)

6b. The breathtaking old <u>priory</u> was set in the most beautiful grounds. (1)

7a. The biologist examined the <u>ground</u> carefully and decided to carry out some further tests. (177)

7b. The biologist examined the <u>seabed</u> carefully and decided to carry out some further tests. (1)

8a. His late, beloved <u>father</u> was a wonderful man who could never do enough for his family. (474)

8b. His late, beloved <u>godson</u> was a wonderful man who could never do enough for his family. (1)

9a. It was a serious <u>matter</u> which caused both parties a lot of grief. (279)

9b. It was a serious <u>fracas</u> which caused both parties a lot of grief. (1)

10a. Sadly, the <u>window</u> was broken and would be very expensive to repair. (200)

10b. Sadly, the <u>dynamo</u> was broken and would be very expensive to repair. (1)

11a. On receiving the <u>letter</u>, the solicitor stood up and shook his client's hand firmly. (206)

11b. On receiving the <u>refund</u>, the solicitor stood up and shook his client's hand firmly. (1)

12a. The little girl only had one <u>friend</u> but she didn't seem to mind. (356)

12b. The little girl only had one <u>mitten</u> but she didn't seem to mind. (1)

13a. As the wind grew stronger, the <u>family</u> shivered in the cold. (420)

13b. As the wind grew stronger, the <u>puffin</u> shivered in the cold. (1)

14a. When they finally arrived, the <u>school</u> was empty. (513)

14b. When they finally arrived, the <u>kennel</u> was empty. (1)

15a. Admittedly she was a <u>mother</u>, but she was also an extremely beautiful young woman. (474)

15b. Admittedly she was a <u>tomboy</u>, but she was also an extremely beautiful young woman. (1)

16a. I heard the <u>others</u> singing in the garden and it made me smile. (312)

16b. I heard the <u>cuckoo</u> singing in the garden and it made me smile. (1)

17a. The pressure of the <u>system</u> was too much for her and she finally gave in. (373)

17b. The pressure of the <u>sprain</u> was too much for her and she finally gave in. (1)

18a. It's a lovely little <u>street</u> and it has real character. (321)

18b. It's a lovely little bistro and it has real character. (1)

19a. He made his way to the office which was behind the main building. (281)

19b. He made his way to the prefab which was behind the main building. (1)

20a. He is a despicable parent and loves nothing more than to cause his children harm. (317)

20b. He is a despicable sadist and loves nothing more than to cause his children harm. (1)

21a. She was little but her self-confidence made her appear taller. (408)

21b. She was petite but her self-confidence made her appear taller. (1)

22a. If you leave the door open, any old person can just walk in off the street. (262)

22b. If you leave the door open, any old punter can just walk in off the street. (1)

23a. Success in the business world is largely due to chance according to many. (172)

23b. Success in the business world is largely due to acumen according to many. (1)

24a. He liked her figure but wasn't so keen on her personality. (191)

24b. He liked her libido but wasn't so keen on her personality. (1)

10.2.2 Experiment 3: Child frequency

First number in parentheses is word frequency (counts per million); second number is Age-of-Acquisition (months).

1a. They could see the table in the kitchen in the middle of the floor. (241, 20.2)

1b. They could see the knife in the kitchen in the middle of the floor. (11, 17.4)

2a. She rushed downstairs in a blue dress to see what was happening. (154, 32.2)

2b. She rushed downstairs in a blue towel to see what was happening. (16, 32.1)

3a. The kitchen was empty except for a <u>glass</u> and a few plates. (211, 34.1)

3b. The kitchen was empty except for a <u>stool</u> and a few plates. (16, 34.7)

4a. All the children loved the <u>fairy</u> because she was kind and beautiful. (146, 44.6)

4b. All the children loved the <u>nurse</u> because she was kind and beautiful. (11, 44.4)

5a. My mum has lost her <u>watch</u> and she can't find it anywhere. (300, 33.7)

5b. My mum has lost her <u>glove</u> and she can't find it anywhere. (3, 34.8)

6a. The beautiful <u>queen</u> was admired by everyone in the country. (552, 37.7)

6b. The beautiful <u>piano</u> was admired by everyone in the country. (30, 39.5)

7a. The man told us that the <u>kangaroo</u> had disappeared and no-one knew where it could be.(138, 34.1)

7b. The man told us that the <u>necklace</u> had disappeared and no-one knew where it could be. (8, 38.4)

8a. I'd love to have either a <u>tortoise</u> or a stick insect as a pet. (103, 28.5)

8b. I'd love to have either a <u>ladybird</u> or a stick insect as a pet. (57, 13.6)

9a. They ate all the <u>jelly</u> and asked if there was any more. (146, 26.0)

9b. They ate all the <u>toast</u> and asked if there was any more.(51, 36.7)

10a. He broke the <u>camera</u> although he didn't mean to. (160, 44.1)

10b. He broke the <u>cooker</u> although he didn't mean to. (8, 42.9)

11a. The little boys played with the <u>bottle</u> all afternoon. (151, 30.1)

11b. The little boys played with the <u>jigsaw</u> all afternoon. (11, 29.7)

12a. They looked in the <u>castle</u> but couldn't find what they were looking for. (297, 37.5)

12b. They looked in the <u>drawer</u> but couldn't find what they were looking for. (22, 33.5)

13a. They could see the mountain in the distance with the sun setting behind it. (173, 52.1)

13b. They could see the cockerel in the distance with the sun setting behind it. (11, 57.4)

14a. Emily had a beautiful sledge which she loved more than anything. (143, 61.0)

14b. Emily had a beautiful violin which she loved more than anything. (5, 56.9)

15a. The handsome doctor always had a smile for everyone. (170, 36.4)

15b. The handsome cowboy always had a smile for everyone. (14, 60.7)

16a. Of all the animals, the boy liked the tiger best because he had never seen one before. (157, 35.0)

16b. Of all the animals, the boy liked the camel best because he had never seen one before . (22, 44.0)

17a. They could just make out the rocket flying through the air. (141, 44.4)

17b. They could just make out the bullet flying through the air. (3, 64.0)

18a. They thought that the dragon didn't look very nice. (425, 37.8)

18b. They thought that the grapes didn't look very nice. (22, 41.9)

19a. He only realised that he had forgotten his umbrella when it was too late. (95, 24.2)

19b. He only realised that he had forgotten his scissors when it was too late. (16, 23.3)

20a. All she needed was the orange and the milk to finish making breakfast. (146, 26.8)

20b. All she needed was the kettle and the milk to finish making breakfast. (22, 33.0)

10.3 Appendix 3: Materials for Experiment 4

Experimental sentences divided into six regions (/). Region 3 was the target word, and Region 5 was the disambiguating region. There were four versions of each experimental sentence: condition (a): ambiguous word (dominant meaning); condition (b): unambiguous control for sentence (a); condition (c) ambiguous word (subordinate meaning); and condition (d): unambiguous control for sentence (c). Only sentences in conditions (c) and (d) were analysed and reported in Chapter 5. Note also that sentences 5, 6, 10, 16 and 22 were not analysed as they were rated as predictable. Word frequencies in parentheses (counts per million): first number is frequency indexed by an adult.corpus (CELEX) and the second number is frequency indexed by a child corpus (CPWD).

1a. He put down/ the/ bat/ and picked up/ the ball/ from the ground./ (9, 122)
1b. He put down/ the/ net/ and picked up/ the ball/ from the ground./ (35, 92)
1c. He put down/ the/ bat/ and it/ flapped its little wings/ and ran away./
1d. He put down/ the/ hen/ and it/ flapped its little wings/ and ran away./ (6, 241)

2a. The little boy played with/ the/ straw/ instead of/ drinking his juice/ yesterday./ (22, 41)
2b. The little boy played with/ the/ spoon/ instead of/ drinking his juice/ yesterday./ (11, 38)
2c. The little boy played with/ the/ straw/ inside/ the barn/ yesterday./
2d. The little boy played with/ the/ chick/ inside/ the barn/ yesterday./ (2, 35)

3a. The/ awful/ marks/ in the/ exam made the/ teacher angry./ (29, 41)
3b. The/ awful/ grade/ in the/ exam made the/ teacher angry./ (12, -)
3c. The/ awful/ marks/ on the/ wallpaper wouldn't/ come off at all./
3d. The/ awful/ stain/ on the/ wallpaper wouldn't/ come off at all./ (6, -)

4a. It wasn't a/ big/ match/ but the/ players were/ brilliant./ (55, 81)

4b. It wasn't a/ big/ pitch/ but the/ players were/ brilliant./ (11, 22)

4c. It wasn't a/ big/ match/ but it was enough to/ light all the candles/ we needed./

4d. It wasn't a/ big/ flame/ but it was enough to/ light all the candles/ we needed./ (15, 8)

5a. The woman put/ the/ bulb/ into/ the ground/ carefully./ (6, 11)

5b. The woman put/ the/ seed/ into/ the ground/ carefully./ (28, 70)

5c. The woman put/ the/ bulb/ into/ the lamp/ carefully./

5d. The woman put/ the/ wire/ into/ the lamp/ carefully./ (35, 16)

6a. The children laughed at/ the/ horn/ of the/ car because/ it was so squeaky./ (9, 68)

6b. The children laughed at/ the/ beep/ of the/ car because/ it was so squeaky./ (1, 3)

6c. The children laughed at/ the/ horn/ of the/ rhino because/ it looked really funny./

6d. The children laughed at/ the/ neck/ of the/ rhino because/ it looked really funny./ (72, 87)

7a. The music club needed a/ new/ speaker/ because the old one/ had retired./ /(17, -)

7b. The music club needed a/ new/ drummer/ because the old one/ had retired./ /(3, 5)7c. The music club needed a/ new/ speaker/ because the old one/ had broken./ /

7d. The music club needed a/ new/ trumpet/ because the old one/ had broken./ /(5, 16)

8a. She hated having/ long/ nails/ and always/ cut them/ straightaway./ (14, 8)

8b. She hated having/ long/ hairs/ and always/ cut them/ straightaway./ (9, 3)

8c. She hated having/ long/ nails/ in the/ toolbox because/ they were dangerous./

8d. She hated having/ long/ wires/ in the/ toolbox because/ they were dangerous./ (35, 16)

9a. The woman noticed/ the/ chest/ of/ treasure and/ money was full./ (43, 78)

9b. The woman noticed/ the/ trunk/ of/ treasure and/ money was full./ (46, 20)

9c. The woman noticed/ the/ chest/ of/ the man when/ he took his t-shirt off./

9d. The woman noticed/ the/ tummy/ of/ the man when/ he took his t-shirt off./ (51, 10)

10a. The/ huge/ trunk/ of the/ tree was/ covered in tiny green leaves./ (20, 46)

10b. The/ huge/ branch/ of the/ tree was/ covered in tiny green leaves./ (54, 87)

10c. The/ huge/ trunk/ of the/ elephant looked/ even bigger than his ears./

10d. The/ huge/ knees/ of the/ elephant looked/ even bigger than his ears./ (54, 43)

11a. He looked at/ the/ line/ he had/ drawn and/ couldn't decide if it was straight./ (218, 114)

11b. He looked at/ the/ edge/ he had/ drawn and/ couldn't decide if it was straight./ (76, 87)

11c. He looked at/ the/ line/ of/ washing and/ wondered if it was dry yet./

11d. He looked at/ the/ pile/ of/ washing and/ wondered if it was dry yet./ (25, 41)

12a. When the girl saw/ the/ table/ she/ sat down/ ready for her tea./ (203, 241)

12b. When the girl saw/ the/ chair/ she/ sat down/ ready for her tea./ (107, 208)

12c. When the girl saw/ the/ table/ in the/ maths book/, she knew it would be a hard question./

12d. When the girl saw/ the/ graph/ in the/ maths book/, she knew it would be a hard question./ (164, -)

13a. They had waited a long time for/ the/ ball/ and/ kicked off/ immediately./ (93, 346)

13b. They had waited a long time for/ the/ game/ and/ kicked off/ immediately./ (148, 160)

13c. They had waited a long time for/ the/ ball/ and/ danced/ happily./

13d. They had waited a long time for/ the/ song/ and/ danced/ happily./ (33, 124)

14a. A really/ good/ coach/ has a high-quality/ engine and/ drives fast./ (28, 8)

14b. A really/ good/ truck/ has a high-quality/ engine and/ drives fast./ (25, 59)

14c. A really/ good/ coach/ /teaches students/ to think for themselves./

14d. A really/ good/ tutor/ /teaches students/ to think for themselves./ (17, -)

15a. She loved/ the/ band/ which she/ listened to/ everyday on the radio./ (32, 100)

15b. She loved/ the/ song/ which she/ listened to/ everyday on the radio./ (33, 104)

15c. She loved/ the/ band/ and/ wore it/ to school every single day./

15d. She loved/ the/ ring/ and/ wore it/ to school every single day./ (66, 95)

16a. The boy put/ the/ plug/ into the/ bath before/ he ran the water./ (6, 127)

16b. The boy put/ the/ soap/ into the/ bath before/ he ran the water./ (20, 49)

16c. The boy put/ the/ plug/ into the/ socket to/ make the lamp work./

16d. The boy put/ the/ wire/ into the/ socket to/ make the lamp work./ (35, 16)

17a. There was/ a/ tap/ that was/ broken in/ the kitchen sink./ (22, 81)

17b. There was/ a/ pot/ that was/ broken in/ the kitchen sink./ (23, -)

17c. There was/ a/ tap/ on the/ door which/ gave them all a fright./

17d. There was/ a/ rap/ on the/ door which/ gave them all a fright./ (1, 3)

18a. She saw/ the/ mole/ /digging up/ the garden in the middle of the night./ (4, 57)

18b. She saw/ the/ ants/ /digging up/ the garden in the middle of the night./ (8, 43)

18c. She saw/ the/ mole/ on his/ face as/ soon as he turned round./

18d. She saw/ the/ lump/ on his/ face as/ soon as he turned round./ (13, 19)

19a. He gave her/ a/ ring/ /made of /silver/ to celebrate their anniversary./ (66, 95)

19b. He gave her/ a/ gift/ /made of /silver/ to celebrate their anniversary./ (31, 41)

19c. He gave her/ a/ ring/ on her/ mobile to/ arrange a time to meet./

19d. He gave her/ a/ call/ on her/ mobile to/ arrange a time to meet./ (238, 254)

20a. They found/ the/ note/ difficult to/ read because/ the writing was so bad./ (82, 35)

20b. They found/ the/ list/ difficult to/ read because/ the writing was so bad./ (75, 46)

20c. They found/ the/ note/ difficult to/ sing because/ it was so high./

20d. They found/ the/ tune/ difficult to/ sing because/ it was so high./ (16, 38)

21a. She liked/ the/ shade/ especially as/ blue was/ her favourite colour./ (22, 19)

21b. She liked/ the/ skirt/ especially as/ blue was/ her favourite colour./ (20, 11)

21c. She liked/ the/ shade/ because it/ was cooler/ than outside in the sun./

21d. She liked/ the/ hotel/ because it/ was cooler/ than outside in the sun./ (125, 24)

22a. The girl knew/ every/ letter/ he had/ written off/ by heart./ (121, 54)

22b. The girl knew/ every/ email/ he had/ written off/ by heart./ (-, -)

2c. The girl knew/ every/ letter/ in the/ alphabet off/ by heart./

22d. The girl knew/ every/ sound/ in the/ alphabet off/ by heart./ (165, 419)

23a. At Christmas,/ the/ pipe/ often/ leaks because/ the water turns to ice./ (22, 41)

23b. At Christmas,/ the/ sink/ often/ leaks because/ the water turns to ice./ (26, 46)

23c. At Christmas,/ the/ pipe/ is/ played by/ somebody almost every night./

23d. At Christmas,/ the/ tune/ is/ played by/ somebody almost every night./ (16, 38)

24a. Everyone liked/ the/ bar/ which had/ good beer/ and played good music./ (66, 19)

24b. Everyone liked/ the/ pub/ which had/ good beer/ and played good music./ (21, 8)

24c. Everyone liked/ the/ bar/ made of/ chocolate which/ was very yummy./

24d. Everyone liked/ the/ pie/ made of/ chocolate which/ was very yummy./ (13, 89)

25a. In the middle of/ the/ court/ the two/ tennis players/ shook hands politely./ (128, 5)

25b. In the middle of/ the/ crowd/ the two/ tennis players/ shook hands politely./ (48, 81)

25c. In the middle of/ the/ court/ the/ lawyer asked/ an important question./

25d. In the middle of/ the/ trial/ the/ lawyer asked/ an important question./ (55, -)

26a. A lot/ of/ fans/ /cool you/ down in the summer when it's hot./ (9, 8)

26b. A lot/ of/ ices/ /cool you/ down in the summer when it's hot./ (3, 1)

26c. A lot/ of/ fans/ /support Manchester/ United but I don't./

26d. A lot/ of/ lads/ /support Manchester/ United but I don't./ (3, 9)

27a. She couldn't find/ her/ glasses/ so she/ couldn't see/ a thing in the dark./ (51, 124)

27a. She couldn't find/ her/ candles/ so she/ couldn't see/ a thing in the dark./ (9, 14)27a. She couldn't find/ her/ glasses/ so she didn't have/ any wine/ with her dinner./

27a. She couldn't find/ her/ bottles/ so she didn't have/ any wine/ with her dinner./ (33, 51)

28a. He used/ the/ key/ to/ open the door/ of the haunted house./ (71, 581)

28b. He used/ the/ bar/ to/ open the door/ of the haunted house./ (66, 19)28c. He used/ the/ key/ to/ shoot the baddie/ on his computer game./

28d. He used/ the/ gun/ to/ shoot the baddie/ on his computer game./ (63, 16)

10.4 Appendix 4: Materials for Experiments 5 and 6

10.4.1 Experiment 5

Experimental sentences divided into five regions (/). Region 3 was the target word, and Region 4 was the post-target region. Sentence (a) is always the high-attached condition, and sentence (b) is the low-attached condition. Word frequencies of target words are in parentheses (counts per million): first number refers to frequency indexed by an adult.corpus (CELEX) and the second number is frequency indexed by a child corpus (CPWD).

1a. Emily ate the cake with/ the silver/ spoon/ on her/ birthday./ (11, 38)
1b. Emily ate the cake with/ the silver/ icing/ on her/ birthday./ (2, 14)

2a. Sam bought the land with/ the/ money/ from/ his father./(403, 365)
2b. Sam bought the land with/ the/ river/ from/ his father./ (108, 435)

3a. The man hunted the tiger with/ the sharp/ spear/ in the/ jungle./ (8, 3)
3b. The man hunted the tiger with/ the sharp/ claws/ in the/ jungle./ (7, 57)

4a. The nurse injected the patient with/ the horrible/ needles/ at the/ hospital./ (7, 8)
4b. The nurse injected the patient with/ the horrible/ disease/ at the/ hospital./ (63, 8)

5a. The boy poked the elephant with/ the long/ stick/ from/ outside the cage./ (54, 146)
5b. The boy poked the elephant with/ the long/ trunk/ from/ outside the cage./ (20, 46)

6a. The carpenter fixed the cupboard with/ the new/ drill/ because/ it was broken./ (10, 5)
6b. The carpenter fixed the cupboard with/ the new/ shelf/ because/ it was broken./ (14, 27)

7a. My dad cleaned the room with/ the wooden/ brush/ yesterday./ / (16, 119)

7b. My dad cleaned the room with/ the wooden/ floor/ yesterday./ / (161, 314)

8a. Mum wiped the floor with/ the dirty/ cloth/ this/ morning./ (45, 35)

8b. Mum wiped the floor with/ the dirty/ patch/ this/ morning./ (17, 62)

9a. The bakers cut the pies with/ the sharp/ knives/ at the/ back of the shop./ 9, 14)

9b. The bakers cut the pies with/ the apple/ filling/ at the/ back of the shop./ (21, 5)

10a. The explorer found the campsite with/ the new/ compass/ just/ before nightfall./ (5, -)

10a. The explorer found the campsite with/ the new/ caravan/ just/ before nightfall./ (7, 30)

11a. Jack drank his lemonade with/ the strange/ straw/ from/ his glass./ (22, 41)

11b. Jack drank his lemonade with/ the strange/ taste/ from/ his glass./ (56, 65)

12a. Tom hit the ball with/ the blue/ racket/ straight/ into the net./ (11, 11)

12b. Tom hit the ball with/ the blue/ stripe/ straight/ into the net./ (2, 3)

13a. The waiter served the soup with/ a white/ tray/ in the/ restaurant./ (20, 30)

13b. The waiter served the soup with/ a white/ roll/ in the/ restaurant./ (28, 49)

14a. The plumber fixed the sink with/ the special/ tool/ in less/ than an hour./ (16, 5)

14b. The plumber fixed the sink with/ the broken/ pipe/ in less/ than an hour./ (22, 41)

15a. Katie opened the present with/ the blue/ scissors/ excitedly./ / (4, 16)

15b. Katie opened the present with/ the blue/ wrapping/ excitedly./ / (4, 3)

16a. The gardener planted the tree with/ the huge/ spade/ at the/ end of the garden./ (3, 16)

16b. The gardener planted the tree with/ the huge/ trunk/ at the/ end of the garden./ (20, 46)

10.4.2 Experiment 6

Experimental sentences divided into six regions (/). Region 5 was the target word, and eye movement measures from Regions 2 and 4 were also taken. Sentence (a) is always the early closure condition, and sentence (b) is the late closure condition.

1a. I think/ I'll wear/ the new skirt/ I bought/ tomorrow./ It's really nice./

1b. I think/ I'll wear/ the new skirt/ I bought/ yesterday./ It's really nice./

2a. Mum says/ we'll eat/ the spaghetti/ she cooked for dinner/ tomorrow./ I love spaghetti./

2b. Mum says/ we'll eat/ the spaghetti/ she cooked for dinner/ last night./ I love spaghetti./

3a. The students/ are going to hand in/ the homework/ they did/ next week./ They hate doing their homework./

3b. The students/ are going to hand in/ the homework/ they did/ last week./ They hate doing their homework./

4a. Adam/ will marry/ the beautiful girl/ he met/ next May./ Her name is Mia./

4b. Adam/ will marry/ the beautiful girl/ he met/ last May./ Her name is Mia./

5a. /I'll show/ everyone the photos/ I took/ tomorrow./ They're of my holiday in Greece./

5b. /I'll show/ everyone the photos/ I took/ yesterday./ They're of my holiday in Greece./

6a. Ella/ is going to drink/ the orange juice/ she made herself/ yesterday./ She loves juice./

6b. Ella/ is going to drink/ the orange juice/ she made herself/ tomorrow./ She loves juice./

7a. /I'll take back/ the library books/ I borrowed/ next Monday./ I haven't read them yet./

7b. /I'll take back/ the library books/ I borrowed/ last Monday./ I haven't read them yet./

8a. Dad says/ he'll finish/ the job/ he started/ next Saturday./ He always says that!/

8b. Dad says/ he'll finish/ the job/ he started/ last Saturday./ He always says that!/

9a. Suzie/ will watch/ the programme/ she recorded/ last week./ It's about Britney Spears./

9b. Suzie/ will watch/ the programme/ she recorded/ last week./ It's about Britney Spears./

10a. Peter/ is going to meet/ the friend/ he phoned/ tomorrow evening./ They're going to the theatre./

10b. Peter/ is going to meet/ the friend/ he phoned/ yesterday evening./ They're going to the theatre./

11a. /They'll listen/ to the CD/ they got/ later on./ It's by the Arctic Monkeys./

11b. /They'll listen/ to the CD/ they got/ earlier./ It's by the Arctic Monkeys./

12a. /He'll play/ the new game/ his friend gave him/ next weekend./ It's a really easy one./

12b. /He'll play/ the new game/ his friend gave him/ last weekend./ It's a really easy one./

13a. Sarah/ will start/ the new job/ they offered her/ next Thursday./ She's a bit nervous./

13b. Sarah/ will start/ the new job/ they offered her/ last Thursday./ She's a bit nervous./

14a. The children/ will eat/ the meal/ their mother made/ tomorrow evening./ She's a good cook./

14b. The children/ will eat/ the meal/ their mother made/ yesterday evening./ She's a good cook./

15a. The teacher/ will give back/ the homework/ she marked/ in the next class./ I hope I get a good mark./

15b. The teacher/ will give back/ the homework/ she marked/ in the last class./ I hope I get a good mark./

16a. /I will say/ thank you for the present/ my friend bought me/ tomorrow./ It's a pencil case./

16b. /I will say/ thank you for the present/ my friend bought me/ yesterday./ It's a pencil case./

10.5 Appendix 5: materials for Experiment 7

Experimental sentences divided into five regions. For each set of three, the sentence (a) is the anomalous condition, sentence (b) is the implausible condition, and sentence (c) is the plausible (control) condition. Region 3 was the target region. Adult word frequencies for the instrument of the verb, and the infinitival verb are given in parentheses (counts per million).

1a. Beatrice used a towel to dry/ the important/ programme/ on the/ computer./ (15, 91)
1b. Beatrice used a key to open/ the important/ programme/ on the/ computer./ (71, 295)
1c. Beatrice used a password to open/ the important/ programme/ on the/ computer./ (1, 295)

2a. The man used the formula to explain/ the beautiful/ boat/ after the/ trip./(25, 84)
2b. The man used the shoelace to tie up/ the beautiful/ boat/ after the/ trip./ (1, 35)
2c. The man used the rope to tie up/ the beautiful/ boat/ after the/ trip../ (31, 35)

3a. Robert used a radio to play/ the horrible/ mouse/ that was/ very scared./ (84, 274)
3b. Robert used a hook to catch/ the horrible/ mouse/ that was/ very scared../ (31, 71)
3d. Robert used a trap to catch/ the horrible/ mouse/ that was/ very scared./ (21, 71)

4a. Justin used a needle to sew/ the spotted/ Dalmatian/ that he/ was walking./ (9, 4)
4b. Justin used a joystick to control/ the spotted/ Dalmatian/ that he/ was walking./ (-, 218)
4c. Justin used a muzzle to control/ the spotted/ Dalmatian/ that he/ was walking./ (3, 218)

5a. Jenny used a hose to water/ the small/ butterfly/ flying past./ / (3, 158)
5b. Jenny used a mousetrap to catch/ the small/ butterfly/ flying past./ / (-, 71)
5c. Jenny used a net to catch/ the small/ butterfly/ flying past./ / (32, 71)

6a. Sarah used a fork to eat/ the fresh/ water/ extremely/ carefully./ (12, 136)
6b. Sarah used a purse to carry/ the fresh/ water/ extremely/ carefully./ (9, 100)

6c. Sarah used a bucket to carry/ the fresh/ water/ extremely/ carefully./ (13, 100)

7a. Matthew used a brush to sweep/ the bright/ star/ in the/ sky./ (16, 13)
7b. Matthew used a microscope to watch/ the bright/ star/ in the/ sky./ (6, 109)
7c. Matthew used a telescope to watch/ the bright/ star/ in the/ sky./ (6, 109)

8a. Dad used a fork to eat/ the purple/ flowers/ in the/ garden./ (12, 136)
8b. Dad used a sword to protect/ the purple/ flowers/ in the/ garden./ (13, 48)
8c. Dad used a fence to protect/ the purple/ flowers/ in the/ garden./ (22, 48)

9a. Todd used a hammer to nail/ the heavy/ shopping/ from Tesco./ / (13, 14)
9b. Todd used a helicopter to carry/ the heavy/ shopping/ from Tesco./ / (11, 100)
9c. Todd used a trolley to carry/ the heavy/ shopping/ from Tesco./ / (5, 100)

10a. Ben used a car to climb/ the highest/ branch/ of the/ tree./ (276, 37)
10b. Ben used a map to reach/ the highest/ branch/ of the/ tree./ 30, 93)
10c. Ben used a ladder to reach/ the highest/ branch/ of the/ tree./ (13, 93)

11a. The witch used a cloth to polish/ the special/ liquid/ for the/ magic potion./ (45, 7)
11b. The witch used a basket to hold/ the special/ liquid/ for the/ magic potion./ (18, 156)
11c. The witch used a bowl to hold/ the special/ liquid/ for the/ magic potion./ (30, 156)

12a. John used a straw to drink/ the carrots/ for dinner/ last night./ (22, 119)
12b. John used an axe to chop/ the carrots/ for dinner/ last night./ (7, 8)
12c. John used a knife to chop/ the carrots/ for dinner/ last night./ (35, 8)

13a. The man used a feather to tickle/ the thin/ spaghetti/ yesterday/ evening./ (5, 1)
13b. The man used a kettle to boil/ the thin/ spaghetti/ yesterday/ evening./ (11, 21)
13c. The man used a pan to boil/ the thin/ spaghetti/ yesterday/ evening./ (22, 21)

14a. The woman used the expensive video to record/ the birthday/ present/ yesterday./ (7, 78)

14b. The woman used the fluffy towel to wrap/ the birthday/ present/ yesterday./ (15, 9)

14c. The woman used the pretty paper to wrap/ the birthday/ present/ yesterday./ (174, 9)

15a. Phillip used the match to light/ the smelly/ cheese/ from Italy./ / (55, 289)

15b. Phillip used the scissors to cut/ the smelly/ cheese/ from Italy./ / (4, 176)

15c. Phillip used the knife to cut/ the smelly/ cheese/ from Italy./ / (35, 176)

16a. Mum used a CD player to hear/ the dirty/ dishes/ in the/ sink./ (-, 188)

16b. Mum used a hoover to clean/ the dirty/ dishes/ in the/ sink./ (1, 87)

16c. Mum used a sponge to clean/ the dirty/ dishes/ in the/ sink./ (6, 87)

17a. The man used a submarine to attack/ the front/ porch/ for the/ party./ (10, 110)

17b. The man used a toothbrush to clean/ the front/ porch/ for the/ party./ (2, 87)

17c. The man used a mop to clean/ the front/ porch/ for the/ party./ (6, 87)

18a. Mum used the music to calm/ the hot/ beans/ for dinner./ / (133, 40)

18b. Mum used the toothpick to serve/ the hot/ beans/ for dinner./ / (1, 68)

18c. Mum used the spoon to serve/ the hot/ beans/ for dinner./ / (11, 68)

19a. Dad used the fork to eat/ the overgrown/ grass/ in the/ garden./ (12, 136)

19b. Dad used the scissors to cut/ the overgrown/ grass/ in the/ garden./ (4, 176)

19c. Dad used the lawnmower to cut/ the overgrown/ grass/ in the/ garden./ (-, 176)

20a. Mum used a spoon to feed/ the chocolate/ cake/ in the/ oven./ (11, 52)

20b. Mum used a crane to put/ the chocolate/ cake/ in the/ oven./ (2, 684)

20c. Mum used an oven glove to put/ the chocolate/ cake/ in the/ oven./ 18, 684)

21a. Richard used a ghost to scare/ the different/ heights/ of children/ in the class./ (20, 7)

21b. Richard used a stopwatch to measure/ the different/ heights/ of children/ in the class./ (1, 49)

21c. Richard used a ruler to measure/ the different/ heights/ of children/ in the class./ (8, 49)

22a. The woman used a mobile phone to ring/ her valuable/ books/ while she/ was away./ (15, 66)

22b. The woman used a jewellery box to store/ her valuable/ books/ while she/ was away./ (-, 57)

22c. The woman used a cardboard box to store/ her valuable/ books/ while she/ was away./ (78, 57)

23a. The man used a pencil to write/ the expensive/ bottle/ of wine/ in the restaurant./(15, 128)

23b. The man used a tin opener to open/ the expensive/ bottle/ of wine/ in the restaurant./ (-, 295)

23c. The man used a corkscrew to open/ the expensive/ bottle/ of wine/ in the restaurant./ (1, 295)

24a. Liz used her coffee mug to drink/ her friend's/ phone number/ yesterday./ / (6, 119)

24b. Liz used her compass to find/ her friend's/ phone number/ yesterday./ / (5, 516)

24c. Liz used her address book to find/ her friend's/ phone number/ yesterday./ / (273, 516)

25a. The prince used a microphone to sing/ the lovely /princess/ from the/ dragon./ (6, 24)

25b. The prince used a floppy disk to save/ the lovely/ princess/ from the/ dragon./ (10, 68)

25c. The prince used a sword to save/ the lovely/ princess/ from the/ dragon./ (13, 68)

26a. The boy used a stepladder to climb/ the enormous/ fish/ for dinner. / / (1, 37)

26b. The boy used a baseball glove to catch/ the enormous/ fish/ for dinner./ / (5, 71)

26c. The boy used a fishing rod to catch/ the enormous/ fish/ for dinner. / / (28, 71)

27a. The girl used a trumpet to play/ the tiny/ picture/ for her/ mother./ (5, 274)

27b. The girl used a roller to paint/ the tiny/ picture/ for her/ mother./ (4, 40)

27c. The girl used a brush to paint/ the tiny/ picture/ for her/ mother./ (16, 40)

28a. The shop assistant used a spade to dig/ the/ / food/ around the/ supermarket./ (3, 19)

28b. The shop assistant used a wallet to carry/ the/ / food/ around the/ supermarket./ (7, 100)

28c. The shop assistant used a trolley to carry/ the/ / food/ around the/ supermarket./ (5, 100)

29a. The farmer used a saw to cut/ the dirty/ pigsty/ in his/ farm./ (-, 87)

29b. The farmer used a duster to clean/ the dirty/ pigsty/ in his/ farm./ (2, 176)

29c. The farmer used a hose to clean/ the dirty/ pigsty/ in his/ farm./ (3, 176)

30a. The waiter used a seed to grow/ the fresh/ milk/ in the/ teacup./ (28, 92)

30b. The waiter used a bucket to pour/ the fresh/ milk/ in the/ teacup./ (13, 24)

30c. The waiter used a jug to pour/ the fresh/ milk/ in the/ teacup./ (3, 24)

31a. Dad used a blue pen to write/ his shaggy/ beard/ in the/ morning./ (19, 128)

31b. Dad used a blunt knife to shave/ his shaggy/ beard/ in the/ morning./ (35, 6)

31c. Dad used a sharp razor to shave/ his shaggy/ beard/ in the/ morning./ (8, 6)

32a. The student used a piano to play/ the difficult/ word/ he didn't/ understand./ (26, 274)

32b. The student used a roadmap to find/ the difficult/ word/ he didn't/ understand./ (30, 516)

32c. The student used a dictionary to find/ the difficult/ word/ he didn't/ understand./ (6, 516)

33a. Santa Claus used a teaspoon to eat/ his Christmas/ sleigh/ around the/ world./ (3, 136)

33b. Santa Claus used a door handle to pull/ his Christmas/ sleigh/ around the/ world./ (43, 68)

33c. Santa Claus used a reindeer to pull/ his Christmas/ sleigh/ around the/ world./ (4, 68)

34a. Gemma used her frying pan to cook/ her handsome/ boyfriend/ on Valentine's/ Day./ (1, 39)

34b. Gemma used her alarm clock to ring/ her handsome/ boyfriend/ on Valentine's/ Day./ (36, 66)

34c. Gemma used her mobile phone to ring/ her handsome/ boyfriend/ on Valentine's/ Day./ (80, 66)

35a. Dad used a kitchen broom to brush/ his lovely/ family/ on holiday/ to Greece./ (6, 16)

3b. Dad used a wheelbarrow to take/ his lovely/ family/ on holiday/ to Greece./ (1, 768)

35c. Dad used a small plane to take/ his lovely/ family/ on holiday/ to Greece./ (45, 768)

36a. Damien used a stepladder to climb/ the nasty /rat/ in his/ house./ (1, 37)

36b. Damien used a machine gun to kill/ the nasty/ rat/ in his/ house./ (63, 79)

36c. Damien used a mousetrap to kill/ the nasty/ rat/ in his/ house./ (21, 79)