THE UNIVERSITY OF HULL

An examination of poor readers' approaches to recognising printed words.

being a Thesis submitted for the Degree of

Doctor of Philosophy

in the University of Hull

by

Amy Caroline Warhurst

BSc. (Hons) Psychology, The University of Hull

November 2008

ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr Singleton and Dr Lavidor for their advice, and for permission to use their testing stimuli. I would also like to say thank you to Keda Cowling for allowing me to use her book '*Toe-by-Toe*' and to alter it to include a synthetic phonics element. I would also like to thank DB for his patience in enduring the battery of tests conducted, and to all of my participants, particularly those who took part in the intervention programme, for all of their effort and hard-work, and for sticking with it despite sometimes finding it difficult.

More personally, I would like to say thank you to Sarah and Gaston for starting this journey with me, and for sharing all the tears and laughter along the way. To James, thank you for your support and guidance throughout. To Kathrine, thank you for your friendship, your belief in me, your staunch support and your eternal optimism. Thank you also to my Mum, Dad and brother, for your encouragement and support and for helping me to remain calm and committed in the face of stress. To my partner, B, I would like to say thank you for everything; for helping me to always find the funny side of any situation, and for your over-whelming belief in my abilities.

The biggest thank you, however, is to Rhona, without whom this thesis would never have been written. Thank you for your patience, understanding, support, advice and expertise.

"Believe you will succeed at whatever you do, and never forget the value of persistence, discipline, and determination." Edmund O'Neil

Thank you to everyone who helped me to achieve my goals, and who believed in my ability to do so.

ABSTRACT

This thesis aimed to explore the approaches to reading made by poor readers. An examination was made of a compensated adult dyslexic, DB, who did not show the degree of impairment in phonological awareness skills generally reported in studies, yet who read words via a visual route. He did show an atypical pattern of performance on the regularity task, as he showed no advantage in reading regular versus irregular words, but on the other hand he showed only slightly impaired non-word reading. A lexical decision task, split across the hemispheres, revealed that DB had no over-reliance on the right hemisphere (RH) for reading processes, which it has been suggested is characteristic of adult dyslexics, yet on the other hand he did not show the same pattern of results as the non-dyslexic controls.

It was thought probable that due to his age and attendance at university, DB would have become too accustomed to his particular compensatory reading strategies to be able to be taught a more phonological approach within the time constraints of this thesis. Instead, it was deemed more appropriate to examine other, younger, poor readers to see whether they also read visually, and whether they could learn a more phonological approach to improve their reading accuracy.

In a second study, therefore, a group of high-school pupils with reading difficulties took part in a reading intervention programme using a synthetic phonic approach. The inclusion of synthetic phonics was due to recent research showing that this form of phonics is very beneficial for beginning readers. The study was designed to see if it was also effective for children making slow progress in learning to read. In an intervention lasting on average just over 16 hours, mean reading age improved by 24 months in a 14

month period, compared to an improvement of only 5 months over the same period for the control group. The ‘visual’ readers in the experimental group also became much more phonological in their approach to reading. One child, XP, was of particular interest, displaying a strong tendency to read words visually at pre-test, showing patterns of reading similar to DB, although neither XP nor DB showed signs of severe phonological awareness deficits. At post-test, however, XP showed a robust regularity effect, and much improved non-word reading accuracy, indicative of the adoption of a more phonological approach to reading. Synthetic phonics was therefore shown to be beneficial as a remediation for older readers who have already developed difficulties.

TABLE OF CONTENTS

CHAPTER ONE

Reading Models, Reading Difficulties and the Teaching of Reading

Reading is an essential skill, not least because without it many avenues within society become inaccessible. It is necessary to be able to read in order to gain information about the world in which we live, for example road signs, maps, shop names, and newspaper headlines. Reading development, therefore, has been the subject of much research and debate over the last few decades. It is widely accepted that the majority of children learn to read with the help and guidance of parental and school tuition, and emerge as competent and fluent readers by the time they leave secondary school. Unfortunately, however, some children do not learn to read at the normal rate. This can be for a multitude of different reasons: lack of exposure to print, lack of school attendance, unsupportive parents, general lack of belief in the importance of an education, or due to specific reading difficulties, such as dyslexia. Regardless of why such children do not learn to read at the same rate as their peers, there needs to be additional support available to these children, and extra tuition in the skills for competency and fluency in reading. Indeed, all children need to be equipped with the tools necessary for decoding and encoding the printed word.

Words are not simple stimuli, but are complex, not only in terms of their semantics, but also in terms of their lexical properties, including, amongst others, word frequency, orthographic neighbourhood, and spelling regularity. Word frequency refers to how often a word is used. The source of information about words and their frequencies is text; to be precise, only those words that occur in a certain text, or a sampling group of texts, enter a frequency dictionary. According to Karlsen and Snodgrass (2004), most American frequency dictionaries are comprised of tabulations that are based on a text quantity assumed to be reasonably representative of printed American English. It is also thought that the tabulation is positively correlated with occurrences of words in natural speech. For example, 'boy', 'chair', and 'blue' are words that occur with high frequency within natural language, whereas words like 'tandem', 'ocelot', and 'pagoda' occur with low frequency (Karlsen & Snodgrass, 2004, p272).

An orthographic neighbourhood is calculated by seeing how many words can be created by changing one letter of the stimulus word, preserving letter positions (Coltheart, Davelaar, Jonasson & Besner, 1977). For example, 'Marsh' has two neighbours, 'harsh' and 'march'. In contrast, the word 'cover' has 13 neighbours (including coven, covet, cower, hover, lover, mover and rover) (Lavidor, Hayes, Shilcock & Ellis, 2004).

Regular words have regular spelling-sound correspondences, i.e. they are pronounced in the same way that they are spelt e.g. 'dog', 'mint', 'cave'. Irregular words may be 'exception' words, i.e. 'pint', or 'bass', which are spelled similarly to regular words, but pronounced differently, or 'strange' words, i.e. 'piece', or 'friend', which have individual spellings, non-comparable to any others in the English spelling system, and which do not follow spelling-sound correspondences.

This chapter will first examine how the skill of word reading is acquired, and will critically examine the various models of reading available. Following from this, there will be a discussion of what causes some children to have difficulties with the printed word, and some suggestions for overcoming these issues. Subsequently there will be a discussion of the different definitions available for developmental dyslexia, with reference to the specific difficulties encountered, and some suggestions for the cause of these difficulties. An in-depth review of the methods suggested for teaching reading in the early stages of reading development will follow.

Theories and Models of Reading

Several researchers have put forward varying models that attempt to show, in detail, how word reading is carried out. These models can be split into adult models of word reading e.g. dual-route theory and connectionist models, and developmental models of reading e.g. stage models. An additional model of reading will be described, the split-fovea model; however, this model only refers to a special case of reading where there is a central fixation point.

Models of Skilled Adult Reading

Dual-Route Theory

Dual-Route theory states that in skilled adult reading words are read in one of two ways, either by visual sight-word reading, or by phonological recoding (Coltheart, 1978; In Underwood, 1978). Phonological recoding refers to the process of translating letters into sounds by applying letter-sound rules, and identifying the word from its pronunciation. Sight-word reading refers to the process of establishing direct connections between the visual form of the printed words and their respective meanings in memory, acquired through repetitive reading of these words. This theory assumes that these two processes are separate, but that sight-word reading of a word occurs after it has been recoded several times. Regular words are at an advantage over irregular words, as regular words can be read by either route, whereas irregular words can only be read by the visual route.

Coltheart, Rastle, Perry, Langdon, and Ziegler (2001) created a computational dual-route cascaded (DRC) model of word recognition and reading aloud. Unlike the connectionist models reviewed here (see later), rather than using a learning algorithm based on back-propagation, Coltheart *et al*. specified the construction of the DRC model themselves utilising previous empirical and theoretical research on reading. In particular, their model was based around the interactive and competition (IAC) model of McClelland and Rumelhart (1981) and Rumelhart and McClelland (1982). The DRC model was also based around work by Morton (1980) with one important difference; Morton's model was a thresholded processing model, whilst the DRC was a cascaded processing model. According to Coltheart, Rastle, Perry, Langdon, and Ziegler (2001), thresholded processing can be defined as the processing occurring in any module which does not begin to affect subsequent modules at an early point in processing, but where activation is only passed on to the later modules after a threshold is reached in the earlier module. Cascaded processing, however, does not use thresholds within modules, instead, as soon as there is the slightest activation in an early module this flows on to later modules.

Other design choices for the DRC were made such that its spelling-sound correspondence rules for the non-lexical reading route were based on single phonemes only, resulting in the adoption of grapheme-phoneme correspondence (GPC) rules as listed in Rastle and Coltheart (1999). It could also be applied to words of up to eight letters in length, and consisted of three routes, the lexical-semantic route (not yet implemented), the lexical non-semantic route, and the GPC route. Each route consisted of several interactive layers, each containing a set of units. These units represented the smallest individual symbolic parts of the model, such as words in the orthographic lexicon, or letters in the letter unit layer. These units could interact in one of two ways, through either inhibition or excitation. The former is where activation of a unit made it more difficult for another unit to rise; the latter is where the activation of a unit contributed to the activation of other units. A letter caused either an excitation or inhibition of a unit in the orthographic lexicon dependant upon its position. That is, a letter in the Nth set of letter units would excite all units in the orthographic lexicon for every word that contains that letter in the Nth letter position of the word and would inhibit all other units of the orthographic lexicon.

The lexical non-semantic route of the DRC model produced the pronunciation of a word through a series of processes. To begin, the features of a word's letters would activate the word's letter units (in parallel across all letter positions), which would then activate the word's entry in the orthographic lexicon. This word entry in the orthographic lexicon would then activate the corresponding word entry in the phonological lexicon, which in turn would activate the word's phonemes (in parallel across all phoneme positions).

The GPC route of the DRC model would convert a letter string into a phoneme string by applying grapheme-phoneme correspondence rules. Visual features and corresponding letter units would be activated in the same way as in the lexical non-semantic route. The GPC route would then pause for a set number of cycles before beginning to operate on the first letter of the input. The rule set would then be searched until an appropriate rule was found to convert the letter to a phoneme, and then that phoneme's unit in the phoneme system would receive some activation (adding to the activation it was already receiving from the lexical route, which is not yet implemented). After a set number of cycles had passed, the next letter would become available to the GPC route, so that it could translate the first two letters in the input string into a phoneme or phonemes. This process would continue, with one letter being added after every set number of cycles until either the letter string was named, or the final position in the letter units was reached. Thus, the GPC route would assemble the letters into phonology, serially, letter by letter.

The whole DRC model would operate on a series of cycles. During the first cycle, the visual feature units would have been fixed to the features corresponding to the input letter string. This fixing would have meant that during cycle two, activation from the feature level would have reached the letter level. During cycle three, activation would have reached the orthographic lexicon and would have been fed back to the letter level, and thus, to the phoneme level. This process of cascaded activation would eventually lead to a build-up of activation in the phonemic layer, and also to activation feeding back from the phoneme layer to the letter layer. The GPC route would also be contributing activation to the phoneme layer at this time.

The model was evaluated using 7981 words, of which 7898 were read correctly. The incorrectly read words consisted mainly of heterophonic homographs (e.g. 'bow') whereby the model's pronunciation was always the more frequent member of the pair, or of regularisations of irregular words where the first letter was the irregular letter (e.g. 'isle', or 'heir'). The model made these regularisation errors, however, only when simulating speeded reading, but when simulating reading at leisure, these errors were corrected. In terms of non-word reading, the model had a 98.93% accuracy level, based on the reading of 7000 non-words. The majority of the errors made were lexical captures, i.e. where a non-word was given the pronunciation of an orthographically and/or phonologically similar word.

The model was able to successfully simulate many effects of reading aloud behaviour, e.g. frequency effects, regularity effects, and length effects, amongst others; lexical decision behaviours, e.g. word frequency effects, orthographic neighbourhood effects, and pseudo-homophone effects, amongst others; acquired dyslexia behaviours, e.g. surface dyslexia and phonological dyslexia; and other behaviours including Stroop effects. It is the only computational model of reading that can simulate the reading aloud task and the lexical decision task, and which can simulate such a wide variety of reading behaviours.

The developmental application of the dual-route theory (e.g. Frith, 1985; see later) would suggest that reading development first occurs through children using a primitive form of sight-word memory for familiar words, later learning recoding skills, which are used for infrequent or new words, and finally developing a more sophisticated form of sight-word reading. There are, however, several problems with a developmental application of the dual-route theory, the most important one being its inability to explain why novice readers need to learn how to phonologically recode words in order to learn to read. This is because dual-route theory regards sight-word reading as a separate skill and process to phonological recoding. Dual-route theory also holds that sight-word reading is *learned* by rote.

Connectionist Models

Connectionist models of reading challenge the traditional dual-route theory of reading, and instead suggest a 'triangle' model of reading (e.g. Seidenberg & McClelland, 1989). The computerised implementation of the triangle model comprises three sets of simple processing units: a cache of grapheme units representing orthography, a cache of phoneme units representing phonology and a cache of semantic units representing meaning. Words, within this model, are represented as distributed patterns of activity across each set of units. Therefore, when a word is encountered by the network it sparks a pattern of activity across the grapheme units. This activity pattern is transmitted through the network, resulting in a pattern of activity across the phoneme units, thus generating sounds that are consistent with the word's pronunciation.

According to the triangle model, there are two pathways between the written and spoken word; the first being a pathway mapping directly from orthography to phonology, and the second being a pathway which maps from orthography to phonology via semantics. This is not dissimilar to the dual-route theory of reading development; however, there are differences between the two. In the triangle model there is no definite distinction between the types of words processed by each pathway, whereas dual-route accounts suggest that irregular items can only be read by the lexical route and non-words by the phonological route. The dual-route theory also suggests that different types of processing underlie the lexical and sub-lexical routes, whereas it is suggested that a single mechanism underlies all processing in the triangle model. Connectionist models suggest that the connection strengths for patterns that occur frequently are greater than for those that occur less frequently therefore giving regularity and consistency effects in reading.

Seidenberg and McClelland (1989) implemented part of the triangle model, the direct pathway from orthography to phonology, as a three-layer, computerised connectionist network. It learned to pronounce numerous single-syllable orthographic word forms to a level similar to skilled readers; however, its non-word reading remained poor relative to human readers. Subsequently, Plaut *et al.* (1996) improved the model, altering the way in which orthography and phonology were represented. This modified version consisted of a set of grapheme units, coding all single and multi-letter graphemes, connected via a set of hidden units, to a cache of output, phoneme units. This network learned to produce pronunciations for both regular and irregular words and non-words to adult levels. Harm and Seidenberg (1999) investigated the impact of prior exposure to phonological forms of words on subsequent reading training, and discovered that the network was able to learn the dependencies between phonemes occurring in words and, when fully trained, learned to read words and non-words to adult levels.

Connectionist models of reading were initially designed to simulate skilled readers (e.g. Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996), but recently researchers have focused on their performance during training, and how this performance compares with beginning readers (e.g. Powell, Plaut & Funnell, 2006). By switching attention to the training period of these models, the framework can be extended to issues relating to reading development (e.g. Harm & Seidenberg, 1999). Researchers claim that children make use of alphabetic skills, letter-sound knowledge and phonological awareness from their earliest attempts to read, and that these skills assist children in developing partial orthographic representations of words (e.g. Stuart & Masterson, 1992; Savage, Stuart & Hill, 2001). A key feature within this claim is the concept of lexical representations, i.e. unique, internal representations of familiar whole words. The connectionist model proposed by Plaut *et al*. (1996), however, does not include such representations of whole words. Instead, words are represented as distributed patterns of activity across the arrays of grapheme, phoneme, and semantic units. The network thus learns to read by mapping from graphemes to phonemes in a way that reflects the context in which the graphemes occur.

Research by Powell, Plaut and Funnell (2006) aimed to discover whether the connectionist model proposed by Plaut *et al*. (1996) could simulate the same learning behaviours as child readers. At first, the network's non-reading was inferior relative to word reading when compared with the children, and the network made more non-lexical, than lexical errors, which was the opposite pattern to the children. Three adaptations to the training of the network were implemented; an incremental training regime was introduced; grapheme-phoneme correspondence rules were taught; and an array of words found in children's early reading materials was used. These changes were included to replicate more accurately the learning environment of a child. These improvements to the model resulted in significant differences in non-word reading relative to word reading, closely mirroring the children's performance on this measure. Unfortunately, however, the network continued to make non-lexical errors more often than lexical errors, perhaps due to the lack of a semantic pathway.

An alternative connectionist model of reading development was proposed by Zorzi, Houghton and Butterworth (1998), which was based on the dual-route theory. Unlike the majority of connectionist models, the two-layer assembly (TLA) architecture allowed for an additional direct route between orthography and phonology that was not mediated via hidden units. It was argued that this direct contact between orthography and phonology would enable regular grapheme-phoneme correspondences to be learned much more quickly, and therefore the ability to generalise to novel items would be acquired more efficiently. The orthographic representation was strictly position specific, and relative to the onset-rime structure of words (the onset consists of the consonant cluster preceding the vowel (if any) and rime consists of the vowel and any final consonants: e.g. *string* = *str* + *ing; milk* = *m* + *ilk*.) There were no nodes used for complex graphemes, as used by Plaut *et al.* (1996), thus ensuring that any grapheme-phoneme correspondence rules learned by the network emerged only from its attempts to predict onset-rime phonology from onset-rime orthography. The TLA model succeeded in learning novel items very quickly, and as such provided strong support for the dual-route theory in terms of the different underlying processes required for reading familiar versus novel word stimuli. This is in direct contrast to the single-route theory offered by the majority of connectionist models (e.g. Seidenberg & McClellend, 1989; Plaut *et al.* 1996).

Harm and Seidenberg (2004) presented a connectionist model of reading that rather than focusing on the pronunciation of printed words, addresses the processes involved in determining the meanings of printed words. They suggest that it is possible for a proficient reader to be able to decide on the meaning(s) of a word directly from knowledge of its spelling. In alphabetic orthographies, however, where letters represent sounds, an alternative method for obtaining word meanings is available, that of spelling being converted into a phonological representation that is then used to determine a word's meaning. These two routes have traditionally been entitled 'direct' (orthography to semantic) lexical access and 'phonologically mediated' (orthography to phonology to semantic) lexical access. The model proposed by Harm and Seidenberg (2004) varies from other connectionist models in that it focused on how meaning is computed in a system in which both direct and phonologically mediated pathways are available. Their model determines the meaning(s) of a word through activation of a set of semantic units that develops over time based on input from both pathways. This model has an affinity to the dual-route theory in that both visual and phonological processes can activate lexical semantics. Harm and Seidenberg (2004), however, argue that these two components are not independent; instead, together they jointly achieve an efficient solution. What one pathway contributes to the output depends on what the other pathway contributes. This occurs through a back-propagation procedure, whereby if one part of the system fails or is slow for any given item, this generates error. These errors may occur due to inadequate training to have created a mapping, because the mapping is a difficult one, such as spelling to meaning, or because there are ambiguities in the training set that restrict performance (e.g. homophony in the mapping from sound to meaning). The nature of the learning process is such that when one component is slow or unable to perform accurately, this creates pressure for the system to make up the difference elsewhere. This therefore leads each pathway of the system to be sensitive to the successes and failures of the other pathway. Harm and Seidenberg (2004) found that, with both pathways intact, the model computed meanings more efficiently than the paths did independently. The division of labour between the two was affected by lexical properties including frequency and spelling-sound consistency.

Split Fovea Model

One final model of reading which needs to be described is the split-fovea model. This model is only applicable to single word reading when a reader is centrally fixated, however, it is an important model giving valuable insight into hemispheric activity in reading, and the differing reading strategies employed.

It is now generally accepted that the right, as well as the left, hemisphere contributes to the lexical processing in the normal brain (e.g. Chiarello, Shears, Liu, & Kacinik, 2005). In the split-fovea model, the right hemisphere (RH) is assumed to reflect effects generated by the initial letters of English words (presented in the left visual field, LVF), whilst the left hemisphere (LH) would process the end letters (presented in the right visual field, RVF). Specifically, the letters in a centrally fixated word that fall to the left of the fixation point will project initially to the RH while the letters that fall to the right of the fixation point will project initially to the LH. The split-fovea model of word processing also has implications for the dual-route theory, as it proposes that the two hemispheres process printed words in a different way.

Ellis, Young, and Anderson (1988) suggested that a distinctive mode of processing, termed 'Mode A', may be available to the LH but not the RH, if words are presented in a standard format. Mode A involves rapid, parallel processing of the component letters of words and is therefore reasonably insensitive to word-length. Non-words presented in the RVF/LH are unable to benefit from Mode A because it involves rapid transfer of letter information to the visual input lexicon in the LH (which provides access to semantic and lexical–phonological processing) and is therefore only available to words with representations in the lexicon. According to Ellis *et al.*, non-words have to be processed in a different way, termed 'Mode B'. The first step in Mode B processing is to classify each letter form as an instance of an 'abstract letter identity' so that different versions of the same letter, such as 'D' and 'd', will converge upon the same abstract graphemic representation, as will 'E' and' e', 'Q' and 'q', etc. This line of reasoning would suggest that letter-sound conversion procedures located in the LH operate upon abstract graphemic representations, allowing non-words (or unfamiliar words) to be read aloud. Words in unusual formats presented in the RVF/LH are unable to benefit from Mode A processing but their letters can be converted into abstract graphemes (Mode B) and can access the lexicon by that length-sensitive route. The Ellis *et al.* theory further proposed that the RH only ever has access to Mode B processing. It converts all letter inputs into abstract graphemes, which are the level of representation at which information is transmitted from the RH to the LH. Length effects arise in both hemispheres in the process of converting letter forms into abstract letter identities, which explains why length effects of similar magnitude occur for non-words and abnormally formatted words in the RVF/LH and for all stimuli in the LVF/RH.

Models of Reading Development

Stage Models

The dual-route and the connectionist models of reading development have been shown as detailed models of skilled adult reading. These two models can be applied to theories of reading development, in particular, to the two stage theories of reading development; Frith's (1985) theory, which relates closely to the dual-route model, and Ehri's (2004) theory, which relates closely to connectionist models.

Frith's (1985) stage-model of reading development comprises of a three-phase sequence of strategies used in reading, and a six-step model of how these phases relate to reading and writing development. Ehri's (2004) stage-model of reading development is a four-phase theory of sight-word reading. Both models assume that reading skills are acquired through a developmental sequence of steps, where new strategies are introduced at different points in the sequence. Both models hold that their respective phases are completed in strict sequential order and that no phase can be ‘skipped’ or ‘jumped’ over.

The first stage of Frith’s (1985) theory is entitled the ‘logographic’ stage, and suggests that readers in this stage can instantly recognise familiar words, but that letter order is largely ignored and all connections between the printed word and its meaning are purely visual and non-phonetic. Readers in this stage read words by forming connections out of visual cues that are arbitrarily related to the word’s meaning and bear no relationship to the word’s pronunciation. The second stage of Frith’s (1985) theory is entitled the ‘alphabetic’ stage, and suggests that readers at this stage apply their knowledge of letter-sound correspondences to read words. Therefore, the primary association in this stage is between a letter and its pronunciation, not, as in the previous stage, between a spelling and its meaning. This does not mean that readers stop processing word meanings, only that they do not depend on this for sight-word reading.

The first three stages of Ehri’s (2004) model can be closely matched with the first two stages of Frith’s (1985) model. The first stage of Ehri’s theory is the ‘pre-alphabetic’ phase, and suggests that children in this phase generally use visual cues to recognise words even if they know a few letter names or sounds, as they do not know how to use that information at this stage. Ehri’s second phase of reading development is the ‘partial alphabetic’ phase, and suggests that readers in this stage are beginning to relate letter-sounds to printed words, although they do not yet look at all the letters in a printed word in order from left to right. Readers in this stage often pay close attention to the initial and end letters of words, without processing the letters in-between. The third stage of Ehri’s model is the ‘full alphabetic stage’ and here readers are able to make connections between letters and sounds all through the word.

The final stages of these two theories differ strongly in their views of how fluent readers can sight-read familiar words. Frith (1985) entitles the final stage as the 'orthographic' stage, and suggests that readers in this stage can instantly analyse words into orthographic units without phonological conversion (see also McCaughey, Juola, Schadler & Ward, 1980). Ideally, in this stage, the orthographic units will coincide with morphemes (the smallest meaningful unit of language, e.g. '*man'*, or *'ed'* as in the past tense ending in *'walked')*, and words will be internally represented as abstract letter-by-letter strings. This implies that precise orthographic representations are acquired through giving equal attention to all letters in a word.

Ehri's (2004) final stage is entitled the 'consolidated alphabetic' phase. It proposes that whilst in the second and third stages, the systematic visual-phonological connections between letters seen in words and their pronunciations were incomplete; these connections are completed in this final stage. Readers in this phase also fully analyse spellings as visual symbols for determining phonemic features in pronunciation. Therefore, according to Ehri, sight-word reading involves remembering systematic connections between spellings and pronunciations of words. This implies that letter-sound knowledge is a necessity, as this knowledge is needed to form a complete network of visual-phonological connections in lexical memory.

In summary, Frith (1985) assumes that each stage capitalises on the previous ones, due to a merging of the old and new strategies, i.e. certain components of the old strategy will be retained in order to enhance the new one. Ehri (2004) assumes that each stage is determined by the types of associations formed between a visual cue seen in print and the information about the word stored in memory. Her model suggests that readers in the final stage of reading can use the consolidated alphabetic strategy to sight-read short, regularly spelled, monosyllabic words. Readers may regress, however, to the former three stages to process multi-syllabic words, or words that have spellings defying the reader's letter-sound knowledge.

The differences between the final phases of these two strategies are due, in part, to the different emphasis placed on the importance of employing phonological recoding skills in reading. Frith's (1985) orthographic stage of reading development suggests that readers can instantly sight-read words without the use of phonological conversion, and that this is achieved by activating internal representations that are exact in terms of letter-by-letter detail. This implies that readers need only to analyse the letters present in the written word, and not the sounds in its pronunciation, in order to access the word's meaning. This slightly contrasts the alphabetic stage, which stresses phoneme-grapheme correspondences. Frith explains this divergence between strategies, by stating that certain components of each stage are merged into the new strategy at each phase (see also Karminloff-Smith, 1979, 1984; Bryant, 1982). The second and third stage-related skills of Frith's model do not lend themselves to this, other than that both use a systematic approach either to letter-sound correspondences, or to letter order.

Ehri's (2004) consolidated alphabetic phase of reading development suggests that readers can instantly sight-read words due to phonological recoding. This skill is introduced in the partial-alphabetic stage, built-upon in the full alphabetic phase, and completed in the consolidated alphabetic phase, where connections in memory between the entire sequence of letters in a spelling, and its phonemic constituents in the word's pronunciation, are established. The consolidated alphabetic phase assumes that the letters in a spelling fully determine the word's pronunciation, and consequently the meaning; this excludes words with similar pronunciations. Ehri suggests that as letter-sound knowledge is used in the initial stages of reading, this process should be retained, developed, and should participate in the reading-by-memory operation. This will result in a visual route, lined with phonological information, leading to lexical memory. Thus, according to Ehri's model, letter-sound knowledge is a necessity, however, according to Frith (1985) it is not just unnecessary, but part of a strategy discarded by the more fluent, mature reader when reading familiar words.

The underlying assumptions of Ehri's (2004) theory focus on the connections linking spellings to pronunciations. The model assumes that systematic visual-phonological connections exist between spellings and their pronunciations; therefore, word-specific connections are employed to read words. This would imply that readers use all available systematic relations, and do not just memorise the entire form of the word. This is in direct contrast to Frith's (1985) model, which assumes that the printed form of the word instantly accesses its meaning in memory, via internal representations of letter-by-letter detail, implying that the word form in its entirety is rote-memorised.

From these two theories, it is possible to draw several conclusions about reading development. Frith (1985) and Ehri (2004) propose that normal reading skills occur through a developmental sequence of steps, probably fitting into three or four broad phases, each incorporating a particular strategy. It can be suggested that both logographic skills and phonological recoding skills play an important role in a child becoming a fluent reader.

The two theories also leave several questions unanswered as regards the way in which sight-word reading is processed. Frith's (1985) model suggests that once a word has been recoded several times, readers memorise the entire form of the word and further suggests that readers internally represent the word as a letter-by-letter string, approached in a systematic way. Ehri's (2004) theory disagrees with the concept of rote memory entirely, and suggests that words are approached phonologically in a systematic way, and that readers will use all systematic information available to them, eliminating the need to memorise word forms in their entirety. Unfortunately, it is beyond the scope of this thesis to answer these questions; however, this would be an interesting avenue for future research.

Interim Summary

Three adult reading models (dual-route theory, connectionist models, and the split-fovea model) and two developmental theories (one dual-route and one connectionist) have been described and discussed. Applications of the dual-route theory suggest that reading occurs by children initially employing a primitive sight-word memory for familiar words, then learning phonological recoding skills, used for infrequent or new words, and finally developing a morpheme-based form of sight word reading. The split-fovea model suggests that the LH processes words rapidly via parallel processing and the RH processes words serially, letter by letter. Stage models suggests that normal reading skills occur through a developmental sequence of steps, probably fitting into three broad phases each incorporating a particular strategy, with both logographic and phonological recoding skills playing an important role in a child becoming a fluent reader. Connectionist theories mostly adopt the triangle model, which consists of three sets of simple processing units: a cache of grapheme units representing orthography, a cache of phoneme units representing phonology and a cache of semantic units.

A major flaw of the application of the dual-route theory in Frith's (1985) model is its inability to explain why beginner readers need to learn how to phonologically recode words in order to learn to read. Logographic reading might be refined into the orthographic approach without developing phonological skills, the only casualty being non-word reading. This is due to the dual-route theory regarding sight-word reading as a separate skill and process to phonological recoding. Dual-route theory also holds that sight-word reading is learned by rote. Support for the dual-route theory comes from an unusual type of connectionist model, the TLA (Zorzi, Houghton & Butterworth, 1998). This model employed a direct route between orthography and phonology that was not mediated via hidden units. This direct contact between orthography and phonology enabled regular grapheme-phoneme correspondences to be learned quickly, and therefore the ability to generalise to novel items was acquired efficiently. The findings from this model suggest that it was possible to show the reading process as utilising two different underlying processes; one utilising lexical knowledge for familiar words, and a second, separate process utilising spelling-sound mappings for novel words.

The two stage-theories of Frith (1985) and Ehri (1992) offer alternative explanations to the dual-route theory, although there are some similarities between them. Dual-route theory assumes that the processes used for reading familiar words are separate from the processes used for reading novel words, but that sight-word reading occurs when a word has been recoded several times and is therefore familiar. Frith's (1985) theory of reading development proposes that there are connections between the various phases of acquisition. This occurs through a convergence of skills between the strategies, contradicting dual-route theory. Frith (1985), however, does agree with dual-route theory in that sight-reading of familiar words does not involve phonological recoding. Ehri's (1992) theory agrees that letter-sound relations are used initially to read an unfamiliar word; however, she also states that this process would be retained and applied in order to sight-read familiar words as well.

Dual-route theory holds that regular words are at an advantage over irregular words, as regular words can be read by either route, whereas irregular words can only be read by the visual route. Connectionist models also hold that regular words are at an advantage over irregular words; however, they reject the notion that regular and irregular words are processed by individual sub-systems. Rather, these models are sensitive to the statistical regularities inherent in the orthography of the words on which they are trained. The connection strengths for patterns that occur frequently are therefore greater than for those that occur less frequently, and it is this that leads to regularity and consistency effects in reading.

These different theories of reading development suggest different reasons for why some children fail to learn to read, and each advocates different methods of teaching reading. The dual-route theory infers that a difficulty in reading novel words will slow down the expansion of a sight vocabulary. That is, if children have problems learning and generalising letter-sound rules, then the development of word recognition would be compromised. If these difficulties were severe, it could result in the children being forced to learn all words through the visual route, and consequently, these children would show no advantage for reading regular words over irregular words because the phonological route would remain unused (Snowling, 2000). Dual-route theory might suggest that children should be taught to read via a whole-word, meaning emphasis program. The aim of this is to teach children to recognise word meanings by sight; this is accomplished through learning to read visually distinctive words (e.g. farm, rabbit, wagon) by reading and re-reading these words, both on flash cards, and in meaningful contexts (Richek, 1977-1978). This process also leads to children acquiring recoding skills through letter-sound relations. An alternative method of teaching children to read is the phonics instructional program, which would develop the phonological route in the dual-route model. This strategy involves the children learning phonological recoding by reading words which are of regular spelling and contain letter-sound relations that the children have been taught (Richek, 1977-1978). The split-fovea theory would also advocate the use of the whole-word, meaning emphasis program and the phonics instructional program.

Causes of Reading Disorders

The stage theorists, Frith (1985) and Ehri (1992), both attempt to describe the causes of reading failure, and the stage at which classic developmental dyslexia arises. They both agree that this occurs due to an arrest at a particular stage of the normal reading development sequence. However, they disagree about the stage in which this happens. Frith believes that each phase in the normal developmental sequence depicts a particular type of dyslexia. According to her model, classic developmental dyslexia is the failure to develop alphabetic skills, and therefore corresponds to an arrest at stage one. This would imply that dyslexic children are able to use logographic skills, but not alphabetic or orthographic skills. After the point of arrest, the child would be expected to adopt compensatory strategies (Frith, 1985). Ehri believes that dyslexic children suffer an arrest in normal reading development during stage two; the partial alphabetic stage. Therefore, they are not able to phonologically recode words accurately, nor rapidly; this therefore precludes learning to read words by sight using the cipher strategy. According to Ehri's theory, in developing sight-word reading skills, dyslexic children never progress far enough beyond stage two to become fully adept at stage three.

The stage-theorists, Frith (1985) and Ehri (1992), therefore propose that teaching methods in schools should consist of both a whole-word instructional programme and specific phonics instruction. It could be suggested that children would benefit from first learning whole-words, to aid them during the logographic stage, then to move across to phonics training when the child progresses through to the alphabetic stage of development. An alternative would be to utilise teaching methods that incorporate the two elements equally.

Connectionist models of reading development argue that children who struggle to learn to read do so, not because of an inability to read novel words, but because of difficulties learning letter-sound rules, and with generalising these rules to novel words (e.g. Seidenberg & McClelland, 1989; Snowling, 2000). That is, knowing the sound of '*p*' in the word '*tap*' does not necessarily mean that these children would know the sound of the letter '*p*' in '*pot*'. Connectionist theorists emphasise the value of training on the relationships between orthography and phonology at the sub-word level early in literacy training (e.g. Powell, Plaut & Funnell, 2006; Hutzler, Ziegler, Perry, Wimmer & Zorzi, 2004). Therefore, the connectionist models would predict better reading performance when children had been given phonics training.

General versus Specific Reading Difficulties

Developmental dyslexia was first described just over 100 years ago by two British doctors; Kerr (1896) and Morgan (1896). Morgan described a boy of normal intelligence who had failed to learn to read. This boy showed many characteristic signs of dyslexia, including transpositions of letters (e.g. Percy – Precy), dysphonetic spelling errors (carefully – calfuly), and substitutions of phonemes (peg - pag). Unfortunately, however, despite the volume of research surrounding developmental dyslexia, there is no universally accepted definition of this reading disorder. Perhaps the most frequently cited definition of dyslexia comes from the World Federation of Neurology (1968; cited in Critchley, 1970) who stated that dyslexia is "A disorder manifested by difficulty in learning to read despite conventional instruction, adequate intelligence and socio-cultural opportunity. It is dependant upon fundamental cognitive disabilities which are frequently of constitutional origin." Unfortunately, this definition uses vague and ill-defined terms, not explaining what is meant by 'adequate intelligence', 'adequate socio-cultural opportunity', nor how much 'difficulty in learning to read' needs to be apparent to label a child as dyslexic. The definition also relies on exclusion criteria, only stating what dyslexics should not be.

More recent definitions of dyslexia argue that it is a difficulty in reading due to a core phonological deficit. The Orton Dyslexia Society of the USA (now the International Dyslexia Association) (1994; cited in Snowling, 2000, p24-25) states, "Dyslexia is one of several distinct learning disabilities. It is a specific language-based disorder of constitutional origin, characterised by difficulties in single-word decoding, usually reflecting insufficient phonological processing abilities. These difficulties in single-word decoding are often unexpected in relation to age or other cognitive abilities; they are not the result of generalised developmental disability or sensory impairment." This definition also presents with some difficulties, as some terms are still vague which may lead to it missing some of the children it aims to identify, and overall it is difficult to falsify.

Snowling (2000, p137) states, "it is rare to find a dyslexic child who does not have some kind of phonological problem if they are tested using sensitive enough measures." This type of thinking is risky in that it may lead to phonological awareness difficulties being given a higher status in dyslexia than other symptoms shown. For example, the British Dyslexia Association states, ''Dyslexia is a specific learning difficulty, best described as a combination of abilities and difficulties that affect the learning process in one or more of reading, spelling, and writing. Accompanying weaknesses may be identified in areas of speed of processing, short-term memory, sequencing and organisation, auditory and/or visual perception, spoken language and motor skills. It is particularly related to mastering and using written language, which may include alphabetic, numeric and musical notation'' (Crisfield, 2002, p. 67). This definition therefore suggests that other difficulties are manifest in dyslexia, and should be treated with equal importance in its diagnosis and treatment.

Other research, however, has shown neurological differences between dyslexics and non-dyslexic readers, with dyslexics showing less activation across the left-hemisphere. This has been shown during rhyme processing and short-term memory tasks (Palesu *et al.*, 1996), reading aloud (Brunswick, McCrory, Price, Frith & Frith, 1999), and other tasks involving various levels of phonological demands, including single letter and non-word rhyme judgements (Shaywitz *et al.*, 1998). During these studies, dyslexics consistently showed an under-activation of the left-posterior regions of the cortex, especially Wernicke's area (associated with the processing of words), angular gyrus (used for storing letter-sound rules) and striate cortex (the main receiving area for visual signals). Dyslexics also showed over-activation of the left frontal (usually used in controlling language related movement) and right posterior regions (usually used in visuo-spatial processing), suggestive of anomalous brain function (possibly compensatory) when compared to non-dyslexic readers. These findings suggest that a definition showing dyslexia to be a specific form of language impairment, affecting the way in which the brain encodes written and phonological information, would perhaps be most accurate.

Another area of debate is how to distinguish between a dyslexic reader and any other poor reader. Some researchers have suggested that in order to do this the relationship between reading skill and intelligence needs to be examined. There is a consensus amongst researchers that, within the general population, there is a positive correlation between an individual's intelligence quotient (IQ) score and their reading level (Snowling, 2000). By applying a regression analysis to this relationship, predictions can be made for any individual's reading age when given their chronological age and IQ score. A child is considered dyslexic if their reading ability is significantly below their expected reading ability. If, however, a child has a reading age significantly below their chronological age, but not out of line with expectation due to their IQ being relatively low, (s)he is considered to have a general reading difficulty. Problems arise from this approach to classification from its reliance on the construct of IQ. IQ is a vague and imprecise label, used to encompass and describe many skills under one broad term, meaning that IQ is perhaps only loosely connected to reading ability. Another possible problem occurs because IQ is a general measure of intelligence, testing both verbal IQ and performance IQ. It can therefore be suggested that if low verbal IQ can be a consequence of reading disability (due to a lack of exposure to print, or due to lowered comprehension of written texts) it could also mask the specificity of a child's reading problem.

An alternative method of distinguishing between general poor readers and those with specific reading difficulties was proposed by Stanovich (1991). Stanovich argued that it is reasonable to expect that a child's ability to understand what they hear to be at a similar level to their ability to understand what they read (providing that they can read). If this line of reasoning is true, then a child whose reading comprehension is below their listening comprehension could be described as having a specific reading difficulty.

Further difficulties in diagnosing dyslexia arise from the assortment of problems it can produce and the varying severity of those problems in individuals. In an attempt to solve this problem, some researchers have attempted to group together individuals with particular symptoms, thus creating sub-groups of dyslexia. Boder (1973) suggested three different sub-categories of dyslexics; 'dysphonetic', 'dyseidetic' and combination dyslexics. 'Dysphonetic dyslexia' manifests as an inability to sound out printed words unless immediately familiar, combined with a limited sight-word vocabulary. Word-attack skills, including phonic analysis and synthesis, present difficulties for dysphonetic dyslexics, who are limited to spelling words 'by eye' alone, and are unable to spell words which are not in their sight vocabulary. Reading errors are usually either visual (e.g. reading *house* as *horse*), or semantic (e.g. reading *laugh* as *funny*) in nature. 'Dyseidetic dyslexia' manifests as an inability to build up a sight-word vocabulary due to difficulties with memorising the visual shapes of words, and therefore an inability to read words as a whole. Dyseidetic dyslexics read laboriously 'by ear', utilising both phonetic analysis and synthesis, sounding out both familiar and novel words. The majority of reading and spelling errors are phonetic in nature, i.e. they usually take the form of phonologically correct pronunciations and misspellings. Combination dyslexics have the most difficulty reading, as they are unable to draw upon either visual or phonic skills. Other researchers have also attempted to create sub-groups within dyslexia diagnosis, e.g. Marshall and Newcombe (1973) who used the terms 'surface dyslexia' (presenting with similar symptoms to Boder's (1973) 'dyseidetic' dyslexics) and 'deep dyslexia' (similar to Boder's (1973) 'dysphonetic' dyslexics).

Teaching Reading

There are almost as many ways to teach reading as there are teachers, however, teaching methods can largely be grouped into three broad categories. These consist of a 'whole-word' approach to reading, a 'whole-language' approach to reading, and 'systematic phonics' teaching. The whole-word approach involves teaching the beginning reader to look at each word as a whole, to memorise its shape, and any distinctive features, and to arbitrarily relate the word form to its meaning. High frequency monosyllabic words would be taught first, regardless of their spelling-sound regularity. Teaching would usually involve the utilisation of 'flash-cards', presenting the words out of meaningful contexts, with the beginner reader using the letter-shapes and overall word shape to recognise the word and its meaning.

The whole-language approach to teaching reading encourages the beginner to use all available cues for word recognition. This would include letter-sound mappings, letter shapes, word-shapes, accompanying pictures and sentence-context. Words would be presented in sentences, and larger text forms, such as paragraphs. Stories and poems would form a major part of the teaching curriculum, supplemented by related pictures, with emphasis given to semantics and pronunciation of unfamiliar words.

The systematic phonics approach teaches novice readers all of the major grapheme-phoneme correspondences in a clearly defined sequence. Letters are introduced by their sounds, not their names (i.e. *A* = *'ah'* as in *'cat'*, *B* = '*bu*' as in '*bat*'), and are first introduced with their most commonly used sounds, and then their alternative sounds are taught (e.g. '*A*' = '*cat*', '*water*', '*bath'*, '*cake*'). Digraphs are then introduced (e.g. '*ie*', '*ea*', '*ou*'), followed by consonant blends (e.g. '*str*', '*fl*', '*br*') and word endings (e.g. '*ing*', '*ed*', '*er*'). There are two main approaches within systematic phonics teaching, that of analytic phonics, and that of synthetic phonics. The American National Reading Panel described analytic phonics methods as teaching children whole-words before teaching them to analyse these into their component parts, and emphasise their larger sub-parts of words (i.e. onsets, rimes, phonograms, spelling patterns) as well as phonemes. They also described synthetic phonics methods as emphasising the teaching of students to convert letters (graphemes) into sounds (phonemes) and then to blend the sounds to form recognisable words (National Institute of Child Health and Human Development, 2000). One of the principles of synthetic phonics is that a reader should never be asked to read something that is too difficult for them, or that they do not have the skills to read.

The *National Literacy Strategy* (NLS) *Framework for Teaching* (Department for Education and Employment (DfEE), 1998) specified the extensive objectives and the comprehensive teaching format of literacy teaching in primary schools in England. Due to the Office for Standards in Education (OfSTED) inspection processes, the content and methodology promoted by the NLS were effectively imposed on all teachers (Wyse & Styles, 2007). In the year 2000, the use of analytic phonics to teach early reading skills was embraced by the NLS, when they introduced *Progression in Phonics* materials (National Literacy Trust (NLT), 2007). This move sparked heated debate amongst researchers, with some advocating a move back to 'mixed methodology', incorporating phonics and a 'whole-language approach' to make sense of text (e.g. Dombey, 2006), and other researchers promoting the exclusive employment of synthetic phonics (e.g. Johnston & Watson, 2004, 2005). Partly due to the controversy created by the publication of Johnston and Watson's (2005) research into synthetic phonics, the English government commissioned an investigation into the teaching of early reading, to be headed by the education consultant, Jim Rose. The resulting review (Rose, 2006) was dubbed 'the Rose report'.

The Rose report (Rose, 2006), was based upon ten schools, and reviewed the teaching of early reading. These schools had been pre-judged as being representative of the best practice in the teaching of phonics. The report concluded that teachers should be required to teach reading through synthetic phonics as the prime approach in learning to decode and encode print. It sparked much debate, with researchers and teachers alike divided over the best way of teaching phonics to children, and how best to include this teaching in the curriculum. In support of synthetic phonics teaching, research by Johnston and Watson (2005), and Foorman *et al.* (1997), found that teaching through a synthetic phonics approach is more effective than employing an analytical phonics approach.

Wyse and Styles (2007), however, suggest that there is insufficient evidence to promote synthetic phonics as the preferred method of teaching of early reading, and propose that any method of systematic phonics teaching would offer equally effective teaching. This view is also supported by a comprehensive study by the US National Reading Panel (NRP) (National Institute of Child Health and Human Development, 2000). The NRP found that systematic phonics programmes are all significantly more effective than non-phonics programmes, although they suggest that these systematic phonics programmes do not appear to differ significantly from each other in their effectiveness (however, this review was carried out nearly 8 years ago, and new studies have been published since then). The NRP also concluded that systematic phonics instruction is best introduced to beginning readers, and that for these children it is highly beneficial. The NRP further suggests, however, that phonics should not become the primary component in a reading programme; neither in the amount of time allocated to it, nor in the significance attached. Work by Torgerson, Brooks and Hall (2006) suggests that although phonics is an essential part of literacy teaching, it should be used in combination with other methods. In support of this, Berninger *et al.* (2003) found that phonics teaching was most effective when combined with reading comprehension training. The Australian government (2005) also carried out a review of the teaching of early reading and advocates an integrated approach, including systematic phonics teaching, comprehension, fluency, and grammar. Synthetic phonics teaching, however, does not support such an integrated approach. Instead, it emphasises discrete teaching of phonemes and graphemes, before introducing sentences or whole texts.

Torgerson, Brooks and Hall (2006) carried out a systematic review of approaches to the teaching of reading, including only randomised controlled trials (RCTs). They concluded that there is not enough evidence to support the view that any one form of systematic phonics is superior to any other, and that there is no significant difference in effectiveness between synthetic and analytic phonics approaches. Johnston (2008), however, points out that Torgerson *et al.* used only three studies in order to carry out a meta-analyses, one of which was unpublished, and that mistakes were made when deciding which studies should be included, and which data should be entered into the calculations.

There have been countless reading intervention studies trialling different types of reading instruction methods for the beginning reader, some of which are reviewed here. Torgesen *et al.* (1999) identified 180 5-year-old children ‘at risk of failure’ based on poor performance of letter naming and phoneme identity. These children were split into four groups, three of which had four x 20 minutes training per week on a one-to-one basis for 2.5 years. Group 1 were taught phonological awareness and synthetic phonics, and were given reading practice with high frequency words. They read stories that needed decoding skills and contextual awareness, and discussed the meaning of words and stories with their teacher. Group 2 were taught via embedded phonics, including tuition in whole-word recognition of a small group of words using drills and word-games to learn these words. They examined letter-sound correspondences in the context of these words, and wrote these words in sentences. Their reading practice focused on acquiring word-level reading skill, but not context, thus extending sight-word vocabulary and single word decoding. They discussed the meaning of words and stories with their teacher. Group 3 were given regular classroom support, based on the curriculum (thus varying slightly across children) including a variety of teaching methods from phonics-based instruction through to approaches focusing on meaning. Their tuition was more similar to group 2’s instruction than to group 1. Group 4 received no intervention. The researchers found that those children receiving phonological awareness and synthetic phonics training (group 1) had the best overall outcome, and were significantly better at phonemic decoding than the other three groups. Groups 1 and 2 were equally good at word identification, and both outperformed groups 3 and 4.

Cunningham (1990) taught three groups of 6-year-old children twice a week for ten weeks. Group 1 received instruction in phonemic awareness with a “skill and drill” approach where the procedural knowledge of segmentation and blending of phonemes were taught. Group 2 were taught using a "meta-level" approach that alongside teaching the procedural knowledge of segmentation and blending also strongly emphasised the application, and value, of phonemic awareness within the context of reading, i.e. letters were associated with the phonemes. Group 3 listened to stories and discussed them with their teacher, effectively acting as a control group. A significant improvement in reading achievement was observed for both experimental groups, however, the children in group 2 performed significantly better on reading achievement than the skill and drill experimental group. Similarly, Ball and Blachman (1991) ran an intervention with 5-year-old children. The group who were provided with training in phonological awareness and spelling-sound correspondences progressed better in reading than another group taught letter names and sounds in the context of more general language activities, when compared to a control group who received no intervention. Bradley and Byrant (1983) carried out a two-year intervention with children beginning at age 6 who had been diagnosed as 'at risk' based on their performances on sound categorisation. Those children who had been taught phonological awareness training combined with letter-sound correspondences had better reading performances than children taught just phonological awareness, children taught just semantic categorisation, and unseen controls.

MacKay and Cowling (2004) carried out two intervention programmes using *Toe-by-Toe*. The first intervention consisted of 24 secondary school pupils with low reading ages who were split into two groups; one receiving normal school learning support, and the other receiving *Toe-by-Toe* instruction for twenty minutes every school day for three months. The control group made average gains of five months, whilst the experimental group gained on average forty-two months. The second intervention used 104 children aged over 11 years, and post-testing after five months, before the intervention programme was complete, pupils showed average gains of 14 months.

The results from these intervention studies indicate that phonological awareness training alone is not as effective in promoting reading as training that also emphasises letter-sound correspondences. It is therefore possible to suggest that children best learn to read by drawing their attention to letters and speech sounds to facilitate the development of mappings between orthography and phonology.

The National Reading Panel (National Institute of Child Health and Human Development, 2000), however, conducted a quantitative meta-analysis evaluating the effects of phonemic awareness instruction on learning to read and spell. The main focus of the study was to test whether phonemic awareness instruction was effective in helping children learn to read, and if so, under what circumstances and for which children it was most effective. The findings of this meta-analysis were reported by Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh and Shanahan (2001). Ehri *et al.* reported that, in order to qualify for the analysis, studies had to meet five criteria. These were that they had to have adopted an experimental or quasi-experimental design with a control group, and to have appeared in a refereed journal. They also had to have tested the hypothesis that phonemic awareness instruction improves reading performance over alternative forms of instruction or no instruction, to have provided instruction in phonemic awareness that was not confounded with other instructional methods or activities and to have reported statistics permitting the calculation or estimation of effect sizes. Fifty-two studies met these criteria. There were three reader groups: normal readers, at-risk children, identified by either low phonemic awareness ability or low socioeconomic status, and reading disabled students. The primary statistic employed in this study was effect size, indicating whether and by how much, performance of the treatment group exceeded that of the control group. Three outcomes were of primary interest: phonemic awareness, reading, and spelling. The meta-analysis showed that phonemic awareness instruction significantly improved acquisition of phonemic awareness, significantly improved reading ability, and significantly improved spelling ability, and in all cases was more effective than other forms of instruction or no instruction. The greatest gains, however, were shown when letters were used in the training.

Summary

As has been shown, reading is not an innate skill, but one that must be learned. There are several current theories as to how a beginner first learns to understand the printed word. Models of skilled adult reading include the dual-route, the connectionist, and the split-fovea models, and there are related models of reading development which propose stages in reading development. Dual-route theory might suggest that children should be taught to read via a whole-word, meaning-emphasis programme to develop the visual route, and also a phonics instructional program to develop the phonological route. Phonics training is also advocated by connectionist theorists. The split-fovea theory, and the two stage theorists, Frith (1985) and Ehri (2004) would also advocate the use of both the whole-word, meaning emphasis program and the phonics instructional programme.

Unfortunately, some children do not become fluent readers. These children can be classified into two broad categories, those with a specific reading disorder, e.g. dyslexia, and those with a general reading disorder. The definition of dyslexia is still under debate, with most researchers split between a definition focusing on a core phonological deficit, and a definition highlighting neurological language impairment. Also under debate is how to distinguish between a dyslexic reader and any other poor reader. Some researchers suggest that a child reading below their IQ-suggested capabilities is dyslexic (e.g. Snowling, 2000), and others suggest that a child whose listening comprehension exceeds their reading comprehension is dyslexic (e.g.

Stanovich, 1991). There have also been attempts to create sub-groups of dyslexics based on grouping together those with particular symptoms (e.g. Boder, 1973; Marshall and Newcombe, 1973).

A further area of debate is how to teach children to read. Broadly speaking there are three main categories of teaching methods; a ‘whole-word’ approach, a ‘whole-language’ approach, and ‘systematic phonics’ teaching. There are two main approaches within systematic phonics teaching; analytic phonics and synthetic phonics. At the beginning of the millennium, analytic phonics was implemented in the majority of schools in England. This resulted in much controversy amongst reading researchers and teachers, and in response to this the English government commissioned an investigation into the teaching of early reading, dubbed the ‘Rose report’ (Rose, 2006). This report concluded that reading should be taught primarily through synthetic phonics, which created further debate, with researchers divided over the best way to teach phonics to children, and how best to include this in the curriculum. Many intervention studies have been carried out to compare and test the effectiveness of synthetic phonics, analytic phonics, and other forms of reading instruction. The majority have found that systematic phonics instruction is most effective, although researchers differ over whether they advocate synthetic phonics over analytic phonics.

This thesis will closely examine the reading strategies adopted by an adult dyslexic, and by secondary school pupils with reading difficulties, with an aim to explain these reading patterns through existing reading models and theories, and to explore whether less helpful reading strategies can be altered in order to improve reading ability. This will be achieved by examining in detail the approach to reading taken by a well-compensated adult with developmental reading problems, and then report on the effects of an intervention with secondary age school pupils based on phonics principles.

CHAPTER TWO

DB: A Dyslexic Case-Study

Although developmental dyslexia was first described in 1896 (see Kerr, 1896; and Morgan, 1896), there is no universally accepted definition of this reading disorder. In fact, there are three main theories of developmental dyslexia; the cerebellar, the magnocellular (auditory and visual), and the phonological theories.

The cerebellar theory (Nicholson & Fawcett, 1990; Nicholson, Fawcett & Dean, 2001) claims that dyslexics have a mildly dysfunctional cerebellum, which is the cause of several cognitive difficulties. The cerebellum plays a role both in motor control and in automatisation of over-learned tasks. Impaired motor control may affect speech articulation, which would therefore lead to deficient phonological representations. Further to this, impairment in the capacity to automatise would affect, amongst other things, the learning of grapheme-phoneme correspondence rules.

The magnocellular theory attempts to unify two other theories of developmental dyslexia, the auditory and the visual theories, as well as the cerebellar theory (Ramus *et al.*, 2003). The auditory theory specifies that dyslexics have a deficit in the perception of short or rapidly varying sounds, whereas the visual theory suggests that dyslexia is due to an impairment of the processing of letters and words on a page of text. The latter is based on the division of the visual system into two distinct pathways that have different roles and properties: the magnocellular and the parvocellular. The visual theory suggests that the magnocellular pathway is selectively disrupted in dyslexia, leading to deficiencies in visual processing, and via the posterior parietal cortex, to abnormal binocular control (Stein & Walsh, 1997). The magnocellular theory suggests that the magnocellular dysfunction is not restricted to the visual pathways, but is generalised to all modalities (visual, auditory and tactile). It further suggests that as the cerebellum receives massive input from various magnocellular systems in the brain, it would also be affected by the general magnocellular defect (Stein & Walsh, 1997).

Many researchers believe that dyslexia is caused by a combination of phonological impairment and magnocellular deficits (e.g. Talcott, Hansen, Willis-Owen, McKinnell, Richardson & Stein, 1998; Valdois, Bosse, & Tainturier, 2004). Talcott *et al.* suggest that the reading and spelling errors of irregular words by dyslexics may be due to deficits in visual coding, representation and memory. They conducted a study of 36 adults, half of which were dyslexic, and asked them to complete two tasks of visual magnocellular function, random dot kinematograms (RDK) and critical flicker fusion (CFF). They found that dyslexics were less sensitive to the CFFs and to detection of coherent motion in RDKs. The motion and flicker sensitivity scores for the two groups were strongly correlated to non-word reading ability, with lower sensitivity correlated with lower non-word reading ability. This therefore implies that visual deficit problems occur in dyslexia, alongside phonological deficits. Other researchers believe that developmental dyslexia is caused by a phonological deficit, with rare cases of non-phonological problems occurring (e.g. Ramus *et al.*, 2003).

The phonological theory of dyslexia suggests that dyslexics have a specific impairment in the representation, storage and/or retrieval of speech sounds (Ramus *et al.*, 2003). This theory explains dyslexics' reading impairment as being due to poorly represented, stored, or retrieved, grapheme-phoneme correspondence rules. The phonological theory is the most widely accepted hypothesis as the origin of dyslexia (Valdois, Bosse, & Tainturier, 2004). Indeed, most recent definitions of developmental dyslexia suggest that it is a difficulty in reading caused by a core phonological deficit (e.g. Stanovich & Siegel, 1994). The Orton Dyslexia Society of the USA (now the International Dyslexia Association) (1994; cited in Snowling, 2000, p24-25) states, that dyslexia is "characterised by difficulties in single-word decoding, usually reflecting insufficient phonological processing abilities." Some researchers go still further, and suggest that the vast majority of dyslexic children have a phonological problem (e.g. Snowling, 2000). Phonological awareness acts as a catalyst to the development of decoding skills in English (Byrne, 1998), and therefore a child whose phonological skills are poorly developed will inevitably encounter problems when reading. This was found to be the case in a study by Dixon and Stuart (2002). They taught ten, six-letter words, all in block capitals, to 46 reception children, eleven of whom had poor phonological skills, e.g. they did not know that the first sound in the spoken word 'mat' is 'mmm'. Once all the words were learned, the children were shown eight different spellings for each word, only one of which was actually correct, and asked which one was correct, and if any of the other variations were also correct. Those with good phonological skills picked an average of only three options, including the correct one, e.g. 'SANDAL', 'SADNAL', 'SARDAL'. Those with poorer phonological skills found it very hard to distinguish between the eight variations. The findings of this study imply that without strong phonological awareness skills children are forced to rely on salient visual features, which are absent when using block capitals. The findings would also imply that phonological awareness has two separate inputs to word reading, the ability to decode a word and be able to read it, and more accurate and easier storage of words.

Additional definitions of dyslexia, however, suggest that other difficulties are manifest in dyslexia, not just phonological difficulties, and all symptoms should be treated with equal importance in its diagnosis and treatment. For example, the British Dyslexia Association states, ''Dyslexia is a specific learning difficulty, best described as a combination of abilities and difficulties that affect the learning process in one or more of reading, spelling, and writing. Accompanying weaknesses may be identified in areas of speed of processing, short-term memory, sequencing and organisation, auditory and/or visual perception, spoken language and motor skills…'' (Crisfield, 2002, p. 67).

Many researchers have investigated the spelling abilities of dyslexic readers; indeed, difficulty with spelling is a common symptom of developmental dyslexia (Snowling, 2000). In normal spelling development, children rely heavily on phonology, quickly learning to use letter sounds, although at first they omit parts of words that are difficult to segment (Treiman, 1993). Alongside this, Treiman states that children are also sensitive to orthographic conventions from a very early stage, e.g. learning that 'ck' is never found at the beginning of a word, or that 'ed' endings are spelt the same even if they sound different (e.g. 'liked', 'jumped'). Bradley and Bryant (1983) conducted a longitudinal study following 400 children from age 4 to 8 years. They found a strong relationship between children's phonological awareness at age four, and their reading and spelling ability at age eight, but not their mathematical performance, even when IQ, memory, and socio-economic status were controlled. Snowling (2000) therefore states that children who have lowered phonological awareness ability will be disadvantaged, and therefore the acquisition of orthographic knowledge will be compromised. Support for this view comes from Miles (1983), who states that developmental dyslexics often have more difficulty in spelling than they do in reading, with even well-compensated dyslexics still showing persistent spelling errors as adults. There are two possible explanations for this, either it is because more time is being spent on reading than on spelling instruction and practice, or because spelling problems are more resistant to remediation than reading problems. Dyslexics can learn to spell words, however, this process is usually gradual, and is on a whole-word basis (Ellis, 2001). Their non-word spelling ability is equally as poor as their non-word reading ability, which suggests that any underlying phonological problems hinder the learning of spelling-sound correspondence rules. This suggestion is supported by the types of spelling errors made by dyslexics, which some researchers say are rarely phonetic, but usually contain several of the correct letters even if the spelling is irregular, e.g. ‘dorgher’ for ‘daughter’ or ‘chorce’ for ‘chorus’ (Temple & Marshall, 1983).

In addition to spelling difficulties, researchers have often reported short-term memory deficits in dyslexic readers, finding that despite having normal memory span for visual information, dyslexics often remember fewer verbal items than expected for their age (Snowling, 2000). Hulme, Newton, Cowan, Stuart, and Brown (1999) found that spoken stimuli could be retained in short-term memory for approximately four seconds in a recall task, with poor readers holding spoken material for even less time. It has also been found that dyslexics typically perform poorly on memory tests, such as the Digit Span task (Snowling, 2000). One possible interpretation of these findings is that dyslexic readers have impaired representations of the phonological forms of words, and this phonological coding impairment restricts the number of items that they can retain in memory, which therefore has an effect on working-memory tasks.

A study by Hulme, Maughan and Brown (1991) found that non-dyslexic adults showed a close relationship between speech rate and memory span for both words and non-words, with a slower speech rate linked to less recall. In addition, a study by McDougall, Hulme, Ellis and Monk (1994) found that poor readers tend to have slower speech rates than normal readers. Speech rate and phoneme deletion ability, however, showed independent predictive relationships with reading skill, which indicates that speech rate (a good predictor of memory span) does not overlap completely with phonological awareness ability.

A study by Swanson and Berninger (1995) found that poor readers’ working memory (WM) performance on visual-spatial measures was at a similar level to chronological age matched (CA) controls, but that WM performance on verbal measures was inferior to CA controls. They also found that poor readers' performances on visual-spatial and verbal WM measures were superior to reading comprehension-matched controls, with and without controlling for reading recognition scores. Those poor readers who had difficulty with both word recognition and comprehension had combined WM and short-term memory (STM) deficits. Poor readers with difficulty in comprehension only showed average phonological STM performance but low WM, and children with poor word-recognition skills showed the opposite pattern. These results seem to imply that less skilled readers have impaired WM, which leads to comprehension problems, independent of their phonological coding skills.

Researchers have shown that dyslexics typically have deficits at the level of underlying phonological representations (e.g. Bradley & Byrant, 1978; Perin, 1983; Bruck, 1990). Stanovich and Siegel (1994) found that poor phonology is related to poor reading performance, irrespective of IQ. In support of this, research has shown that many of the phonological deficits in dyslexia persist into adulthood (e.g. Hulme & Snowling, 1997). Further to this, a study of undergraduate students with developmental dyslexia found that the number of words they spelt correctly was lower than other undergraduates. The phonological skills and the word reading accuracy of dyslexic students were also lower than that of other undergraduates and that of controls taken from a wide range of socio-economic backgrounds (Hanley, 1997). In addition, the reading vocabulary of dyslexics was found to be lower than would have been predicted from their scores on picture naming tasks. Contrary to these findings, however, research by Daryn (2000) found that adult dyslexics who make it to university had managed to compensate for their phonological deficits.

This chapter will aim to investigate and describe the reading, spelling and phonological skills, and the visual and verbal memory skills of an adult dyslexic. It is predicted, in accordance with the prevailing phonological deficit theory of reading disorders, that he will show signs of uncompensated or compensated phonological deficits, and a verbal memory impairment.

Case Study, DB

DB is a 22 year old post-graduate who intends be a commercial researcher. He was diagnosed as suffering from developmental dyslexia when he was 8 years of age, but because the family moved often, he did not receive consistent remedial teaching. DB was given a number of standardised tests of language and intellectual abilities. On the Wechsler Adult Intelligence Scales WAIS Block Design test, DB obtained a score of 115, and on the WAIS Vocabulary test, he obtained a verbal score of 115. This compares with his Wechsler Intelligence Scales for Children (WISC) scores as a child of 6 years, where he scored high on verbal ability, scoring 130, and well above average in performance IQ, with a score of 121.

Word Reading:

On the *Wide Range Achievement Test of Reading* (*WRAT-III*; blue form, Jastak & Wilkinson, 1993) a test of single word reading, DB achieved a standardised score of 99. DB also obtained a standardised score of 89 on the WRAT spelling test. Twenty male students from the University of Hull (mean age = 19.9 years, SD = 0.97) averaged a standardised score of 110.25 (SD = 7.14) on the WRAT reading test, and 108.1 (SD = 4.24) on the WRAT spelling test.

To assess reading speed, DB was given the *Test of Word Reading Efficiency* (*TOWRE*, Torgesen, Wagner, Rashotte, Rose, Lindamood, & Conway, 1999). On this test, he had to read a set of familiar words, arranged in a list on the page, as quickly as possible in a time limit of 45 seconds. His sight-word efficiency standard score for this test was 84. These scores suggest that DB has below average reading fluency, and is well below that expected for someone with his verbal IQ.

Spelling:

On an essay writing test, DB made spelling errors on 7.5% of the words in the passage. He wrote 475 words in 30 minutes (15.83 words per minute), and made 36 spelling errors. Eight undergraduate male controls, all from the University of Hull Psychology department, completed this task for course credits. Spelling errors were made, on average, for 2.86% (SD=2.62) of words written. They averaged 356.25 words (SD=99.35) in 26.64 minutes (SD=5.52), resulting in a mean writing speed of 13.37 words per minute (SD=3.08). Therefore, DB makes over double the number of spelling errors in free writing than controls, but his speed is normal. This, coupled with his low score on the WRAT spelling test, suggests that DB has below average spelling ability.

Non-Word Reading:

To assess the speed of pseudo-word reading, DB completed the Phonemic Decoding Efficiency subtest of the *TOWRE*. He was asked to read a set of pseudo-words, arranged on a list on the page, as quickly as possible in a time limit of 45 seconds. DB obtained a standardised score of 89. These tests show DB to be in the low-average range for non-word reading, indicating a few problems in pronouncing unfamiliar letter-sound sequences.

Reading Comprehension:

DB was presented with a comprehension test devised by Singleton and Simmons (2001) consisting of a 1,125 word passage about a factory. He was then given 20 multiple-choice questions based on the passage, and 20 minutes in which to answer them. Five of these questions were literal, with the answers being stated explicitly within the text. The other 15 questions required DB to integrate the information from different parts of the text, or to draw conclusions where the necessary information is not stated explicitly. DB was free to read back over the passage while answering the questions. On this test DB obtained a standardised score of 100, as based on the norms of 80 first year psychology students from the University of Hull. This suggests that despite his poor word reading skills (compared to controls) DB had average reading comprehension.

Phonological and Graphophonemic Skills:

DB was presented with a phoneme deletion task, consisting of 21 words, and asked questions such as "What is school without the 's'?". He achieved an accuracy rate of 95.24% on this test. Twenty male University of Hull students (mean age = 19.9 years, SD = 0.97) averaged an accuracy rate of 92.3% (SD = 9.4). DB was then given a graphophonemic awareness test (Scarborough, Ehri, Olson, & Fowler, 1998). This test consisted of 19 printed words and required him to determine which letter or letters in the words correspond to sounds in the word, and to record the number of sounds in the word. An example of this would be:

T H R O U G H - 3. He scored 78.94% on this test. Ten male controls from the University of Hull, aged between 18 and 23 (mean = 19.6, SD = 1.7) completed this task also, averaging 46.84% (SD = 29.61) accuracy.

A further test of his phonemic awareness skills, i.e. phonological distinctiveness, was presented (adapted from Elbro, Nielson, & Petersen, 1994). For this test, DB was asked to read 28 sentences aloud, very slowly and carefully, as if reading to a child, breaking the words into syllables. The experimenter demonstrated by reading the first sentence to him. Responses were scored according to the pronunciation of vowels in certain key words (see Appendix 1). If the correct form of the vowel was used, the participant scored two points. If an incorrect form of the vowel was used, but otherwise the word was pronounced correctly, the participant scored one. A score of zero was given if the participant was unable to break the word down into its individual syllables, and/or if an incorrect pronunciation was given. DB was unable to break his words down into their syllables, e.g. read-er, thereby scoring zero; however, he did manage to score 89.29% accuracy on the pronunciation of the words given. Ten male controls from the University of Hull, aged between 18 and 26 (mean = 20.5, SD = 2.46) also completed this task. All ten controls could easily break words down into syllables, and with a maximum score of 56, they scored an average of 53.3 (SD = 3.37). They averaged 95.18% (SD = 6.02) in accuracy of reading the items aloud. Thus DB had no problems in reading the words aloud, but he could not break them down into syllables. The impression given was that he could not slow down his pronunciation of the words to do this.

Visual Search Skills:

DB was set a visual search task (Mesulam, 1985), to find as many incidents of one target symbol (max. 60) as he could amongst 314 other symbols. He was also asked to find as many incidents of one target letter (max. 60) as he could amongst 314 other letters. After he had circled ten targets, he was given a different coloured pencil to track his search pattern. Eleven male undergraduate controls, aged between 18 and 22 years old (M = 19.09, SD = 1.14), also took part in this study. All controls were from the University of Hull Psychology department and took part for course credits.

There were two different versions of this test, a random version, with the symbols/letters in a random fashion across the page (all upright and correctly orientated), and a lines version, with the symbols/letters arranged in horizontal lines across the page. DB first completed one random version for letters, followed by a random version for symbols. His accuracy was 100% on both tests, but for the letter search, his search pattern was highly erratic. His search pattern for the symbols was less erratic, but still not ordered.

After this preliminary testing, DB was tested a second time to explore further his search patterns in greater depth. This time, however, the two additional lines versions of the test were completed. Controls were tested to compare their search pattern with that of DB's to see if DB was using a unique search pattern. The presentation order was counter-balanced across the controls. DB's accuracy for both the symbols and letters version of the lines task was at ceiling, higher than the majority of controls. Z-scores show that DB is slower than controls on both of the letters version of this task (random and lines), but faster than controls on both of the symbols versions (random and lines). DB shows slightly higher accuracy levels on both of the letters versions of this task, and higher accuracy on both of the lines versions of this task. (See Table 1 for mean accuracy levels, standard deviations and z-scores for DB's second testing; see Table 2 for mean reaction times, standard deviations and z-scores for DB's second testing).

Table 1:

Normal readers, and DB's second testing, accuracy levels for the visual search task.

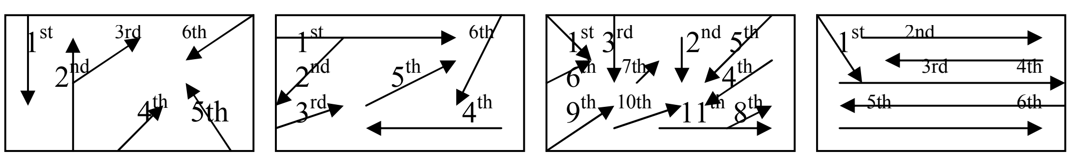
	Random Letters	Random Symbols	Lines Letters	Lines Symbols
Accuracy Level: DB	98.33%	98.33%	100%	100%
Accuracy Level: Controls	98.09% (0.02%)	99.64% (0.01%)	96% (0.05%)	98.73% (0.02%)
DB's Z-Score	-0.04	-1.66	0.78	0.71
% of control's scores below DB's	50%	5.48%	78.81%	78.81%

Table 2:

Normal readers, and DB's second testing, reaction times for the visual search task.

	Random Letters	Random Symbols	Lines Letters	Lines Symbols
Reaction Times (MS): DB	123.00	79.00	109.00	69.00
Reaction Times(MS): Controls	115.09 (SD 35.172)	89.64 (SD 15.81)	96.45 (SD 11.05)	87.64 (SD 12.192)
DB's Z-Score	0.22	-0.63	1.03	-1.33
% of controls faster than DB	57.93%	27.43%	84.13%	8.08%

In the repeated second testing**,** on the random tests for both shapes and letters, DB's search pattern was still erratic, however, on the second testing with the lines versions of shapes and letters, DB's search pattern for shapes was much more ordered, however, his search pattern for letters remained chaotic (see Figure1).



a)Random letters b)Random symbols c)Lines Letters d)Lines Symbols

Figure 1: *DB's visual search patterns*

The control group's search patterns were also analysed. Two of the controls used an erratic search pattern throughout, similar to DB's random symbols search pattern (shown in Figure 1, above). Five of the controls used a top to bottom, bottom to top, zigzag approach, or a top to bottom, top to bottom approach throughout, and the remaining four controls used a left to right, right to left, zigzag approach, or a left to right, left to right approach. Not one of the controls showed a search pattern similar to DB's lines version of the letters task (see Figure 1 (c)).

Visual Reading Abilities

DB was given an 'e' deletion task to test his visual reading abilities. This task consisted of a short passage about the highlights of Hull city (see Appendix 2). DB's task was to read through the passage, and to cross out every 'e' he came across. DB showed an accuracy rate of 87%, with a possible tendency to ignore certain letters when reading, specifically 'e' in 'the'. Ten male controls from the University of Hull (mean age = 19.5 years, SD = 1.27) gained an average accuracy rate of 84.98%, and also showed a tendency to ignore the 'e' in 'the', although this was less pronounced than DB (See Table 3 for mean accuracy levels).

Table 3:

DB's and normal readers' accuracy levels for the 'e' deletion task

	DB	Controls	
E' Deletion Scores	Mean	Mean	SD
Overall Accuracy Level	87.00	84.98	12.22
Percentage of 'e' missing from left of a word	10.53	12.36	10.12
Percentage of 'e' missing from right of a word	89.47	87.65	10.12
Percentage of 'e' missed that were stressed	73.68	58.84	9.92
Percentage of 'e' missed that were unstressed	5.26	13.54	11.02
Percentage of 'e' missed that were silent	21.05	27.62	10.77

Visual Skills:

DB was presented with the embedded figures test (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962) where his average reaction time for correct responses was 28 seconds. This is well within the average range when compared to his peers (based on the norms of 150 college males, Witkin *et al.*, 1971) who averaged 54.3 seconds (SD = 36.8) indicating that his visuo-spatial skills are normal.

Visual and Verbal Memory:

DB completed the Doors and Peoples test of visual skills (Baddeley, Emslie, & Nimmo-Smith, 1994). The test is comprised of four subcomponents: visual recognition, visual recall, verbal recognition, and verbal recall. Presentation order was as follows: verbal recall, nonverbal recognition, delayed verbal recall, nonverbal recall, verbal recognition and delayed nonverbal recall.

Non-Verbal Visual recognition (the doors test):

For nonverbal recognition, coloured photographs of doors were shown to DB. The test was split into two parts (A and B). In each, 12 'target' doors were shown individually to DB at 3 second intervals, followed by the same targets presented in a 2 x 2 array with three other distractor items, all of which fitted the same general label (e.g. church door). DB was asked to pick out the one, out of the four, that had been shown before. Set B was harder than A. One mark was given for each correct response and scaled scores derived from the combined set A and B score. DB achieved a score at the 17th percentile, indicating low-average ability.

Nonverbal Visual Recall (the shapes test):

Four simple drawings were shown individually to DB; each for 5 seconds. DB was then required to draw each of them immediately from memory. Following a delay DB was requested to draw the shapes again from memory. All four comprised a readily drawn version of a cross, varying in overall shape (square or elongated), and in the characteristic features of the crux and the end of the four arms. As with the 'doors', they had an obvious but unhelpful verbal label, e.g. a 'cross'. The drawings were scored according to criteria described by Baddeley *et al.* (1994). DB achieved a total score at the 17th percentile, indicating low-average ability.

Verbal recognition (the names test):

Names were used as the stimulus material, with each item comprising a forename and a surname (e.g. Diane Neeson). As with the nonverbal recognition test there were two subtests (A and B). In set A, 12 female names were shown to DB at 3 second intervals. Then each name was shown in a 2 x 2 array, together with three distractors, and DB was asked to recognise the items presented before. The three distractor items always had the same forename as the experimental item. Set B used male names and was harder than set A. As with the nonverbal recognition test, one mark was given for each correct response and scaled scores were obtained from the combined set A and B score. DB achieved a score in the 17th percentile indicating low-average ability.

Verbal recall (the person test):

This required DB to learn the full names of four characters. Each name comprised a forename and a surname (e.g. Cuthbert Cattermole) and was presented with a photograph, representing the person. The name and photograph were linked by telling DB the occupation of the person (e.g. "This is the minister. His name is Cuthbert Cattermole"). All four names were presented for three seconds followed by the recall stage in which DB was cued to recall the name, given the occupation (e.g. "What was the minister's name'?"). There were three trials to learn the names, unless DB recalled all the names on a particular trial, at which point the test was terminated. Subsequently, delayed recall of the names was tested. One mark was given for each forename and surname correctly recalled, plus an extra mark for a correct pairing. Scores from the three trials were combined and used to derive a scaled score. If the test was terminated early, as indicated above, maximum marks were given for the remaining trials. DB achieved a score in the 95th percentile indicating above average ability.

WAIS Digit Span:

Due to such a large difference in abilities between verbal recall and the other sub-tests of the Doors and People test, DB was presented with the WAIS Digit Span test (Wechsler, 1997), where he had to repeat forward and backward a series of auditorily presented digits. He obtained a standardised score of 90. This shows DB to be in the low-average range for verbal memory, much lower than expected for his verbal IQ.

Summary

The reading, spelling, phonological skills, and the visual and verbal memory skills of DB, a 22 year old post-graduate student who was diagnosed as suffering from developmental dyslexia when he was eight years of age, were evaluated. DB's word recognition was average for the general population (standard score 99) but significantly below average for someone of his educational background and verbal IQ. DB's phoneme awareness skills were close to average for his age. His reading comprehension skills were average compared to other university students. His visual skills were well within the average range, although he was slow on visual search tasks involving letters, and his search pattern for a random array of letters was erratic. His visual and verbal memory skills, however, were below average.

Overall, he does not present the pattern of performance of a phonological dyslexic, as although his non-word reading was mildly impaired he showed little sign of a phonemic awareness deficit. He was found to have a subtle problem in being unable to break words into syllables when reading aloud sentences. This, however, may reflect his short term memory deficit, shown by his low-average performance on the digit span task.

CHAPTER THREE

An In-Depth Examination of DB's Phonological Reading Skills: A Case-Study

A widely used model of reading is the dual-route model (Coltheart, 1978; In Underwood, 1978). This model states that words are read in one of two ways, either by phonological recoding, or by sight. Phonological recoding refers to the process of translating letters into sounds by the application of grapheme to phoneme conversion, and then recognising the identities of words from their pronunciation. Sight-word reading refers to the process of establishing direct connections between the visual form of the printed words and their meanings in memory, as a result of much practice reading the words. A frequently used test of this model is the regularity task, where participants read regular and irregular words. This task generally shows that even adults read regular words more accurately and quickly than irregular words (Seidenberg, Waters, Barnes & Tanenhaus, 1984). It is deduced from this that regular words are at an advantage because as well as being readable by the visual route, information from the phonological route can also assist word recognition. It has been suggested that adults and children differ on the effects of spelling-sound correspondences, with adults showing a regularity effect on low frequency exception words and strange words, and children showing regularity effects with high frequency exception words as well (Waters, Seidenberg & Bruck, 1984). Irregular words generate errors in spellings and pronunciations, especially with exception words, e.g. 'have', and strange words, e.g. 'aisle'. Often these errors are regularisations e.g. pronouncing 'have' as 'gave'.

A direct test of the phonological route is non-word reading. If a reader performs well on a non-word task it can be said that they have good phonological recoding skills, and therefore are capable of processing words via this route. If, however, a reader were to perform less well on a non-word task, they could be relying more on the sight-word route. Examining dyslexics on their ability to read phonologically, via regularity and non-word tasks, is standard practice, and as such forms the experiments in the current study.

For Experiment 1, the regularity task, it is hypothesised that the non-dyslexic controls will show a robust regularity effect, with responses being fastest and most accurate to regular words then to exception words, and responses being slowest and least accurate to strange words. It is also hypothesised that normal readers will have faster and more accurate responses to high frequency words than to low frequency words. As DB shows only mild phonological awareness problems (see Chapter Two), it is hypothesised that he will also show a robust regularity effect in the same way, and that he too will be faster and more accurate to high frequency words, compared to low frequency words. Experiment 2, the non-word reading task, was carried out in addition to the TOWRE non-word reading task of phonemic decoding efficiency (see Chapter Two), to confirm that DB's phonological deficits are only mild, an unusual trait for a dyslexic. It also recorded individual response times per item, rather than just giving an overall response time over a list. It is hypothesised that DB will show similar performance levels to the non-dyslexic controls on this task.

Experiment 1:

Regular words consist of ‘gave’, ‘this’, ‘cat’, etc. and have regular spelling-sound correspondences, i.e. they are pronounced in the same way that they are spelt. Irregular words may be ‘exception’ words, i.e. ‘have’, or ‘watch’, which are spelled similarly to regular words, but pronounced differently, or ‘strange’ words, i.e. ‘aisle’, or ‘eyes’, which have individual spellings, non-comparable to any others in the English spelling system, and which do not follow spelling-sound correspondences. It is hypothesised that the controls will show a significant regularity effect and they will have faster and more accurate responses to high frequency words than to low frequency words (as detailed in the introduction). As DB shows only mild phonological problems, it is hypothesised that he will also show a robust regularity effect in the same way, and that he too will be faster and more accurate to high frequency words, compared to low frequency words.

Participants:

DB, a male dyslexic case-study, and 20 male students from the University of Hull, (mean age = 19.9 years, SD = 0.97) took part in this experiment**.** All completed the *Wide Range Achievement Test of Reading* (*WRAT-III*; blue form, Jastak & Wilkinson, 1993), a test of single word reading, and the control group averaged a standardised score of 110.25 (SD = 7.14). DB achieved a standardised score of 99.

Procedure:

One hundred and twenty unrelated mono-syllabic words of similar length (see appendix 3) (taken from Kuĉera and Francis (1967), cited in MRC Psycholinguistic Database, 2004) were presented to participants. Twenty were high frequency regular words, 20 were high frequency exception words, 20 were high frequency strange words, 20 were low frequency regular words, 20 were low frequency exception words and 20 were low frequency strange words (see Table 4 for mean word length and frequencies). A two-way ANOVA was carried out on the MRC word frequencies, with word type (regular and irregular) and frequency (high or low) as the between subjects factors.

There was no main effect of word type (F(1,38) = <1), but there was a main effect of frequency ($F(1, 38) = 61.76$, $p<0.001$), showing that the matching had been effective.

The computer used was a Dell laptop with a 60hz refresh rate. The screen was 14 inches, with a resolution of 1024x768. The program used was MemWord version 1.4. Each word appeared individually in the centre of the screen. DB was asked to pronounce each word as quickly and as accurately as possible. At the onset of his verbal response to the word, the item disappeared from the screen. Accuracy was recorded, as was the actual pronunciation given, and the reaction times for each correct response. Twenty controls also took part, following the same procedure.

Table 4:

Mean word-length frequencies for stimuli used in the regularity task

	High Frequency			Low Frequency		
	Regular	Strange	Exception	Regular	Strange	Exception
Mean Frequency	356.7	251.4	357.8	8.3	5.5	8.6
SD	334.5	299.0	523.9	7.1	6.3	6.9
Mean Length	4.3	4.5	4.5	4.1	4.6	4.15
SD	0.7	0.8	0.8	0.6	0.8	0.5

Results:

As DB is a case-study, all data for both DB and controls were analysed across-stimuli (i.e. by items as opposed to by subject).

Reaction Times:

DB:

A factorial ANOVA was carried out on DB's mean reaction time data, with word type (regular, strange and exception) and frequency (high or low) as the main factors. There was no main effect of frequency, ($F(1, 19) = 1.724$, $p>0.05$). There was a main effect of word type ($F(2,38) = 3.271$, $p<0.05$). Newman-Keuls post-hoc tests showed that regular words were responded to significantly slower than exception words ($p<0.05$). There was also a significant interaction between word type and word frequency ($F(2, 38) = 3.655$, $p<0.05$). Newman-Keuls post-hoc tests showed that amongst high frequency words, regular words were responded to significantly slower than strange words ($p<0.01$) and exception words ($p<0.01$). There were no significant differences between word types in low frequency words. (See Table 5 for means and standard deviations; see Figure 2 for mean reaction times for word type and frequency, with standard deviation bars).

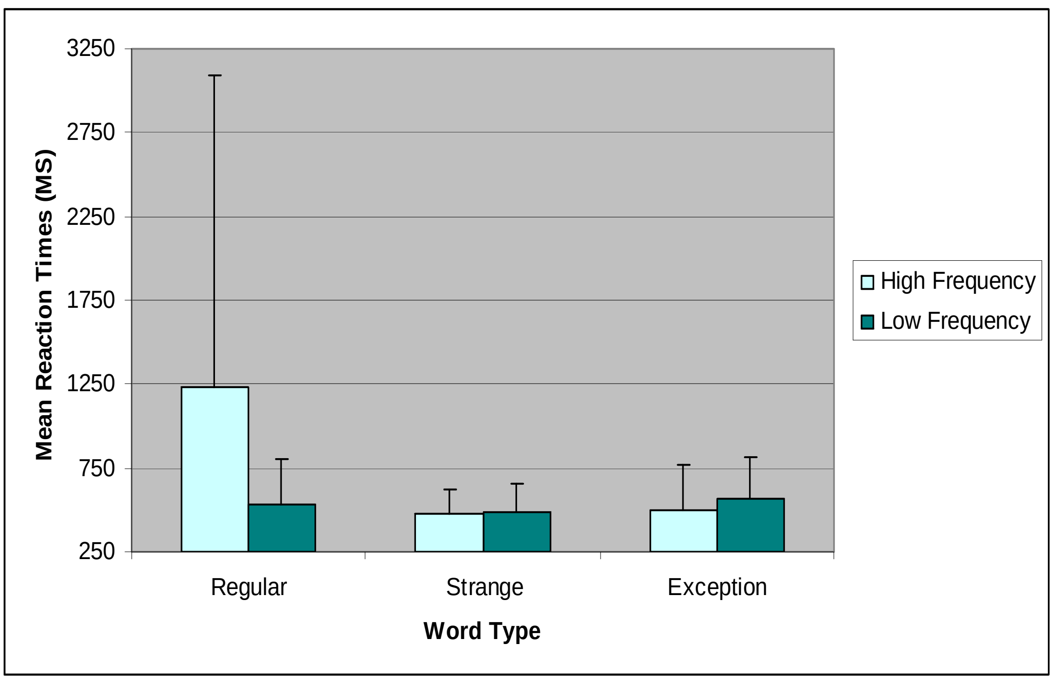


Figure 2: *Mean reaction times (ms) for word type and word frequency for DB*

As DB had large standard deviations for high frequency regular words, this warranted removal of two anomalous reaction time data, both of which were above 6000ms (the words he struggled with were ‘smile’ and ‘life’). This data-removal was not done initially, as, being a case-study, the data-cleansing process may have hidden some of the true patterns of DB's reading. A new factorial ANOVA was carried out on DB's cleansed mean reaction time data, with word type (regular, strange and exception) and frequency (high or low) as the main factors. There was no main effect of frequency, ($F(1, 17) = >1$). There was no main effect of word type ($F(2,34) = 1.41$, $p>0.05$). There was also no significant interaction between word type and word frequency ($F(2, 34) = 1.90$, $p>0.05$). (See Table 5 for means and standard deviations).

Controls:

A factorial ANOVA was carried out on the control group's mean reaction time data, with word type (regular, strange and exception) and frequency (high or low) as the main factors. There was a significant main effect of frequency, $F(1, 19) = 41.966$, $p<0.001$. This showed that high frequency words were responded to significantly faster than low frequency words. There was a significant main effect of word type, $F(2, 38) = 17.531$, $p<0.001$. Newman-Keuls post-hoc tests showed that the control group responded significantly faster to regular words than exception words ($p<0.01$) and strange words ($p<0.01$). There was a significant interaction between word type and frequency, $F(2, 38) = 26.679$, $p<0.001$. Newman-Keuls post-hoc tests showed that response times to high frequency regular words were significantly faster than to high frequency strange words ($p<0.05$) and high frequency exception words ($p<0.01$). Response times to low frequency regular words were significantly higher than to low frequency exception words ($p<0.01$) and low frequency strange words ($p<0.01$). Response times to low frequency exception words were significantly faster than to low frequency strange words ($p<0.01$). No other interactions were significant. (See Table 5 for means and standard deviations; see Figure 3 for mean reaction times by word type and frequency with standard deviation bars).

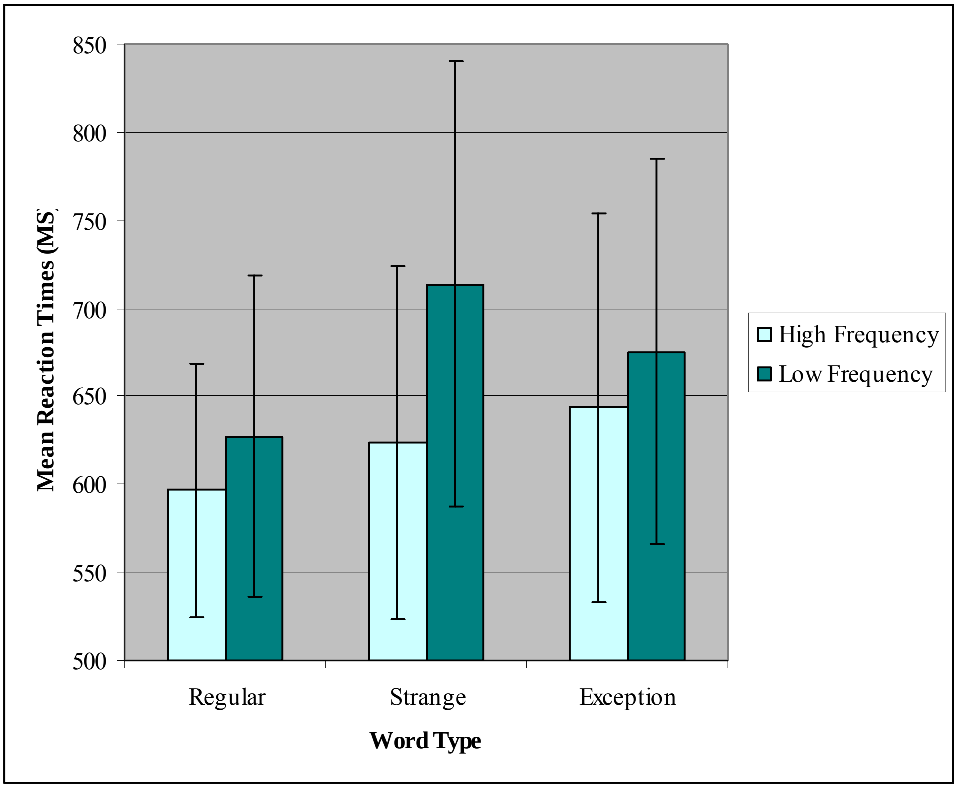


Figure 3: *Mean Reaction Times for Word Type and Word Frequency for Controls*

Table 5:

Mean reaction times for word type for DB (cleansed data in brackets) and controls

		High Frequency			Low Frequency		
		Regular	Strange	Exception	Regular	Strange	Exception
DB	Mean RT	1231.05 (660.97)	471.25	500.65	533.63	490.12	569.45
	SD	1864.57 (512.95)	151.36	269.48	272.07	166.34	249.28
Controls	Mean RT	596.65	623.95	643.55	627.50	713.93	675.38
	SD	72.47	100.09	110.69	90.89	126.00	109.48

Accuracy:

DB:

A factorial ANOVA was carried out on DB's mean accuracy data, with word type (regular, strange and exception) and frequency (high or low) as the main factors. There was no significant main effect of frequency, $F(1, 19) = 1$, $p>0.05$, and there was no main effect of word type, $F(2, 38) = 0.487$, $p>0.05$. There was a significant interaction between word type and word frequency on accuracy, $F(2, 38) = 3.353$, $p<0.05$.

Newman-Keuls post-hoc tests failed to show where these differences reached significance. (See Table 6 for means and standard deviations; see Figure 4 for mean accuracy rating for word type and word frequency).

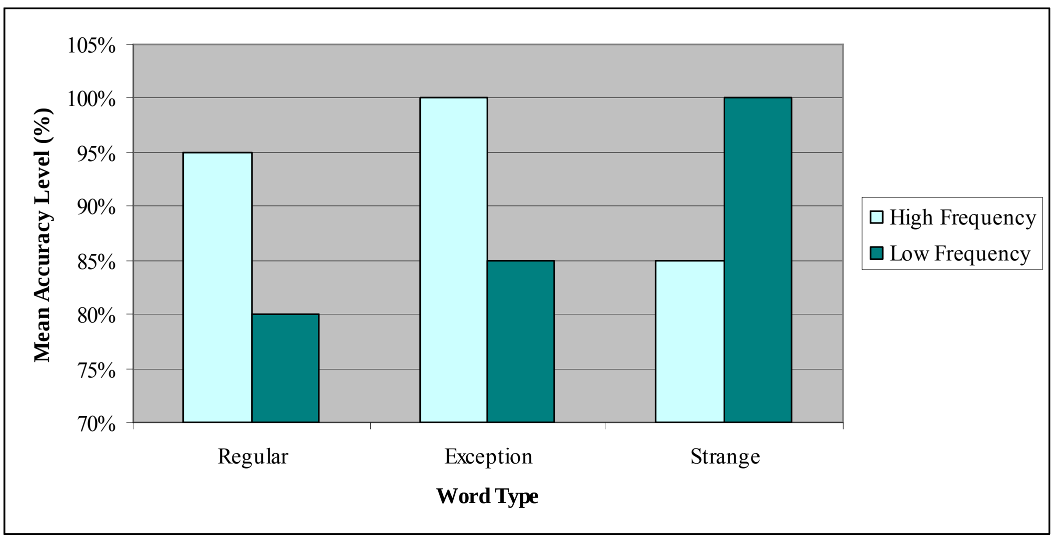


Figure 4: *Mean accuracy rating for word type and word frequency for DB*

Controls:

A factorial ANOVA was carried out on the control's accuracy data, with word type (regular, strange and exception) and frequency (high and low) as the main factors. There was a main effect of frequency, $F(1, 19) = 6.565$, $p<0.05$. High frequency words were responded to significantly more accurately than low frequency words. There was a main effect of word type, $F(2, 38) = 26.679$, $p<0.001$. Newman-Keuls post-hoc tests showed that regular words were responded to significantly more accurately than exception words ($p<0.01$) and strange words ($p<0.01$). Exception words were responded to significantly more accurately than strange words ($p<0.05$). There was a significant interaction between frequency and word type, $F(2, 38) = 13.403$, $p<0.001$. Newman-Keuls post-hoc tests showed that high frequency regular words were responded to significantly more accurately than high frequency strange words, ($p<0.01$) and high frequency exception words were responded to more accurately than high frequency strange words ($p<0.01$). Low frequency regular words were responded to more accurately than low frequency strange words ($p<0.01$), and low frequency exception words ($p<0.01$). No other interactions were significant. (See Table 6 for means and standard deviations; see Figure 5 for mean accuracy rating for word type and word frequency with standard deviation bars).

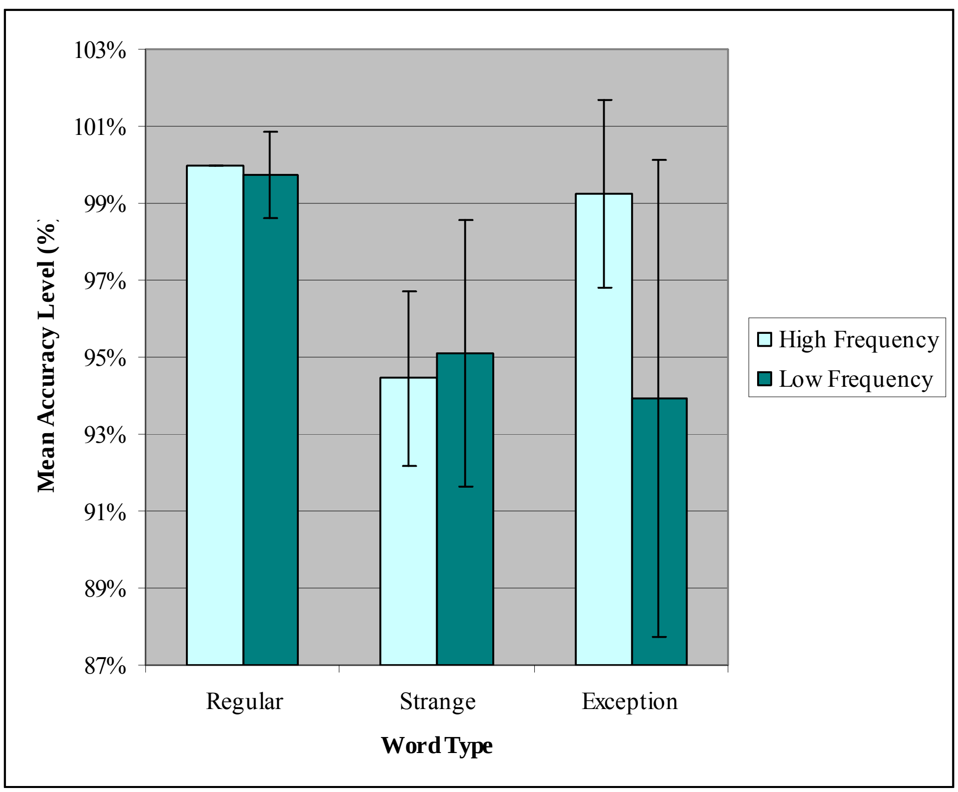


Figure 5: *Mean accuracy rating for word type and word frequency for controls*

Table 6:

Mean accuracy rating for word type for DB and controls

		High Frequency			Low Frequency		
	Accuracy	Regular	Strange	Exception	Regular	Strange	Exception
DB	Mean	95%	100%	85%	80%	85%	100%
Controls	Mean	100%	94%	99%	100%	95%	94%
	SD	0%	2%	2%	1%	3%	6%

DB's reaction times were then compared to the reaction times of normal readers. Z-scores showed that DB's reaction times were faster than normal readers' reaction times for all word types, with the exception of high frequency regular words, where he was much slower than controls. It must be noted, however, that after his two anomalous reaction times were removed, his reaction times to high frequency regular words were similar to controls. DB's accuracy scores were much lower than the accuracy scores of controls for both high and low frequency regular words, with less than 0.47% of controls scoring lower than DB. DB also had lower accuracy scores for high frequency strange, and low frequency exception words, with less than 10% of controls scoring below DB for these word types. DB reached the ceiling level of accuracy for high frequency strange words, with 38% of controls scoring less than this. He also reached ceiling level for low frequency exception words, with over 90% of controls scoring below DB in this word category. (See Table 7, for z-scores).

Table 7:

Z-scores comparing DB' reaction times (cleansed data in brackets) to controls

	High Frequency			Low Frequency		
	Regular	Strange	Exception	Regular	Strange	Exception
DB's Accuracy Z-Score	-4.36	-0.3	-2.98	-4.23	-1.34	1.31
% of control's scores below DB	<0.47%	38.27%	<0.47%	<0.47%	9.8%	90.21%
DB's RT Z-Score	3.89 (0.11)	-1.44	-1.16	-0.98	-1.55	-1.09
% of controls faster than DB	>99.53% (54.38%)	8.08%	11.51%	15.87%	6.78%	13.69%

DB made eleven reading errors in total, accounting for only 9.17% of the word stimuli. His reading errors (see Table 8 for error analysis; see Appendix 4 for full account of responses) often appear to occur at the ends of words, i.e. the right-hand side of words. Five of his errors were visual errors of this kind, for example 'best' was read as 'bes' by DB, likewise 'comb' was read as 'come' and 'flock' was read as 'flow'. Three of his errors were visual errors made in the middle of the word, for example 'maths' was read as 'mass', and 'steak' was read as 'streek'. One word, 'corps' was read as 'corpse', which suggests a grapheme-phoneme impairment, such that DB had regularised the word's pronunciation. Two words, both high frequency exception words, resulted in no response being given, and it was noted that DB struggled with these words, seemingly unable to read them at all. The words that DB did not pronounce correctly were used as a basis for a multiple-choice questionnaire to check

DB understood their meanings. DB knew the meanings to all of the words, except for one - ‘corps’ (see Appendix 5 for questions).

Table 8:

DB’s word reading errors according to word type

Word Type	Word	Response	Error Type
High Frequency Regular	best	Bes	Visual error at end of word
High Frequency Exception	corps	Corpse	Grapheme-phoneme impairment
High Frequency Exception	eye	None	No response
High Frequency Exception	two	None	No response
Low Frequency Regular	flock	Flo	Visual error at end of word
Low Frequency Regular	lump	Lum	Visual error at end of word
Low Frequency Regular	maths	Mass	Visual error in middle of word
Low Frequency Regular	plank	Plan	Visual error at end of word
Low Frequency Strange	comb	Com	Visual error at end of word
Low Frequency Strange	fete	Feet	Visual error in middle of word
Low Frequency Strange	steak	Streek	Visual error in middle of word

Discussion:

The controls showed a very robust regularity effect, with significantly more accurate responses to regular words, then to exception words, and then to strange words, as was predicted. For words of both high and low frequency, regular words were responded to more accurately than either strange or exception words. The controls also showed significantly faster and more accurate responses to high frequency words than low frequency words, as hypothesised. DB’s results were unexpected. Despite having only mild phonological problems, he did not show any effect of word type in his accuracy scores, and his reaction times were in fact faster to strange words, then to exception words, and slowest of all to regular words, with high frequency regular words having the slowest responses of all. This changed somewhat once his anomalous high frequency reaction time data were removed, after which there was no effect of word type, or frequency, and no interaction was found. On two of the high frequency regular words 'smile' and 'life' DB took over 6000ms to respond, seemingly struggling to find suitable pronunciations. It was this data that was removed in the data cleansing. Even after this data was removed from the calculations, DBs reaction times to high frequency regular words remained longer then to any other word type.

His accuracy scores did show an interaction between word type and frequency but further analysis failed to reveal the particulars. His response times did not show a reliable effect of frequency, and although his accuracy levels were higher to words of high frequency, and lower to words of low frequency, this trend failed to reach significance. DB often omitted the endings of words, which appears to be an unusual trait.

DB's response times tended to be faster than that of controls, and perhaps this shows a speed-accuracy trade-off, resulting in these surprising findings. This, however, is not supported by his responses to high frequency regular words, where his accuracy levels are low despite slow response times. DB's accuracy scores were lower than those of controls for all word types, except for high frequency strange, and low frequency exception words, where DB reached ceiling level.

It is possible that DB's phonological difficulties, though only mild, may have interfered with his ability to perform this task. An additional non-word task is needed to further examine DB's phonological reading skills, to ensure that his TOWRE score of pseudo-word reading (on which DB achieved a standardised score of 89) is an accurate reflection of his ability to phonologically decode written text. If his phonological difficulties are only mild, and his performance on a non-word task is similar to that of controls, this will be indicative of an unusual word reading strategy being used by DB, and hints at the possibility of additional underlying problems in his reading. Experiment 2, therefore, is a non-word reading task.

Experiment 2:

According to the dual-route model of reading, words are read in one of two ways, either by phonological recoding, or by sight. The most direct way to assess a reader's ability to phonologically recode text is by using non-words, as they have not encountered them before. As DB appears to have only mild phonological difficulties, it is hypothesised that he will show similar performance levels to the non-dyslexic controls both in terms of speed of response and accuracy levels.

Participants:

The same participants took part as in Experiment 1.

Procedure:

Twenty non-words were devised and were presented in the same order to each participant. The computer used was a Dell laptop with a 60hz refresh rate. The screen was 14 inches, with a resolution of 1024x768. The program used was MemWord version 1.4. Each non-word appeared individually in the centre of the screen. DB was asked to pronounce each non-word as quickly and as accurately as possible. When he started to verbally respond to the non-word, the item disappeared from the screen. Accuracy was recorded, as was the actual pronunciation given, and the reaction times for each correct response.

Results:

Descriptive statistics showed that DB had a 65% accuracy rating for non-words, and a mean reaction time of 1774.14 ms (see Table 9 for means, z-scores and standard deviations). The controls had a higher accuracy rating of 82% for non-word reading, and a slower mean reaction time of 1919.33 ms. Z-scores showed that DB was only just below average in terms of accuracy, with 42% of controls scoring lower than he did. DB was also only slightly faster than average, with 46% of controls responding faster than he did. (See Table 9 for means, z-scores and standard deviations).

Table 9:

Reaction times, accuracy levels and z-scores for non-words for DB and controls

		Accuracy	Reaction Times (MS)
DB	Mean	65.00%	1778.38
	SD		747.94
Controls	Mean	81.50%	1919.33
	SD	13.70%	967.15
DB Z-Score		-0.26	-0.08
% of Controls scores below DB/ % of Controls faster than DB		42.07%	46.04%

DB made seven errors in total. His four reading errors (see Table 10 for error analysis; see Appendix 6 for full account of responses) appear to be different to the type of errors made in the regularity reading task in Experiment 1. These errors were visual errors near the beginning or middle of the stimuli. In all cases, DB correctly pronounced the endings to all non-words shown. DB gave no response to three of the non-words presented, seemingly unable to attempt a pronunciation for such novel stimuli.

Table 10:

DB's non-word reading errors

Non-Word	Response	Error Type	% of Stimuli Present in Response
go am strak	groam strak	Visual error at beginning of word	100%
sul grim toab	slur grim too ab	Visual error at beginning of word	100%
mun tel klin	mun tek lin	Visual error in middle of word	90%
brem tad lum	brim tu lum	Visual error in middle of word	70%
fair krum dup	None	No response	0%
jig lum tem	None	No response	0%
swa bla nap	None	No response	0%

Discussion:

DB's accuracy for this task was slightly lower than that of the controls. His reaction times, however, were slightly faster than those of controls, so there is possible evidence of a speed-accuracy trade-off, although this appears unlikely. These results were as expected as DB shows mild phonological problems only.

The reading errors made were either towards the beginning or the middle of the stimuli. This is in contrast to his errors in Experiment 1, where he often omitted word endings completely. It seems likely that the use of novel stimuli forced DB to process all the letters sequentially, whereas more familiar words may not. In earlier tests (see Chapter Two) DB scored in the low-average range on several tests of visual and verbal memory, therefore it is possible that the reading errors made here were due to losses from short-term memory, not an impairment in letter scanning.

General Discussion:

On the regularity reading task the controls' pattern of results were as expected, as they displayed a robust regularity effect in terms of response speed and accuracy. DB, however, did not show a regularity effect, nor did he show any effect of word type in his accuracy scores. His responses were less accurate than controls on almost all words, with the exception of high frequency strange, and low frequency exception, words.

His reaction times did not show a recognised pattern, failing even to show an effect of frequency, and being faster to strange words, then to exception words, and slowest of all to regular words, with high frequency regular words having the slowest responses of all. His response times were faster than controls to all words, with the exception of high frequency regular words, where he was much slower than controls. This suggests that DB does not read via a phonological route, but is more likely to rely on a visual route, even for high frequency regular words.

This is confirmed by DB's reading errors, which were mostly visual errors, accounting for 73% of errors made. He was unable to read the words 'eye' and 'two', both of which are high frequency exception words. DB also often missed the endings of words altogether; this occurred five times in total, accounting for 4% of the stimuli read. This appears to be an unusual trait, even for those with reading difficulties. These findings do not appear to be due to an impairment of his phonological ability, as his responses in the non-word reading task were similar to those of controls. His reading errors on the non-word task were only near the beginning or the middle of the stimuli, not at the ends of the non-words. This suggests that when faced with novel stimuli DB is forced to scan and process all the letters, whereas perhaps he does not do this when familiar words are presented. It is possible that DB has compensated for his reading difficulties by scanning only enough of the word to minimise the possible choices, and, in combination with context, to select the correct word without having to scan, process and retain letter order of all letters present.

Further investigation is required to discover what reading strategies DB is employing, as there are currently no documented cases of dyslexics showing this reading pattern. A lexical decision task (LDT) split across the two hemispheres may help to clarify DB's reading performance, as his word reading errors seem to be mostly at the right-hand side of words. This unusual trait could possibly be explained in terms of faulty scanning of words, or by an impairment in the left-hemisphere processing of word stimuli. This issue will be addressed in Chapter Four.

CHAPTER FOUR

Differences in Reading Strategies and Hemispheric Processing in Dyslexics and Non-Dyslexics: A Case-Study

Research suggests that some dyslexics read high frequency words more visually than reading age controls (e.g. Johnston & Morrison, 2007). Johnston and Morrison (2007) examined whether IQ level was associated with differing reading strategies, and it was concluded that low IQ poor readers tend to read via a phonological route, whereas high IQ poor readers (often described as dyslexic) tend to read via a visual route with high frequency words. The present study therefore will also explore more deeply DB's use of orthographic reading, via lexical decision tasks that assess regularity effects and the influence of orthographic neighbourhood.

An orthographic neighbourhood is any word that can be created by changing one letter of the stimulus word, preserving letter positions (Coltheart, Davelaar, Jonasson & Besner, 1977). For example, 'marsh' has two neighbours, 'harsh' and 'march'. In contrast, the word 'cover' has 13 neighbours (including coven, covet, cower, hover, lover, mover and rover) (Lavidor, Hayes, Shilcock & Ellis, 2004).

Lavidor and Ellis (2001) found that lexical decision performance in normal adult readers was affected by orthographic neighbourhood size (N) when stimuli were presented to the RH (or left visual field, LVF), but not when stimuli were presented to the LH (right visual field, RVF).

Although it is now generally accepted that the right, as well as the left, hemisphere contributes to the lexical processing in the normal brain (e.g. Chiarello, Shears, Liu, & Kacinik, 2005), it is less clear whether this is still the case for developmental dyslexics. A case-study carried out by Lavidor, Johnston and Snowling (2006), on FM, a male adult with developmental dyslexia and severe phonological deficits, contrasted performance on words with many orthographic neighbours, with words with few orthographic neighbours (N). It was found that the dyslexic's performance on a lexical decision task (LDT) revealed a greater reliance on RH orthographic processing strategies than normal readers, and poorer LH performance, suggesting less efficient reading processes in the LH. Lavidor, Johnston and Snowling (2006) found facilitation effects in a LDT for words with many-'N' when presented to LVF/RH for normal readers and less severely disabled dyslexics, however, two of their findings were unique to the dyslexics. Firstly, the pattern of faster and more accurate responses to words with many-'N' when compared to words with few-'N' was significantly larger for the dyslexics, and were limited to the LVF/RH. Secondly, orthographic neighbourhood size had an inhibitory effect on RVF/LH words for dyslexics but not for normal readers.

In the split-fovea model, when words are presented centrally, the RH is assumed to reflect 'N' effects generated by the initial, or lead, letters of English words (presented in the LVF), this is tested by the presentation of either high lead (HN) or low lead 'N' items (LN). A high lead 'N' refers to words where the first three letters, e.g. in the word banker, are frequently used at the beginning of words, e.g. banter, banner. A high end 'N' means that the last three letters of a word i.e. manage, are frequently used at the end of words, e.g. forage, garage. Low lead, and low end, words are words that do not begin, or end, with a trio of letters which are used frequently, i.e. carrot, puzzle. The LH would process the end letters (presented in the RVF), these would be either high or low end 'N' items (EN). Words however, are not simple symmetrical stimuli and initial letters of words are possibly more informative than their endings. Orthographic neighbours that share the initial letters with a target (LN) e.g. anger, angel, may facilitate target recognition more than orthographic neighbours that share final letters with a target (EN), e.g. aloud, cloud. It is arguable however, that it is likely that the RH activates LN but that the LH does not necessarily activate EN (Lavidor, Hayes, Shilcock & Ellis, 2004). It is important to note that while both lead N and end N may have an effect on reading performance when words are centrally presented, this is not the case when words are presented to the left or the right of fixation (Lavidor, Hayes, Shilcock & Ellis, 2004).

In this chapter, two lexical decision experiments were carried out to examine DB's use of orthographic reading. Due to his unusual trait of omitting word endings (see Chapter Three), i.e. the right hand-side of words, appearing in the RVF, it is likely that DB has a LH impairment indicating inefficient reading processes in this hemisphere. Experiment 3 was a regularity lexical decision task, and Experiment 4 was an orthographic lexical decision task; both were split across the hemispheres to assess his reading strategies according to hemispheric processing.

In both experiments it is hypothesised that DB will respond more slowly, and make more errors in the RVF/LH, as his reading patterns on a regularity reading task (see Chapter Three) suggest impaired processing in the LH. His pattern of reading on this earlier task, where he showed better performance on irregular words compared to regular words, also suggests that in Experiment 3 DB will show higher accuracy levels, and faster response times, to strange words than to regular or exception words when stimuli are presented to the LVF/RH. It is hypothesised, however, that this pattern of results will be reversed in the RVF/LH, with DB showing a regularity effect. This is partly because research shows dyslexics to have impaired performance in LH processing (e.g. Lavidor, Johnston & Snowling, 2006), suggesting that the LH relies more on a phonological route to reading than a visual route, and partly because DB has only mild phonological difficulties, which should not impede phonological processing of words. It is hypothesised that normal readers will show a regularity effect in terms of both speed of response, and accuracy levels for all word types in all presentation positions (left, centre and right).

For Experiment 4 it is hypothesised that normal readers will show no significant difference in response time or accuracy between words of high-lead 'N' and low-lead 'N' words when they are presented to either the LVF/RH, or the RVF/LH. It is hypothesised, however, that on central presentation, normal readers will show significantly faster response times, and higher accuracy levels to high-lead 'N' words compared to low-lead 'N' words. It is also hypothesised that DB will show greater sensitivity to high-lead 'N' words compared to low-lead 'N' words on central presentation, and no significant difference between high-lead 'N' words and low-lead 'N' words when presented to either LVF/RH or RVF/LH.

Experiment 3:

Research suggests that for reading tasks, dyslexics rely more on right hemisphere processing than left (e.g. Lavidor, Johnston & Snowling, 2006). This has also been shown in regularity lexical decision tasks (LDT) (e.g. Lavidor & Ellis, 2001), which are speeded tasks consisting of regular and irregular words and corresponding non-words, shown individually to a participant, requiring the participant to decide whether the letter string seen was a word or a non-word. This advantage for RH processing in LDTs is possibly due to the activation of many real words, leading to a very fast 'word' response. In pronunciation tasks, however, DB is very accurate with low frequency strange words, where there are few competing possible responses. It is therefore possible to suggest that the visual distinctiveness of these words may be an advantage for him. When words are presented in the central location in a pronunciation task (as in Experiment 1, Chapter Three), DB is very accurate with strange words, possibly because they are visually distinct and the beginning letters are not activating many competing words. It was therefore thought necessary to use a lexical decision task, varying word position, using low frequency regular, exception, and strange words, to follow up the first experiment.

By using a speeded lexical decision task that splits the letter-string across the visual fields, it is expected that DB will make more errors in the right visual field (RVF), the left hemisphere (LH). This is expected partly because most dyslexics seem to have difficulty processing words in the LH (Lavidor, Johnston & Snowling, 2006), but also because DB's previous errors appear to be in the right-hand-side of words, which would fit with this expectation. It is also hypothesised that DB will show higher accuracy levels, and faster response times, to strange words than to regular or exception words when stimuli are presented to the LVF/RH. It is hypothesised, however, that this pattern of results will be reversed in the RVF/LH, with DB showing a regularity effect. It is hypothesised that normal readers will show a regularity effect in terms of both speed of response, and accuracy levels for all word types in all presentation positions (left, centre and right).

Participants:

DB, the dyslexic case-study, and 20 undergraduate male participants (mean age = 26.45, SD = 5.44), who self-reported that they had never been diagnosed with reading difficulties, nor had ever had additional help with their reading during their schooling, took part in this study for course-credit.

Procedure:

DB was presented with a lexical decision task, consisting of 60 words of low frequency, all taken from Experiment 1, 20 of which were regular, 20 were exception, and 20 were strange words (see Appendix 7 for words). The variance of frequency between word groups was not significantly different $F(2,47) = 1.032$, $p>0.05$, meaning that the words were all of similar word-frequency. It must be noted, however, that as these words were the same stimuli as in Experiment 1, orthographic neighbourhood size (N) had not been controlled. Unfortunately, there was a significant difference between the three word types (regular, strange and exception) in terms of overall N ($F(2,57) = 15.25$, $p<0.001$. Newman Keuls post-hoc tests showed that the strange words had smaller overall N (mean = 2.75, SD = 3.24) than the regular words (mean = 9.25, SD = 4.31) ($p<0.05$) and the exception words (mean = 9.1, SD = 5.00) ($p<0.05$).

Sixty non-words were created from these words by changing one or two letters in the words, usually the vowels, to make pronounceable, but non-legal English words, although all digraphs used were legal (see Appendix 8 for non-words). The 120 letter-strings were then presented to DB on a computer screen as a lexical decision task. The individual letter-strings all appeared randomly on the left, the right and in the centre of the screen, 360 (3x120) letter strings being presented. DB was asked to press the "0" for a word and "2" for a non-word on the number pad on the right of a "qwerty" keyboard. The letter-strings all appeared for a period of 150 milliseconds, preceded by a central fixation point, shown for 400 milliseconds. The letter-strings were then followed by a decision screen; where DB was given time to respond word/non-word on the keyboard. This screen lasted for 1800 milliseconds. If a response was not given within this time, it was recorded as an incorrect response.

The letter-strings were presented centrally, to the left of the screen at a visual angle of 2.5º and to the right of the screen at a visual angle of 2.5º. The computer used had a 1 GHz Intel Celeron processor, with 256 MB of RAM, a CTX 16" monitor, with a refresh rate of 100 Hz, and a Logitech 'qwerty' keyboard. DB was sat 57 centimetres away from the screen. Twenty undergraduate male non-dyslexic controls were used for comparison, half using '0' as a word response and '2' as a non-word response, and half with these response keys reversed.

Results:

As DB is a case-study, all data for both DB and controls were analysed across-stimuli.

Reaction Times:

DB:

An across-stimuli ANOVA was carried out on DB's median reaction time data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (regular, strange and exception). Only correct responses were analysed, which resulted in some missing data. This was solved using the median reaction time per condition (i.e. left presentation, regular word) for each missing item. There was no significant main effect of word position F_2 (2, 38) = <1, and there was no significant effect of word type F_2 (2, 38) = 1.02, $p<0.05$. There was no significant interaction, F_2 (4, 76) = <1. (See Table 11 for means and standard deviations, and comparison to controls; see Figure 6 for mean reaction times by word type and position with standard deviation bars).

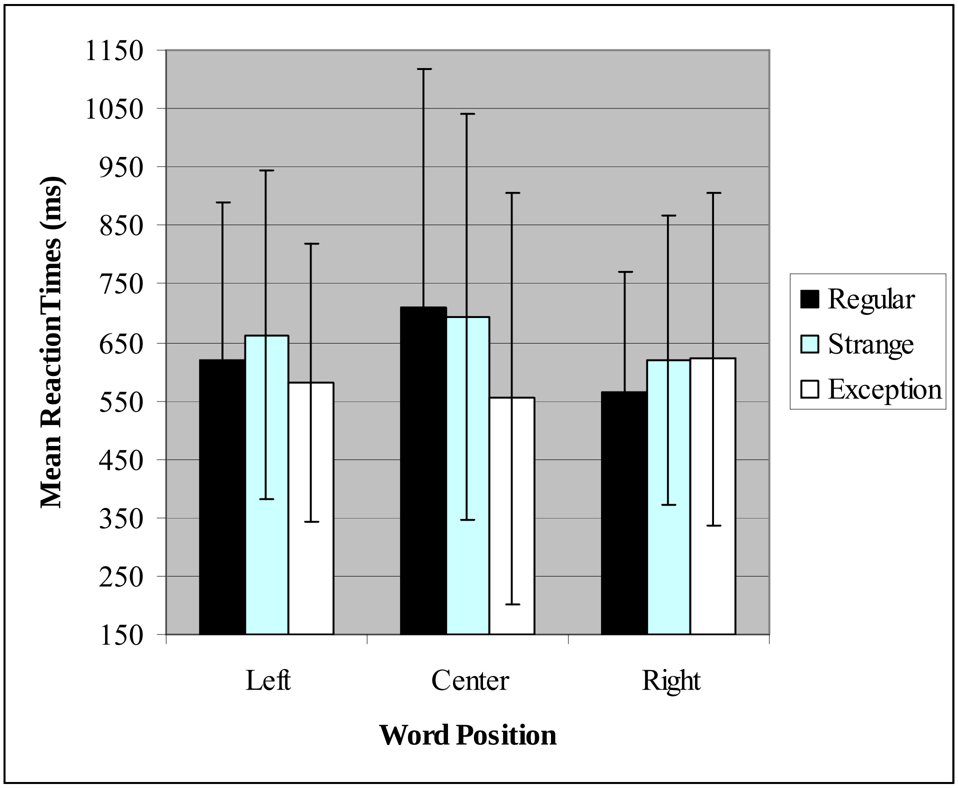


Figure 6: *Mean reaction times for word type and word position for across-stimuli ANOVA for DB*

Controls:

An across-stimuli ANOVA was carried out on the control group's median reaction time data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (regular, strange and exception). Only correct responses were analysed. There was a significant main effect of word position F_2 (2, 38) = 17.36, p<0.001. There was no significant effect of word type, F_2 (2, 38) = <1. There was no significant interaction, F_2 (4, 76) = 2.34, p>0.05.

Newman-Keuls post hoc tests on word position showed that the control group's reaction times were significantly faster to words presented centrally than to words presented to the RVF/LH (p<0.01), and words presented to the LVF/RH (P<0.01). Words presented to the RVF/LH were also responded to significantly faster than words presented to the LVF/RH (p<0.01). (See Table 11 for means and standard deviations, see Figure 7 for mean reaction time by word type and word position with standard deviation bars).

Table 11:

DB's and control group's mean reaction times for word type and word position

		DB		Controls	
		Mean	SD	Mean	SD
Left	Regular	619.10	271.24	491.50	77.35
	Strange	662.55	282.56	533.38	85.21
	Exception	580.55	237.03	514.93	94.55
Centre	Regular	709.65	407.29	426.58	44.07
	Strange	693.65	347.54	454.35	84.92
	Exception	554.00	352.75	443.10	104.75
Right	Regular	566.35	204.65	493.65	93.38
	Strange	619.60	248.23	478.65	62.18
	Exception	621.70	283.79	461.10	50.23

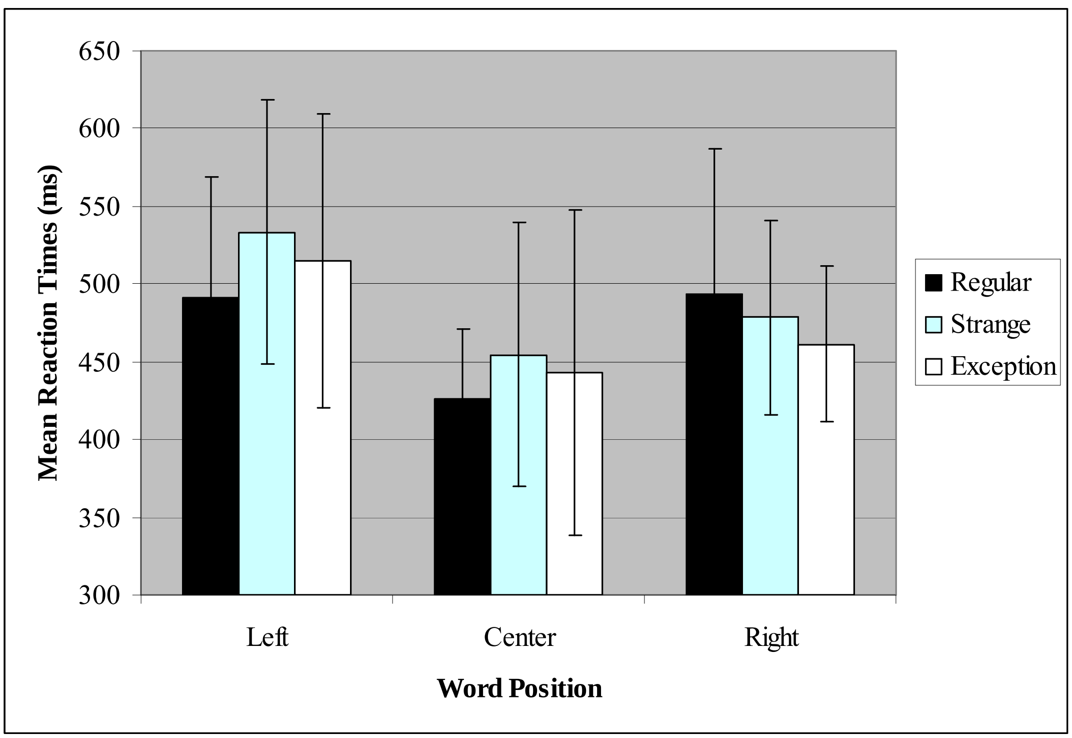


Figure 7: *Mean reaction times for word type and word position for across-stimuli analysis for the control group*

Accuracy:

DB:

A further across-stimuli ANOVA was carried out on DB's mean accuracy data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (regular, strange and exception). There was no significant main effect of word position F_2 (2, 38) = <1, and there was no significant effect of word type F_2 (2, 38) = <1. There was no significant interaction, F_2 (4, 76) = <1. (See Table 12 for means and standard deviations, and comparison to controls; see Figure 8 for mean accuracy level by word type and position).

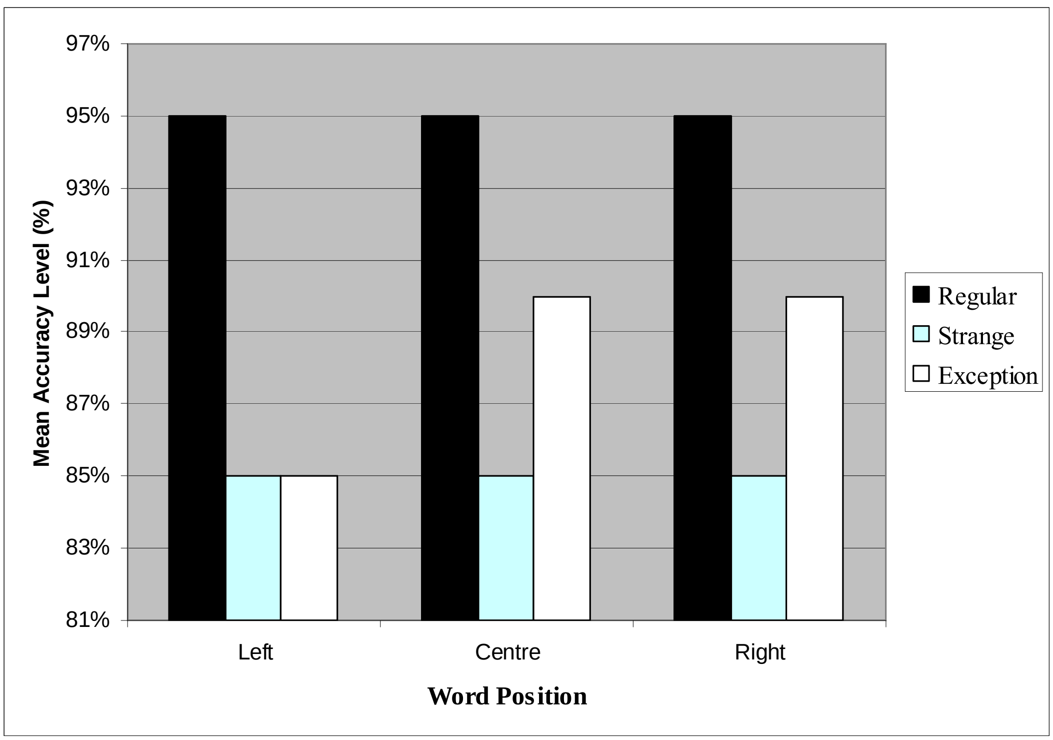


Figure 8: *Mean accuracy level for word type and word position for across-stimuli ANOVA for DB*

Controls:

An across-stimuli ANOVA was carried out on the control group's accuracy data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (regular, strange and exception). There was a significant effect of position $F(2, 38) = 14.536$, $p<0.01$. There was no significant effect of word type, $F(2, 38) = 1.02$, $p>0.05$. There was no significant interaction, $F(4, 76) = 1.23$, $p>0.05$.

Newman-Keuls post-hoc tests showed that words presented centrally were responded to significantly more accurately than words presented to the RVF/LH ($p<0.01$), and words presented to the LVF/RH ($p<0.01$). Words presented to the RVF/LH were responded to significantly more accurately than words presented to the LVF/RH ($p<0.05$).

Table 12:

DB's and control group's mean accuracy levels for word type and word position

		DB	Controls (Across subjects)	
		Mean	Mean	SD
	Regular	95.00%	83.50%	10.65%
	Strange	85.00%	77.75%	15.00%
Left	Exception	85.00%	79.00%	12.73%
	Regular	95.00%	92.00%	6.37%
	Strange	85.00%	80.00%	11.81%
Centre	Exception	90.00%	86.00%	7.00%
	Regular	95.00%	88.75%	10.99%
	Strange	85.00%	77.75%	16.18%
Right	Exception	90.00%	81.00%	10.46%

DB's median reaction times were then compared to the median reaction times of normal readers. DB's reaction times to all word types, in all word positions were slightly slower than normal readers' reaction times (see Table 13 for z-scores).

Table 13:

Z-Scores comparing DB's reaction times to controls

		Regular	Strange	Exception
Left	DB's Z-Score	0.87	0.90	0.57
	% of people faster than DB	81.47%	81.47%	72.57%
Centre	DB's Z-Score	2.17	1.98	0.87
	% of people faster than DB	98.61%	97.72%	81.47%
Right	DB's Z-Score	0.66	1.17	1.02
	% of people faster than DB	75.69%	88.49%	84.13%

Discussion:

DB's reaction times were slightly slower than that of controls for all word types in all positions, and he failed to show an effect of word type, i.e. no regularity effect. There was no significant effect of word position for DB however, which is very surprising, as the literature shows that dyslexics usually have difficulties with the reading processes in the LH (Lavidor, Johnston & Snowling, 2006). DB also showed no significant difficulties with accuracy in the LH. It must be noted, however, that DB had seen the word stimuli before, in Experiment 1, albeit several weeks beforehand, and this may have affected his accuracy and reaction times.

The controls did show an effect of word position, as expected, performing best (faster and more accurately) to words presented centrally, then to words presented to the RVF/LH, and worst (slower and less accurately) to words presented to the LVF/RH. The controls also showed a regularity effect independent of word position, with regular words having the highest accuracy rating, then exception words, then strange words.

It is possible, however, that these results were confounded by the presence of other lexical properties, such as orthographic neighbourhoods, which had not been controlled. This was further examined in Experiment 4.

Experiment 4:

Orthographic Neighbourhood size (N), i.e. the number of words that can be created by changing one letter of the stimulus word, preserving letter positions, can facilitate performance in a LDT for words with many ‘N’ when presented to the LVF/RH (Lavidor, Johnston & Snowling, 2006). The same research also found that dyslexics show greater sensitivity to ‘N’ size than non-dyslexics when presented to the LVF/RH, but that ‘N’ size had an inhibitory effect on the RVF/LH words for dyslexics but not for non-dyslexics.

It is hypothesised in the present study, that, on central presentation, normal readers will show significantly faster response times in making correct responses to high-lead ‘N’ words compared to low-lead ‘N’ words, as demonstrated by Lavidor, Hayes, Shilcock and Ellis (2004). This is because the initial letters of centrally presented words are projected to the RH, and words with high-lead ‘N’ activate not only its own lexical entry but also the entries for other words with similar appearance, thus facilitating the processing of the target word.

In a case study, a dyslexic male, FM, showed greater orthographic sensitivity in the LVF/RH (Lavidor, Johnston & Snowling, 2006), therefore, it is hypothesised that DB will show more accurate and faster response times to high-lead ‘N’ words than low-lead ‘N’ words when items are centrally presented. Their case study, FM, also showed no orthographic sensitivity in the RVF/LH (Lavidor, Johnston & Snowling, 2006), therefore, it is hypothesised that DB will show no significant difference in accuracy or response time between high-lead ‘N’ and low-lead ‘N’ words, when centrally presented, as this would indicate a lack of orthographic sensitivity in the RVF/LH.

As FM showed greater LVF/RH orthographic sensitivity (Lavidor, Johnston & Snowling, 2006), it was argued that this was compensation for his phonological problems. If, however, DB shows RH sensitivity to orthography, and has only mild phonological problems, it will be difficult to argue that his orthographic sensitivity is a compensation for an underlying phonological disorder.

As lead-N and end-N cease to be meaningful when presented to the left or right of fixation, and as overall N size was controlled, it is hypothesised that normal readers will show no difference between high-lead 'N' words and low-lead 'N' words when presented to either the left (LVF/RH) or right (RVF/LH) visual fields (Lavidor, Hayes, Shilcock & Ellis, 2004). It is also hypothesised that DB will show no significant difference between high-lead 'N' words and low-lead 'N' words for items presented to the left visual field (LVF/RH), or the right visual field (RVF/LH).

Participants:

The same participants took part as in Experiment 3.

Procedure:

DB was presented with a lexical decision task, consisting of 96 six-letter words; 32 of which were words with high lead and high end neighbours (e.g. *banker, manage),* 15 were high lead low end neighbour words (e.g. *deceit, impair)*, 21 were low lead high end neighbour words (e.g. *dipper, rotate)*, and the remaining 28 words were low lead low end neighbour words (e.g. *muzzle, freeze)* (taken from Lavidor & Walsh, 2003) (see appendix 9). Due to the difference in n size, the reaction time data used the median average to extend all categories (i.e. left presentation, high lead 'N', low end 'N' words) to 32, and the accuracy data used the mean average. The sets of words did not differ on word frequency $F(3, 60) = 0.228$, $P>0.05$. The words were also matched for overall orthographic neighbourhood size (N), so that no word type (HLHE (mean = 1.88, SD = 1.68), HLLE (mean = 1.60, SD = 1.64), LLHE (mean = 2.24, SD = 1.70), LLLE (mean = 1.43, SD = 1.26)) had a significantly higher or lower overall N size than another ($F(3,92) = 1.174$, $p>0.05$).

Ninety-six non-words were created from these words by changing one or two letters in the words, usually the vowels, to make pronounceable, but non-legal English words, although all digraphs used were legal (see appendix 10). These 192 letter-strings were then presented to DB on a computer screen as a lexical decision task. The individual letter-strings all appeared on the left, the right and in the centre of the screen, randomly, which resulted in 576 (3x192) letter strings being presented. The procedure was as in Experiment 3.

Results:

As DB is a case-study, all data for both DB and controls were analysed across-stimuli.

Reaction Times:

DB:

An across-stimuli ANOVA was carried out on DB's median reaction time data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (high lead 'N' high end 'N' words; high lead 'N' low end 'N' words; low lead 'N' , high end 'N' words; low lead 'N' low end 'N' words). Only correct responses were analysed, which resulted in some missing data. This was solved using the median reaction time per condition (e.g. left presentation, high lead 'N', high end 'N') for each missing item. There was no significant main effect of word position, F_2 (2, 62) = <1. There was a significant effect of word type, F_2 (3, 93) = 3.866, $p<0.05$. Newman Keuls post-hoc tests revealed that high lead 'N' low end 'N' words were responded to significantly faster than low lead 'N' low end 'N' words ($p<0.01$) and significantly faster than high lead 'N' high end 'N' words ($p<0.05$). There was a significant interaction, F_2 (6,186) = 3.45, $p<0.01$. Newman-Keuls post-hoc tests showed that on central presentation, high lead 'N' low end 'N' words were responded to significantly faster than low lead 'N' high end 'N' words ($p<0.05$), high lead 'N' high end 'N' words ($p<0.05$), and low lead 'N' low end 'N' words ($p<0.01$). When presented to the LVF/RH, DB responded significantly faster to high lead 'N' low end 'N' words than to low lead 'N' high end 'N' words ($p<0.05$), low lead 'N' low end 'N' words ($p<0.01$) and high lead 'N' high end 'N' words ($p<0.01$). (See Figure 9 for mean reaction time by word type and position with standard deviation bars; see Table 14 for means and standard deviations).

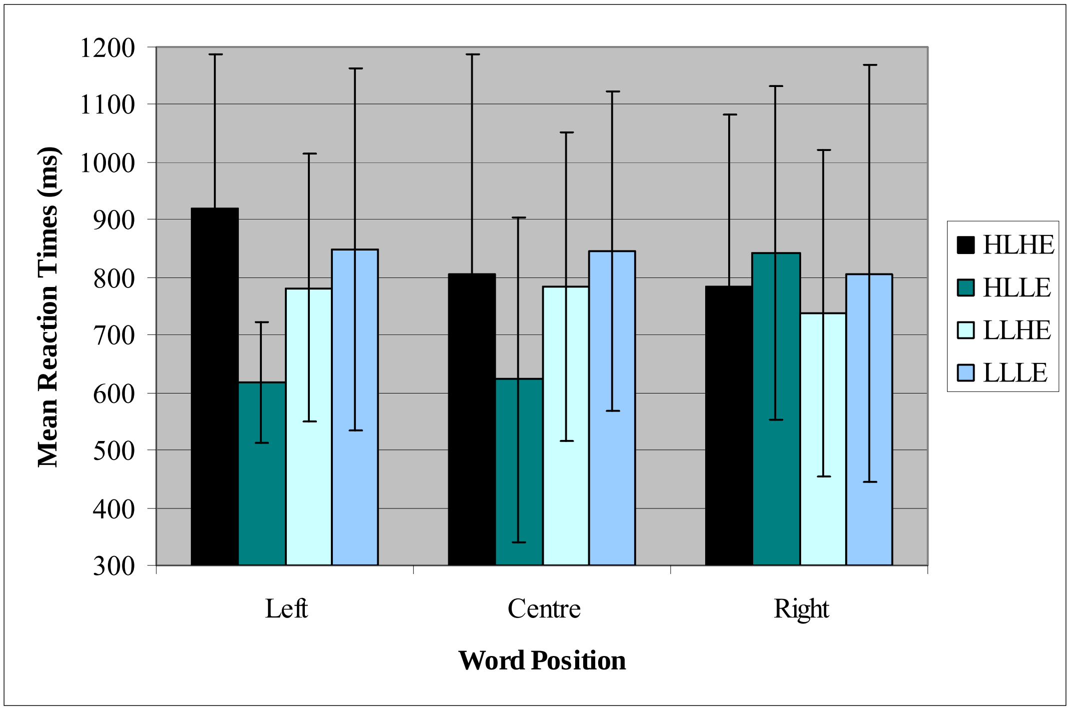


Figure 9: *Mean reaction times for word type and word position for DB*

Table 14:

DB' s and control group's mean reaction times for word type and word position

		DB		Controls	
		Mean	SD	Mean	SD
Left	HLHE	919.00	267.68	568.422	77.863
	HLLE	617.03	105.76	542.563	45.0947
	LLHE	781.31	233.02	522.594	51.435
	LLLE	848.78	314.19	566.547	120.168
Centre	HLHE	805.47	382.00	459.297	96.6509
	HLLE	623.03	281.50	417.953	39.3256
	LLHE	784.47	267.77	431.875	46.5134
	LLLE	844.91	277.07	443.781	59.5049
Right	HLHE	784.63	297.16	520.141	93.5602
	HLLE	843.09	289.41	479.078	47.8624
	LLHE	737.28	284.37	471.641	50.5087
	LLLE	806.44	361.50	534.719	152.498

Controls:

An across-stimuli ANOVA was carried out on the control group's median reaction time data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (high lead 'N' high end 'N' words; high lead 'N' low end 'N' words; low lead 'N' high end 'N' words; low lead 'N' low end 'N' words). Only correct responses were analysed. There was a significant main effect of word position, F_2 (2, 62) = 131.269, $p<0.001$. There was a significant effect of word type, F_2 (3, 93) = 4.2, p>0.01. There was no significant interaction, F_2 (6,186) = 1.442, p>0.05.

Newman-Keuls post-hoc tests showed that words presented centrally were responded to significantly faster than words presented to the RVF/LH (p<0.01), and words presented to the LVF/RH (p<0.01). Words presented to the RVF/LH were responded to significantly faster than word presented to the LVF/RH (P<0.01). Low lead 'N' high end 'N' words, were responded to significantly faster than low lead 'N' low end 'N' words (p<0.05), and high lead 'N' high end 'N' words (p<0.05). High lead 'N' , low end 'N' words were responded to significantly faster than low lead 'N' low end 'N' words (p<0.05), and high lead 'N' high end 'N' words (p<0.05). (See Figure 10 for median reaction time by word type and position with standard deviation bars; see Table 14, above, for means and standard deviations, and comparisons with DB).

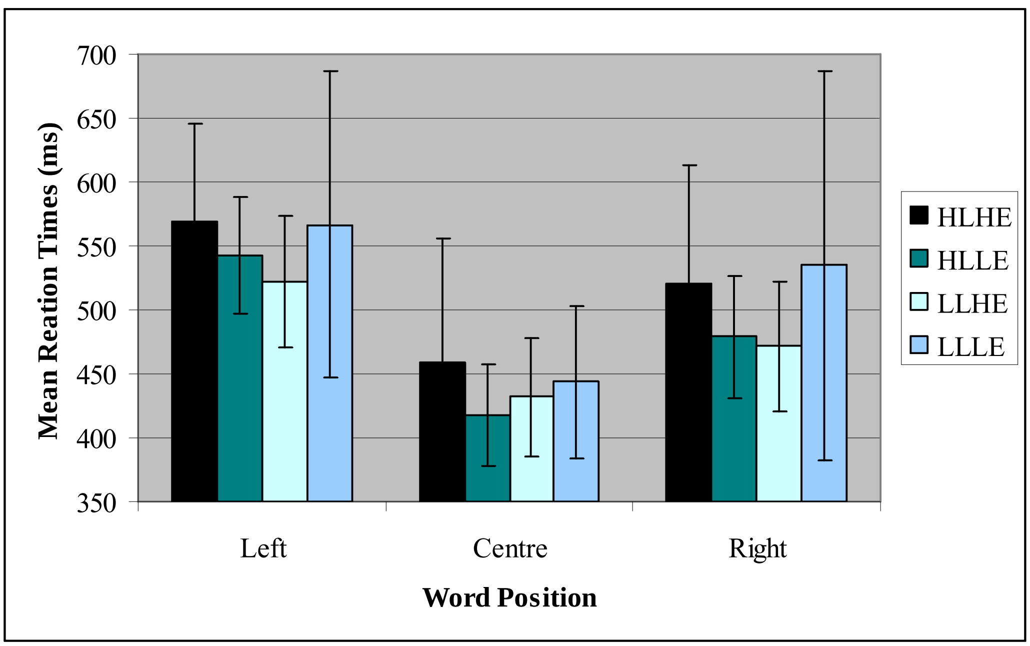


Figure 10: *Mean reaction times for word type and word position on across-stimuli analysis for the control group*

Accuracy:

DB:

A further across-stimuli ANOVA was carried out on DB's accuracy data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (high lead 'N' high end 'N' words; high lead 'N' low end 'N' words; low lead 'N' , high end 'N' words; low lead 'N' low end 'N' words). There was no significant main effect of word position F_2 (2, 62) = <1, and there was no significant effect of word type F_2 (3, 93) = <1. There was no significant interaction, F_2 (6, 186) = 1.77, p>0.05. (See Figure 11 for mean accuracy level by word type and position; see Table 15 for means and standard deviations, and comparison to controls).

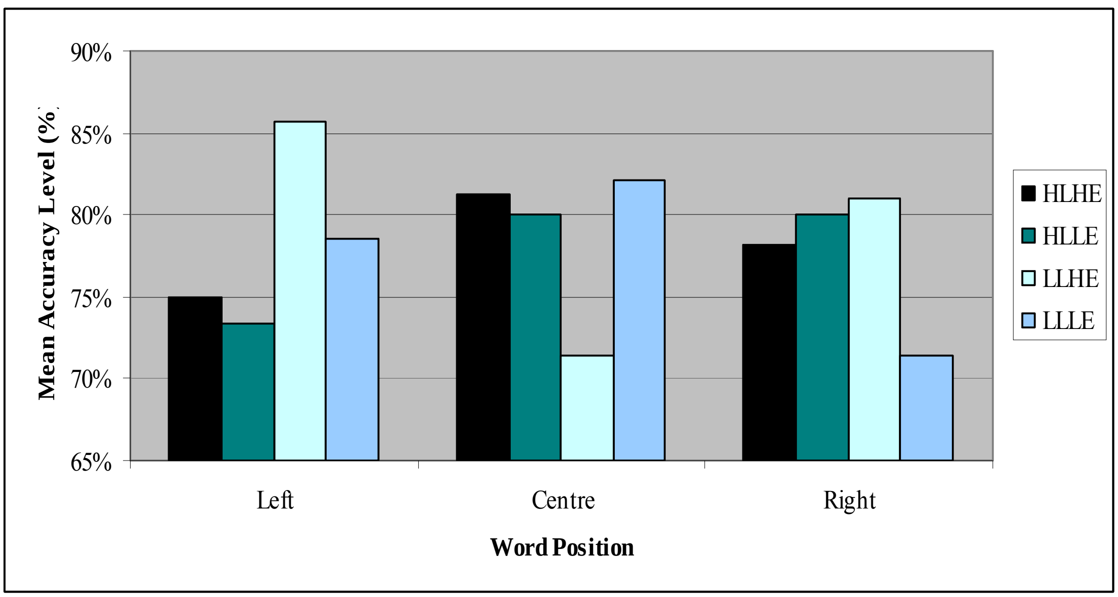


Figure 11: *Mean accuracy level for word type and word position on across-stimuli analysis for DB*

Table 15:

DB's and control group's mean accuracy levels for word type and word position

		DB	Controls (Across Subjects)	
		Mean	Mean	SD
Left	HLHE	75.00%	76.41%	16.46%
	HLLE	73.33%	76.33%	15.67%
	LLHE	85.71%	80.00%	13.78%
	LLLE	78.57%	70.36%	14.89%
Centre	HLHE	81.25%	87.66%	12.88%
	HLLE	80.00%	88.33%	11.42%
	LLHE	71.43%	87.62%	09.45%
	LLLE	82.14%	79.82%	12.97%
Right	HLHE	78.13%	84.38%	14.83%
	HLLE	80.00%	82.67%	12.12%
	LLHE	80.95%	83.10%	10.98%
	LLLE	71.43%	78.21%	13.56%

Controls:

A further across-stimuli ANOVA was carried out on the control group's accuracy data, with individual words compared across positions. The main factors used were word position (left, centre and right) and word type (high lead 'N' high end 'N' words; high lead 'N' low end 'N' words; low lead 'N' , high end 'N' words; low lead 'N' low end 'N' words). There was a significant main effect of word position, F_2 (2, 28) = 29.33, $p<0.001$. There was no significant effect of word type, F_2 (3, 42) = 1.59, $p>0.05$. There was no significant interaction, F_2 (6, 84) = 1.63, $p>0.05$.

Newman-Keuls post hoc tests showed that words presented centrally were responded to significantly more accurately than words presented to the RVF/LH ($p<0.01$), and words presented to the LVF/RH ($p<0.01$). Words presented to the RVF/LH were responded to significantly more accurately than words presented to the LVF/RH ($P<0.01$). (See Figure 12 for mean accuracy level by word type and position with standard deviation bars; see Table 15, above, for means and standard deviations, and comparisons with DB).

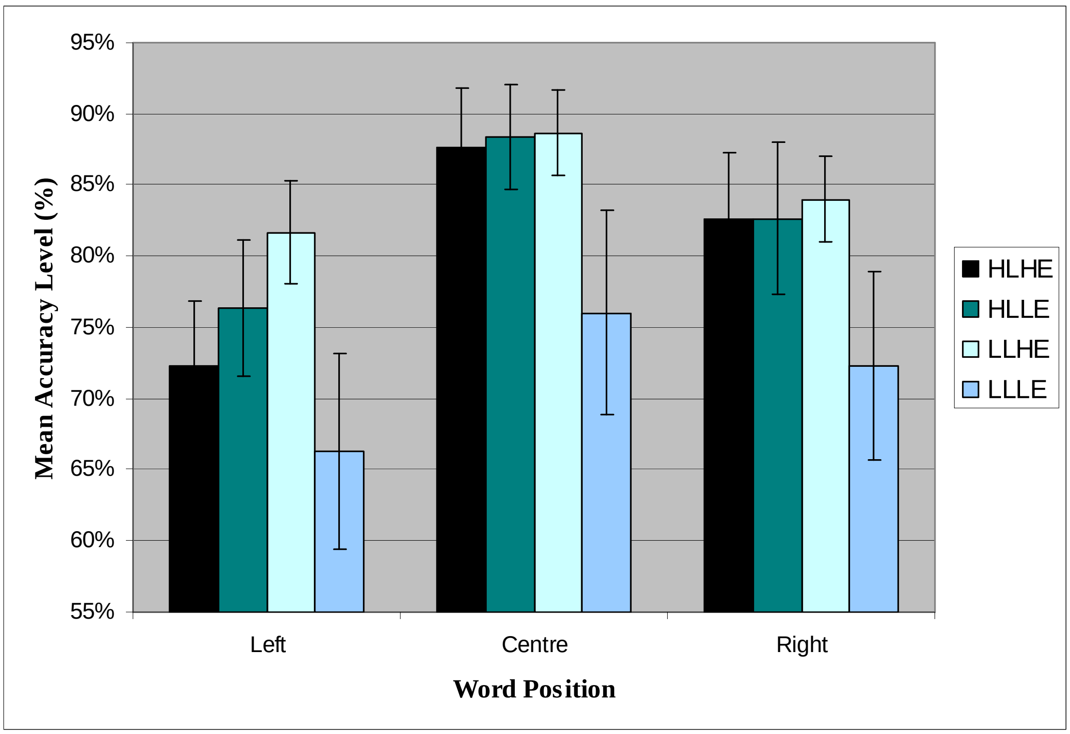


Figure 12: *Mean accuracy level for word type and word position on across-stimuli analysis for the control group*

DB's median reaction times were then compared to the median reaction times of normal readers. DB's reaction times to all word types, in all word positions were slower than normal reader's reaction times (see Table 16 for z-scores).

Table 16:

Z-Scores comparing DB's reaction times to controls

		High Lead High End	High Lead Low End	Low Lead High End	Low Lead Low End
Left	DB's Z-Score	2.08	0.23	2.19	2.37
	% of people faster than DB	98.17%	57.93%	98.61%	99.18%
Centre	DB's Z-Score	2.23	1.33	1.99	2.72
	% of people faster than DB	98.61%	85.85%	97.72%	99.53%
Right	DB's Z-Score	1.73	2.38	1.46	2.03
	% of people faster than DB	92.87%	99.18%	93.22%	97.72%

Discussion:

The controls showed a significant and reliable effect of word position, with words presented centrally being responded to significantly faster, and more accurately, than words presented to the RVF/LH, and words presented to the LVF/RH responded to slowest and least accurately of all. This could be due to the LH's dominance for lexical processing.

Overall the controls responded fastest to low lead 'N' high end 'N' words, then to high lead 'N' low end 'N' words, then to low lead 'N' low end 'N' words, and slowest to high lead 'N' high end 'N' words. Overall, DB responded fastest to high lead 'N' low end 'N' words, then to low lead 'N' high end 'N' words, then to low lead 'N' low end 'N' words, and slowest to high lead 'N' high end 'N' words.

When presented centrally, the controls were fastest to respond to low lead 'N' low end 'N' words, then to high lead 'N' low end 'N' words, then to low lead 'N' high end 'N' words and slowest to high lead 'N' high end 'N' words. This is very surprising, and is the opposite of what was hypothesised. When presented to the LVF/RH, the controls responded fastest to low lead ‘N’ high end ‘N’ words, then to low lead ‘N’ low end ‘N’ words, then to high lead ‘N’ high end ‘N’ words, and slowest to high lead ‘N’ low end ‘N’ words. When presented to the RVF/LH, controls responded fastest to high lead ‘N’ low end ‘N’ words, then to low lead ‘N’ high end ‘N’ words, then to low lead ‘N’ low end ‘N’ words and slowest to high lead ‘N’ high end ‘N’ words. Again, these results were unexpected, as it was predicted that there would be no effect of word type for words presented to either the left or the right of fixation. These results, however, could be due to the high lead ‘N’ high end ‘N’ words activating too many other lexical entries, thus slowing performance rather than facilitating it.

When presented centrally, DB responded fastest to high lead ‘N’ low end ‘N’ words, then to low lead ‘N’ high end ‘N’ words, then to high lead ‘N’ high end ‘N’ words, and slowest to low lead ‘N’ low end ‘N’ words. These results were unexpected, as it was predicted that high lead N words would facilitate performance over low lead N words. When presented to the LVF/RH, DB was fastest to respond to high lead ‘N’ low end ‘N’ words, then to low lead ‘N’ high end ‘N’ words, then to low lead ‘N’ low end ‘N’ words, and slowest to high lead ‘N’ high end ‘N’ words. When presented to the RVF/LH, DB was fastest to respond to low lead ‘N’, high end ‘N’, then to high lead ‘N’ high end ‘N’ words, then to low lead ‘N’ low end ‘N’ words, and slowest to respond to high lead ‘N’ low end ‘N’ words. Again, this is contrary to what was expected, as no effect of word type was predicted for words presented to either side of the fixation point.

DB’s performance on this task is highly unusual for a dyslexic. He shows no over-reliance on the RH for reading processes, and yet he does not have the same pattern of results as the non dyslexic controls. This suggests that his compensatory behaviour for the difficulties he faces with reading have not led to solitary usage of visual cues, but have incorporated the phonological strategies used by the majority of non-dyslexics.

General Discussion of Chapters Two, Three and Four:

DB, a 22-year-old male with developmental dyslexia, has undergone a battery of tests to try to determine the main differences between his reading strategies and non-dyslexics. DB was shown to have average word recognition skills, average phonological decoding skills, average comprehension levels and visual skills for his age group although reading and spelling was below what would be expected for his IQ. He is therefore a well compensated dyslexic. He was, however, below average for visual and verbal memory skills, and reading fluency.

DB is a particularly interesting case as there is currently a great deal of controversy surrounding the role of phonological difficulties in dyslexia (e.g. Nicholson, 2005). The four experiments that were reported in chapters 2, 3 & 4, aimed to discover what reading strategies DB was using. According to the Committee of the International Dyslexia Association (Ryan, 1994; Shaywitz, 1996), “Dyslexia is … a specific language-based disorder of constitutional origin characterised by difficulties in single word decoding”. The literature also suggests that the typical developmental dyslexic has difficulties with phonological decoding and thus relies heavily on visual processing (Frith, 1985). As DB does not appear to have much difficulty with phonological processing, this raises the possibility that he might use the same reading strategies as non-dyslexics. This question was examined in Experiment 1, the regularity pronunciation task.

In Experiment 1, DB’s results revealed a very different pattern of results to that of controls. The non-dyslexics showed a robust regularity effect, finding regular words easier to pronounce than exception words, and having the most difficulty with strange words. Their performance was also facilitated by high word frequency. DB, however, was faster to respond to strange words than exception words, and he had the most difficulty with regular words. His performance was also not facilitated by high word frequency. After removing two anomalous reaction times, however, DB showed no effect of word type, or frequency, and no interaction. DB's pronunciation errors also tended to be at the end of words, often omitting them. It was thought possible that DB was perhaps being inhibited on this task either due to a speed-accuracy trade-off, or due to his phonological reading difficulties.

Experiment 2, the non-word pronunciation task, was used to discover if DB's phonological difficulties were more severe than first thought, or if they could be ruled out as an explanation for his performance on Experiment 1. In Experiment 2, DB's results were as expected. His mild phonological problems led to lower accuracy but he did show a slightly faster response time when compared with the controls. This, however, was not a large enough discrepancy to undermine the original findings from the TOWRE.

DB, however, seems to be using different reading strategies to non-dyslexics, as he does not show the same regularity effects. This, however, cannot be explained in terms of severe phonological difficulties. It was thought that a lexical decision task that split the words across the visual fields, and therefore across the hemispheres, may help to clarify the underlying differences between DB's reading strategy and that of non-dyslexics. This led to Experiment 3, the regularity lexical decision task (LDT). DB's results were once again different to the non-dyslexics; however, his results were also atypical for a developmental dyslexic. The literature shows that dyslexics typically have difficulties with reading processes in the LH (Lavidor, Johnston & Snowling, 2006) and this is possibly due to the LH using phonological decoding in order to read.

DB did not follow this pattern, and indeed his performance showed no such inhibitory effects in the RVF/LH.

The controls did show an effect of word position, as expected, performing best to words presented centrally, then to words presented to the RVF/LH, and worst to words presented to the LVF/RH. The controls also showed a regularity effect, with regular words having the highest accuracy levels, then exception words, then strange words. It was thought possible that DB's results were being confounded by the presence of other lexical properties, such as orthographic neighbourhood size (N). Experiment 4, a neighbourhood LDT, was aimed at exploring this possibility and at further clarifying the reasons for DB's performance.

In Experiment 4, DB still showed no inhibitory effects of LH processing. A significant interaction between word type and position was found. When presented centrally DB responded fastest to high lead 'N' low end 'N' words and slowest to low lead 'N' low end 'N' words. When presented to the LVF/RH, DB was fastest to respond to high lead 'N' low end 'N' words, and slowest to respond to high lead 'N' high end 'N' words. When presented to the RVF/LH, DB was fastest to respond to low lead 'N', high end 'N', and slowest to respond to high lead 'N' low end 'N' words. Again, this is contrary to what was expected, as no effect of word type was predicted for words presented to either side of the fixation point. This pattern of results is atypical for a developmental dyslexic and yet it does not follow the pattern of non-dyslexics either.

The controls did show an effect of word position, with words presented centrally having a faster and more accurate response, then the LH, with the RH having the slowest and least accurate response. The controls also showed a significant effect of word type with low lead 'N' , high end 'N' words being responded to more accurately than high lead 'N' high end 'N' words, which were responded to more accurately than high lead 'N' low end 'N' words, and low lead 'N' low end 'N' words were responded to least accurately of all. The controls also showed a significant interaction between word position and word type, although the control groups' pattern was very different to DB's. The control group, on central presentation, was fastest to respond to low lead 'N' low end 'N' words and slowest to high lead 'N' high end 'N' words. When presented to the LVF/RH, instead of showing no effect, the controls were significantly faster to respond to low lead 'N' high end 'N' words, and slowest to high lead 'N' low end 'N' words. When presented to the RVF/LH, controls also showed an effect of word type, contradictory to what was hypothesised. They responded fastest to high lead 'N' low end 'N' words, and slowest to high lead 'N' high end 'N' words. These results are very surprising, and are the opposite of what was hypothesised. The results may have been due to the high lead 'N' high end 'N' words activating too many other lexical entries, thus slowing performance rather than facilitating it.

Once again, DB's pattern of results appears to be unique. He does not follow the expected pattern for a developmental phonological dyslexic, nor does he follow the pattern of non-dyslexics. He shows a preference for visually distinct words, in Experiment 1, which would fit the profile of a developmental dyslexic; however, there was little evidence that this was due to phonological problems, as he had only mildly impaired non-word reading skills. He does not show inhibited processing of words in the LH, nor facilitatory effects of RH processing, as would be expected for a developmental dyslexic. DB also does not show LH facilitatory effects or RH inhibitory effects as a non-dyslexic might.

His reading errors on the non-word task were near the beginning or the middle of the stimuli, not at the ends of the non-words. It is therefore conceivable to suggest that when faced with novel stimuli DB is forced to scan and process all the letters, whereas perhaps he does not do this when familiar words are presented. It is possible that DB has compensated for his reading difficulties by scanning only enough of the word to minimise the possible choices, and, in combination with context, to select the correct word without having to scan, process and retain letter order of all letters present.

The pattern that DB is showing tells us that he does not read via the same strategies as most developmental dyslexics or via the strategies of a non-dyslexic, but it does not tell us by what strategies he does read. It is possible that there are additional lexical properties, other than regularity, frequency, and orthographic neighbourhood size, which are confounding the results of these experiments; however, other researchers have used the same stimuli and did not appear to have any difficulties in this respect (e.g. Seidenberg, Waters, Barnes & Tanenhaus, 1984; Lavidor & Walsh, 2003).

It was thought probable, however, that due to his age and attendance at university, DB would have become too accustomed to his particular compensatory reading strategies to be able to be taught a more phonological approach within the time constraints of this thesis. Instead, it was deemed more appropriate to examine other, younger, poor readers to see whether they also read visually, and whether they could learn a more phonological approach to improve their reading accuracy.

In Chapters Five and Six, therefore, another study was carried out, involving group of high-school pupils with reading difficulties participating in a reading intervention programme using a synthetic phonic approach. The inclusion of synthetic phonics was due to recent research showing that this form of phonics is very beneficial for beginning readers. The study was designed to see if it was also effective for older children making slow progress in learning to read.

CHAPTER FIVE

A Reading Booster Programme for Older Children

Within literacy research, it is widely accepted that there are three broad categories of methods used to teach reading. These consist of a 'whole-word' approach to reading, a 'whole-language' approach to reading, and 'systematic phonics' teaching. The whole-word approach teaches the beginning reader to look at each word as a whole, to memorise its shape, and any distinctive features, and to arbitrarily relate the word form to its meaning. Children are first taught high frequency monosyllabic words, regardless of their spelling-sound regularity, and teaching methods usually include the use of 'flash-cards', presenting the words out of meaningful contexts, with the beginner reader using the letter-shapes and overall word shape to recognise the word and its meaning. The whole-language approach to teaching reading, however, encourages the beginner to use all available cues for word recognition. This includes letter-sound mappings, letter shapes, word-shapes, accompanying pictures and sentence-context. Words are presented in sentences, and paragraphs, with stories and poems forming a major part of the teaching curriculum, with emphasis given to semantics and pronunciation of unfamiliar words.

The third teaching method is the systematic phonics approach that first teaches novice readers all of the major grapheme-phoneme correspondences in a clearly defined sequence. Letters are introduced by their sounds, not their names (i.e. *A* = *'ah'* as in *'cat'*, *B* = '*bu*' as in '*bat*'), and are first introduced with their most commonly used sounds, and then their alternative sounds are taught (e.g. '*A*' = '*cat*', '*water*', '*bath'*, ‘*cake*’). Digraphs are then introduced (e.g. ‘*ie*’, ‘*ea*’, ‘*ou*’), followed by consonant blends (e.g. ‘*str*’, ‘*fl*’, ‘*br*’) and word endings (e.g. ‘*ing*’, ‘*ed*’, ‘*er*’). There are two main approaches within systematic phonics teaching, that of analytic phonics, and that of synthetic phonics.

The American National Reading Panel described analytic phonics methods as teaching children whole-words before teaching them to analyse these into their component parts, and emphasise their larger sub-parts of words (i.e. onsets, rimes, phonograms, spelling patterns) as well as phonemes. They also described synthetic phonics methods as emphasising the teaching of students to convert letters (graphemes) into sounds (phonemes) and then to blend the sounds to form recognisable words, (National Institute of Child Health and Human Development, 2000). One of the principles of synthetic phonics is that a reader should never be asked to read something that is too difficult for them, or that they do not have the skills to read.

Johnston and Watson (2004) carried out a study of 304 children, aged five years, into the effectiveness of synthetic phonics teaching. There were three groups of children, one group was taught using analytic phonics only, one group received analytic phonics and phonological awareness training, and one group were taught synthetic phonics only. The critical feature of the synthetic phonics programme was that the children were taught to sound and blend for reading, and segment the spoken word for spelling. The children in the analytic phonics programmes were learning letter sounds in the initial position of words, with one of the groups also learning to blend phonemes, and segment spoken words into phonemes, without using letters. Each of the training programmes lasted 16 weeks. At the end of the programme, the synthetic phonics taught group were reading and spelling seven months ahead of chronological age. They read words around seven months ahead of the other two groups, and were eight to nine months ahead in spelling. When a child can segment spoken words into phonemes, for example, *'cat'* into *'c' 'a' 't'*, then they can be said to have achieved phonemic awareness (Snowling, 2000). The synthetic phonics taught group also showed significantly better phonemic awareness skills compared to the other two groups, performing better even than the analytic phonics group that had had a phonological awareness training programme.

The development of phonemic awareness skills is an important factor to consider in reading development; indeed some researchers see this as the driving force behind the development of word recognition skills (e.g. Bradley & Byrant, 1983; Byrne, 1998). In fact, many researchers have shown dyslexics to have deficits in their phonemic awareness skills (e.g. Bradley & Byrant, 1978; Bruck, 1990). Some researchers have suggested that phonological analysis (the ability to segment whole words into their constituent sounds) and phonological synthesis (the ability to blend sounds together to form whole words) are two separate sub-processes, and are distinct enough to be conceptualised as separate aspects of phonological awareness (e.g. Castles & Coltheart, 2004; Wagner & Torgesen, 1987). Share (1995) suggests that phonological synthesis may be the requisite skills for reading and phonological analysis may be the requisite skill for spelling. Phonological awareness, however, has also been shown to be influenced by literacy acquisition (Duncan, Cole, Seymour & Magnan, 2006), thus suggesting a reciprocal relationship between the two.

The National Reading Panel (National Institute of Child Health and Human Development, 2000) conducted a quantitative meta-analysis of 52 studies, evaluating the effects of phonemic awareness instruction on learning to read and spell. The findings of this meta-analysis were reported by Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh and Shanahan (2001). The meta-analysis showed that phonemic awareness instruction significantly improved acquisition of phonemic awareness, significantly improved reading ability, and significantly improved spelling ability, and in all cases was more effective than other forms of instruction or no instruction. The greatest gains, however, were shown when letters were used in the training. Once letters are introduced, however, phoneme awareness training becomes inextricably linked with a phonics approach to teaching.

It can therefore be argued that systematic phonics teaching, particularly synthetic phonics, is the most effective method of instruction when learning to read, and researchers have consistently observed that synthesis skills are much easier to teach than analysis skills (e.g. Torgesen, Wagner & Rashotte, 1994; Torgesen, Morgan & Davis, 1992). The current study therefore utilised a reading intervention programme already found in many secondary schools throughout the UK, *Toe-By-Toe*, and, with the authors permission, added a synthetic phonics element throughout.

The children studied attended a school in an area of significant deprivation, and for some researchers this would raise issues about whether the children were dyslexic or not. Francis, Shaywitz, Stuebing, Shaywitz and Fletcher (1996) carried out a longitudinal study which followed the reading development of 403 children from the age of seven years into adolescence. One group consisted of average level readers, one group consisted of IQ discrepant readers (whereby their reading ability was below average and below that expected for their IQ level) and one group consisted of non-discrepant, or 'garden variety', poor readers (whereby their reading ability was in line with their IQ level, where both were below average). The performance of both groups of poor readers at age seven was substantially below that of the average readers, and this difference remained substantial throughout development. Additionally, growth in word reading skill for the two groups of poor readers was identical. These results provide strong support for the conclusion that both 'classically reading disabled' and 'garden variety' poor readers have similar long-term outcomes for word reading skill, an area in which both groups of poor readers showed substantial deficits throughout development.

Research by Gustafson and Samuelsson (1999) lends further support to this claim, with their results showing that both high and low IQ poor readers showed the same patterns of reading performance, thus indicating that both groups might benefit from the same remedial programmes. The current study therefore employed poor readers fitting either description, and no distinction was made between students who had, or did not have, a discrepancy between their IQ and reading ability. One group received individual reading tuition using *Toe by Toe*, an approach that develops word recognition skills, and another group received no specific remedial reading tuition.

MacKay and Cowling (2004) had previously carried out two intervention programmes using *Toe-by-Toe*. The first consisted of 24 secondary school pupils with low reading ages who were split into two groups; one receiving normal school learning support, and the other receiving *Toe-by-Toe* instruction for twenty minutes every school day for three months. The control group made average gains of five months, whilst the experimental group gained on average forty-two months. The second intervention used 104 children aged over 11 years, and post-testing after five months, before the intervention programme was complete, pupils showed average gains of 14 months. Unfortunately, however, these studies were not subjected to inferential statistical analysis.

Torgesen (2002) stated that in order to be considered successful, a reading intervention programme should substantially increase the efficiency with which pupils identify words in text, thus helping them to acquire efficient word-level skills. It was therefore hypothesised, given Mackay and Cowling's (2004) findings, and Torgesen's statement, that the current intervention group would make not only substantial gains in single-word reading, but also in reading comprehension, spelling and phoneme awareness skills. It was also hypothesised that IQ scores in general would not improve with the intervention, but that given the serial processing method used in the reading intervention (i.e. sounding and blending the letter sounds in unfamiliar words), digit span scores might improve in this condition. It was also hypothesised that the control group would make small, but insignificant, improvements in their literacy and phoneme awareness skills, but not in their IQ scores.

Experiment 5:

Participants

The students in the study had learned to read using the English government programme Progression in Phonics (DfEE, 1998). This is a mixture of an analytic phonics programme and a whole language approach. Although children using this method learn to sound and blend for reading, and to segment spoken words for spelling, they are also taught to use text and pictures to guess unfamiliar words. Thus there is a possibility that sounding and blending for reading is not well-established in children learning by this method.

Twenty-five high school pupils (aged between 11.75 and 12.67; M = 12.14, SD = 0.33) initially took part in this study. Participants were selected based on their reading ability on entry to the school; any pupil in year seven who had a reading ability more than three years below chronological age, but who was not already receiving remedial help, was available for this study. Pupils took part on a voluntary basis. Three pupils expressed a wish to withdraw from the study, and two pupils were reading at their age level or above; this left 20 pupils participating (aged between 11.67 and 12.58; M = 12.13, SD = 0.34). Parental consent was sought before any testing began.

Procedure

Toe-by-Toe is an established reading intervention programme for use in secondary schools for dyslexic pupils reading behind their peers. It is intended for use one-to-one with a teacher or parent. The programme begins with the sounds of the letters and gradually introduces the student to reading, using a combination of word and non-word reading, syllable games and revision exercises. Simple consonant-vowel-consonant words (e.g. 'pot', 'hat', and 'can') are the basis for the first few exercises, with a very gradual build up to more complex word formations, ending with the student being able to read words such as 'chiropodist' and 'psychologist'. Although this is a phonic programme, pupils are not asked to use sounding and blending to read. With the author's permission the approach was modified so that a synthetic phonics element was introduced to aid student learning. This took the form of instructing pupils to elongate the letter sounds, and to blend the individual sounds together to form the word or non-word. Only one student began the programme spontaneously using this method, with all other students opting to 'guess' the word from context or initial letter clues before giving up. As the programme progressed, with much reminding and encouragement, the pupils began using sounding and blending much more frequently and spontaneously during their lessons.

It was recommended, by the author of *Toe-by-Toe,* that the pupil be given lessons of no longer than twenty minutes, on a daily basis, ensuring that at least twenty-four hours had passed between lessons. In the current study, however, it soon became apparent that some students were unable to study continuously for twenty minutes, and so lessons became just fifteen minutes per day per pupil. Lessons occurred every morning at the same time, to maximise consistency, and to minimise disruption to the school's timetables and classes. The lessons were conducted on a one-to-one basis in a partitioned section of the school's library, although other classes of pupils were often being taught on the computers in the main section.

Initially, the intervention was to be run daily for the duration of two school terms to asses the impact on the pupils' reading abilities. Unfortunately, however, due to the poor reading ability of the pupils, and due to their lack of school attendance (median = 59.5%, SD = 14.43%), progress was much slower than anticipated, and it was decided to run the intervention until at least one pupil finished the programme. Once this had happened, that pupil was post-tested and resumed normal school routines. It took that pupil exactly 965 minutes to complete *Toe-by-Toe*. To ensure comparable teaching time, once another pupil had been taught for 965 minutes, independently of whether they had completed the programme or not, they were also post-tested, and then given the option of remaining on the programme (but without further post-testing or attendance monitoring) or returning to usual school routines. Only once all pupils had completed the allotted time of 965 minutes and had been post-tested, did the intervention stop, with no further option to continue.

Two pupils had completed all 287 pages of the intervention before the post-tests were conducted. The remaining eight pupils were at varying stages of the programme (mean = 70.73% completed, SD = 18.29%), with one pupil only 44.25% through the intervention on page 127, and another pupil 93.78% of the way through (page 269), when they were post-tested having completing 965 minutes of tuition. The intervention group were post-tested on average 51.7 weeks (SD = 4.55) after pre-testing had finished, and only once the entire intervention group had been post-tested, were the non-intervention group then post-tested also, 60 weeks after pre-testing had finished. During the intervention period, the non-intervention group received normal school lessons, and minimal extra remedial help. This intervention took four school terms.

Materials

Eight tests were conducted to measure the children's IQs and their abilities on a wide range of skills related to reading.

IQ

Verbal IQ was measured using the WISC Vocabulary Test (Weschler, 2004), where participants were asked to give definitions for 31 words that the examiner read aloud e.g. 'clock'. Two points were available for each word, and full marks were awarded for a good synonym, a major use, a general classification to which the word belongs, one or more primary features, or several descriptive features; e.g. 'it tells you the time' or 'it has numbers and hands'. One point was scored for a correct response with lack of content, a vague synonym, a minor use (not elaborated), a correct attribute which is not a definitive feature (and is not improved after querying), an example using the word itself (not elaborated), a correct definition of a related form of the word, or a concrete interpretation of the word (not elaborated); e.g. 'it tells you when to go home' or 'it goes tick-tock'. No points were scored for an obviously incorrect response, a verbal response with no real understanding shown after query, a physical demonstration that was not elaborated in words, responses that were not totally incorrect, but which, even after questioning, were vague, trivial, or demonstrated a lack of content, or regionalisms and slang not found in dictionaries; e.g. 'wakes you up' or 'hangs on a wall'.

Performance IQ was tested using the WISC Block Design (Weschler, 2004), a task consisting of eleven pictured designs, which the participant was asked to recreate as quickly and as accurately as possible with the blocks given. Scoring varied according to time taken to complete the reconstruction, with an incorrect response or over-long response time scoring nil.

Digit span was tested using the sub test of the CTOPP (Wagner, Torgesen & Rashotte, 1999), *Memory for Digits*. The examiner read out 21 graded number sequences beginning with just two digit sequences and ending with eight digit sequences. After each sequence, participants were required to repeat the numbers in the same order as they heard them in. Responses were marked correct or incorrect and a raw score was given which was then converted into a standard score.

Literacy Skills

The first test administered was the British Abilities Scale Word Reading Test II (BAS II) (Elliott, Smith & McCulloch, 1996). This test consisted of 90 graded words to be read aloud. These ranged from very simple words (i.e. 'up') to words which were more complex (i.e. 'criterion'). Participants were asked to read aloud every word on each line, and were encouraged to guess if they were uncertain of a word. Once they received ten consecutive errors, they were asked to scan the rest of the page for any other word that they knew. The test finished when either they read, or attempted, all words, or when, after ten consecutive errors, they read, or attempted, all other words they recognised. Tests were scored according to how many words were correctly read in total, which was then converted into a reading age, using the conversion table.

Spelling ability was measured using the Schonell spelling test (Schonell, 1932). Participants were asked to spell up to 100 words graded in difficulty from 'see' to 'committee'. The examiner spoke the word, then used the word in a sentence, then spoke the word again before participants were asked to write the word using the correct spelling. Participants were encouraged to write down as much of the word as they could, even if uncertain of its correct spelling. Testing ceased when participants spelled ten consecutive words incorrectly. Correct scores were converted to spelling ages.

Reading comprehension was measured using the Neale Analysis of Reading Ability (NARA) (Neale, 1989), form two. This test consisted of six passages to be read aloud by the participants. These passages were graded in length and difficulty. Once the passage had been read aloud, participants were asked a series of questions to check their comprehension. Scoring was according to the number of correct responses. Whilst participants read, the experimenter marked any errors, and categorised them according to ’mispronunciations’, ‘substitutions’, ‘refusals’, ‘additions’, ‘omissions’, and ‘reversals’. The time taken to read the passage was recorded. Testing ceased when participants made errors in excess to the limit given to that passage (usually 16, but 20 in the last passage), or when the experimenter judged that the participant would be unable to master the next level of difficulty due to the lack of comprehension shown. Scores were derived for reading accuracy, reading rate, and reading comprehension).

Listening Comprehension was also tested with form one of the NARA task. This task consisted of six passages to be read aloud by the experimenter, and listened to by participants. These passages were graded in length and difficulty. Once the passage had been read aloud, participants were asked a series of questions to check their comprehension, scoring was according to the number of correct responses. Testing ceased when the experimenter judged that the participant would be unable to master the next level of difficulty due to the lack of comprehension shown.

Phonological Reading Skill

A non-word reading task consisting of 20 graded non-words for reading (taken from Snowling, Stothard & McLean, 1996) ranging from simple (i.e. ’hast’) to complex (i.e. ‘sloskon’) was administered (see appendix 11). The computer used was a Dell laptop with a 60hz refresh rate. The screen was 14 inches, with a resolution of 1024x768. The program used was MemWord version 1.4. Each non-word appeared individually in the centre of the screen, and disappeared the moment the participant began to speak into the microphone. Responses were carefully recorded phonetically by an experimenter. Accuracy was recorded, as was the actual pronunciation given; the reaction times for each correct response were also recorded.

Phoneme Awareness

A Comprehensive Test of Phonological Processing (CTOPP) (Wagner, Torgesen & Rashotte, 1999) was carried out, to provide a detailed view of the participant's strengths and weaknesses in this area. There are several sub-tasks making up this test, however, and only the following sub-tasks were used:

Elision: Participants were asked to say 20 stimulus words, with part of the word removed, i.e. 'cup' without the 'c'.

Blending Words: The examiner read out words in several parts, and asked participants to say the whole word, i.e. 'm-i-s', participants should respond 'miss'.

Non-Word Repetition: The examiner read out 21 nonsense words, and after each the participant was asked to repeat the nonsense word exactly as they had heard it.

Phoneme Reversal: participants were asked to say 18 made-up words, then to say the same words backwards to give a real word, i.e. 'ood' becomes 'do'.

Blending Non-Words: The examiner read out non-words in several parts, and asked participants to say the whole non-word, i.e. 'n-a-s', participants should respond 'nas'.

Segmenting Words: The examiner read 20 words, and after each, participants were required to repeat the word, then to say each word one sound at a time, i.e. 'pig' becomes 'p-i-g'.

Segmenting Non-Words: The examiner read 20 non-words, and after each, participants were required to repeat the non-word, then to say each non-word one sound at a time, i.e. 'seb' becomes 's-e-b'.

Responses were marked correct or incorrect for each part of each sub-task, and a raw score was given for each sub-task, which was then converted into a standard score.

Results of pre-tests

The eight aforementioned tests were administered to all participants, and then the children were paired, matched as closely as possible, primarily in terms of reading age, reading comprehension, and spelling ability. They were also matched as closely as possible on performance and verbal IQ as well as phonemic awareness ability. A child from each pair was randomly assigned to either the experimental or the non-intervention group. Ten children were in the non-intervention group, and ten children were in the intervention group (see Tables 17, 18 and 19 for means and standard deviations of test scores per group).

Table 17

Pre-tests of general abilities for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
WISC Vocabulary STD Score	70.00	13.54	65.50	7.25
WISC Block Design STD Score	85.50	12.57	81.50	9.73
CTOPP Digit Span STD Score	70.00	33.67	101.00	34.79

One way ANOVAs were carried out on all test scores to compare the non-intervention group to the intervention group, to ensure there were no significant differences in ability prior to the intervention. There were no significant differences between the groups in terms of WISC Vocabulary Standard scores ($F(1, 18) = <1$), WISC Block design Standard scores ($F(1, 18) = <1$), or CTOPP Digit Span Standard Scores ($F(1, 18) = 4.1$, $p = 0.058$, NS).

Table 18

Pre-tests of literacy skills for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
Chronological Age	12.10	0.38	12.20	0.31
BAS Reading Age	8.58	1.08	8.36	0.96
Schonell Spelling Age	8.62	0.73	8.77	1.5
NARA Accuracy STD Score	77.70	8.60	75.8	8.24
NARA Rate STD Score	90.10	12.06	89.30	11.94
NARA Comprehension STD Score	77.60	7.83	77.40	8.19
NARA Listening Comprehension STD Score	78.50	7.78	75.30	6.33
Graded Non-Word Task Accuracy	51.50%	22.98%	56.50%	18.86%
Graded Non-Word Task Mean Reaction Times (ms)	2551.95	1658.75	2192.80	1287.52

One way ANOVAs were carried out on all test scores to compare the non-intervention group to the intervention group, to ensure there were no significant differences in ability prior to the intervention. There were no significant differences between the groups in terms of Chronological age (F(1, 18) = <1), BAS reading age (F(1, 18) = <1), Schonell Spelling Age (F(1, 18) = <1), NARA Accuracy STD Score (F(1, 18) = <1), NARA Rate STD Score (F(1, 18) = <1), NARA Comprehension STD Score (F(1, 18) = <1), NARA Listening Comprehension STD Score (F(1, 18) = 1.02, p>0.05), Graded Non-Word Task Accuracy (F(1, 18) = <1, p = 0.60, NS), or Graded Non-Word Task Reaction Times (F(1, 18) = <1, p = 0.60, NS).

Table 19

Pre-tests of phoneme awareness for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
CTOPP Elision STD Score	65.00	30.28	71.00	39.29
CTOPP Blending Words STD Score	94.00	20.66	101.00	19.12
CTOPP Non-Word Repetition STD Score	97.00	25.41	93.00	21.11
CTOPP Phoneme Reversal STD Score	85.00	22.73	82.00	21.50
CTOPP Blending Non-Words STD Score	101.00	31.78	93.00	20.03
CTOPP Segmenting Words STD Score	83.00	29.46	100.00	23.09
CTOPP Segmenting Non-Words STD Score	93.00	27.91	99.00	31.78

One way ANOVAs were carried out on all test scores to compare the non-intervention group to the intervention group, to ensure there were no significant differences in ability prior to the intervention. There were no significant differences between the groups in terms of CTOPP Elision STD Score ($F(1, 18) = <1$), CTOPP Blending Words STD Score ($F(1, 18) = <1$), CTOPP Non-Word Repetition STD Score ($F(1, 18) = <1$), CTOPP Phoneme Reversal STD Score ($F(1, 18) = <1$), CTOPP Blending Non-Words STD Score ($F(1, 18) = <1$, $p = 0.51$, NS), CTOPP Segmenting Words STD Score ($F(1, 18) = 2.06$, $P=0.168$, NS), or CTOPP Segmenting Non-Words STD Score ($F(1, 18) = <1$).

Results of post-tests

IQ

Two-way mixed ANOVAs were carried out on all test scores and in each case the between subjects factor was group (non-intervention and intervention) and the within subjects factor was test time (pre and post).

WISC Vocabulary STD Score

A two-way mixed ANOVA was carried out on the WISC Vocabulary STD Score results. There was no significant main effect of testing time ($F(1, 18) = 1.424$, $p>0.05$). There was no main effect of groups ($F(1,18) = >1$). There was no significant interaction between testing time and groups ($F(1, 18) = <1$). (See Table 20 for means and standard deviations).

Table 20

Post-Tests of digit span for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
WISC Vocabulary STD Score	70.30	13.53	68.50	11.32
WISC Block Design STD Score	88.00	8.56	82.50	11.61
CTOPP Digit Span STD Score	88.00	25.73	98.00	28.98

WISC Block Design STD Score

A two-way mixed ANOVA was carried out on the WISC Block Design STD Score results. There was no significant main effect of testing time ($F(1, 18) = 1$, $p>0.05$). There was no main effect of groups ($F(1,18) = <1$). There was no significant interaction between testing time and groups ($F(1, 18) = <1$). (See Table 20 for means and standard deviations).

CTOPP Digit Span STD Score

Although mixed ANOVAs showed that there was no significant main effect of testing time on the performance on the CTOPP Digit Span Task ($F(1, 18) = 2.91$, $p>0.05$), there was a main effect of groups ($F(1,18) = 5.69$, $p<0.05$, with the non-intervention group performing better than the intervention group. There was also a significant interaction between testing time and group ($F(1, 18) = 5.69$, $p<0.05$), with the non-intervention group maintaining similar scores across testing time (101 at pre-test, 98 at post-test) and the intervention group gaining across testing time (70 at pre-test, 88 at post-test). (See Figure 13 for mean STD scores for testing time and group with standard deviation bars). Newman-Keuls post hoc tests on CTOPP Digit Span scores showed a significant difference in performance between the intervention group and the non-intervention group at pre-test ($p<0.01$) with the non-intervention group performing better than the intervention group (although a one-way ANOVA had shown this difference not to be significant) and at post-test ($p<0.05$) with the non-intervention group still performing better than the intervention group. Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post test ($p<0.01$), but the non-intervention group did not ($p>0.05$). (See Table 20 for means and standard deviations; see Figure 13 for mean digit span scores at pre-test and post-test with standard deviation bars).

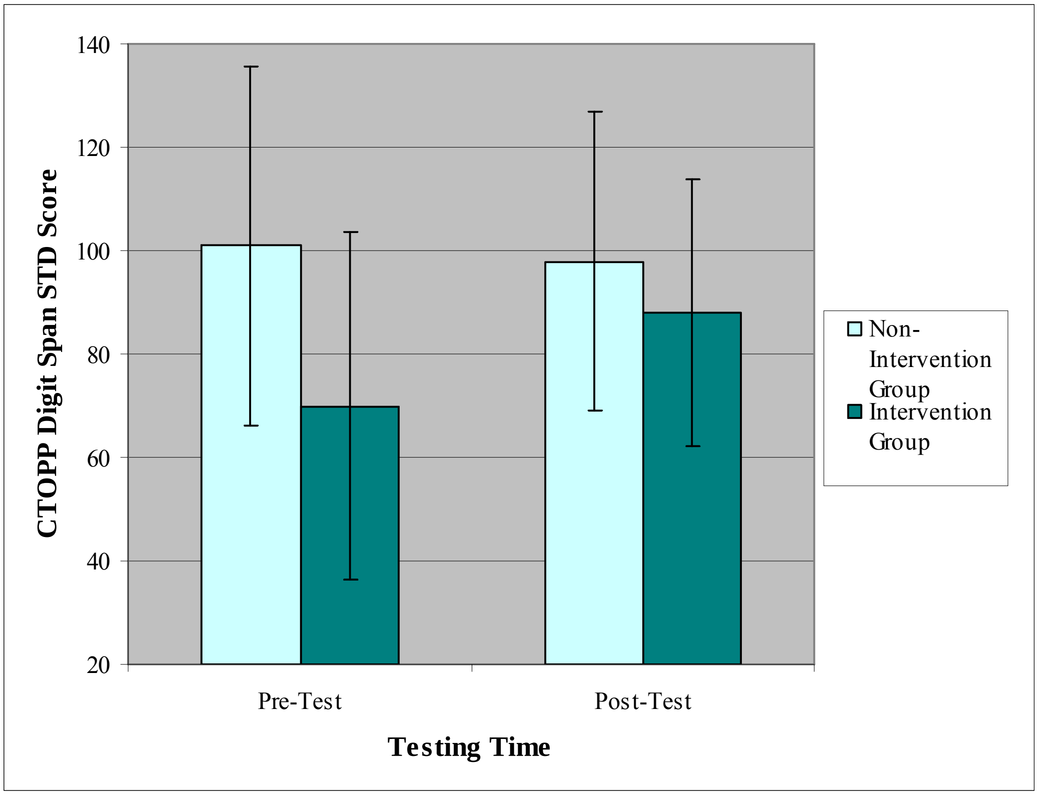


Figure 13: *Pre-and post-test CTOPP Digit Span STD scores for non-intervention and intervention groups*

Chronological Age

A one-way ANOVA was carried out on chronological age at post-test to ensure there were no significant differences between the intervention and the non-intervention group. There were no significant differences between the groups in terms of chronological age ($F(1, 18) = 1.123$, $p>0.05$) at post-test. (See Table 21 for means and standard deviations).

Table 21

Post-tests of literacy skills for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
Chronological Age	13.14	0.43	13.32	0.30
BAS Reading Age	10.57	2.00	8.86	1.48
Schonell Spelling Age	9.40	1.24	9.15	1.71
NARA Accuracy STD Score	90.40	12.35	75.90	10.17
NARA Rate STD Score	84.90	9.27	88.50	11.71
NARA Comprehension STD Score	91.70	7.35	78.00	12.04
NARA Listening Comprehension STD Score	80.00	9.03	74.20	4.76
Graded Non-Word Task Accuracy	77.92%	18.10%	60.00%	19.00%
Graded Non-Word Task Mean Reaction Times (ms)	2096.45	1332.13	1313.40	543.86

Two-way mixed ANOVAs were carried out on all test scores (except chronological age) and in each case the between subjects factor was group (non-intervention and intervention groups) and the within subjects factor was test time (pre and post).

Literacy Skills

BAS Reading Age

Mixed ANOVAs showed that there was a significant main effect of testing time on the performance on the BAS reading age test ($F(1, 18) = 31.69$, $p<0.001$). There was no significant main effect of groups ($F(1,18) = 2.57$, $P>0.05$. However, there was a significant interaction between testing time and group, ($F(1, 18) = 11.36$, $p<0.01$), with the non-intervention group making modest gains of five months across testing time (8.4

at pre-test, 8.9 at post-test) and the intervention group gaining two years across testing time (8.6 at pre-test, 10.6 at post-test). Newman-Keuls post hoc tests showed that the intervention group significantly improved in terms of reading age from pre-test to post test ($p<0.01$), and so did the non-intervention group ($p<0.05$). (See Table 21 for means and standard deviations; see Figure 14 for mean reading ages at pre-test and post-test with standard deviation bars). However, Newman-Keuls post hoc tests showed no significant difference in reading age between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$).

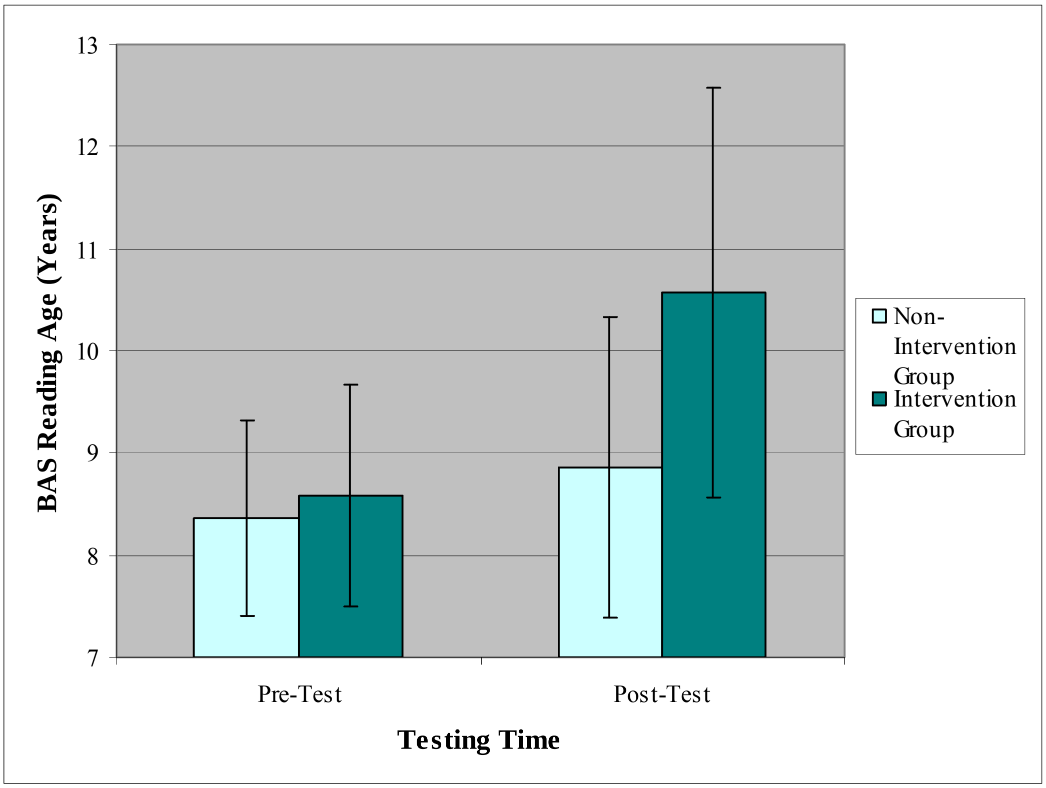


Figure 14: *Pre-and post-test BAS Reading Age for non-intervention and intervention groups*

Schonell Spelling Age

A mixed ANOVA showed that there was a significant main effect of testing time on the performance on the Schonell Spelling Test ($F(1, 18) = 11.08$, $p<0.01$), with performance increasing from pre-test to post-test. There was no main effect of groups

(F(1,18) = <1), and no significant interaction between testing time and group (F(1, 18) = 1.32, p>0.05). (See Table 21 for means and standard deviations).

NARA accuracy

A mixed ANOVA showed that there was a significant main effect of testing time on the performance on the NARA Accuracy STD Score (F(1, 18) = 12.12, p<0.01), but no significant main effect of groups (F(1,18) = 4.07, p<0.05. There was, however, a significant interaction between testing time and group, (F(1, 18) = 11.74, p<0.01), with the non-intervention group maintaining similar performance levels across testing time (75.8 at pre-test, 75.9 at post-test) and the intervention group gaining across testing time (77.7 at pre-test, 90.4 at post-test). Newman-Keuls post hoc tests showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test (p<0.01) with the intervention group performing significantly more accurately than the non-intervention group. Newman-Keuls post hoc tests also showed a significant difference between accuracy scores at pre-test and post-test for the intervention group (p<0.01) but not the non-intervention group (p>0.05). (See Table 21 for means and standard deviations; see Figure 15 for mean accuracy scores at pre-test and post-test with standard deviation bars).

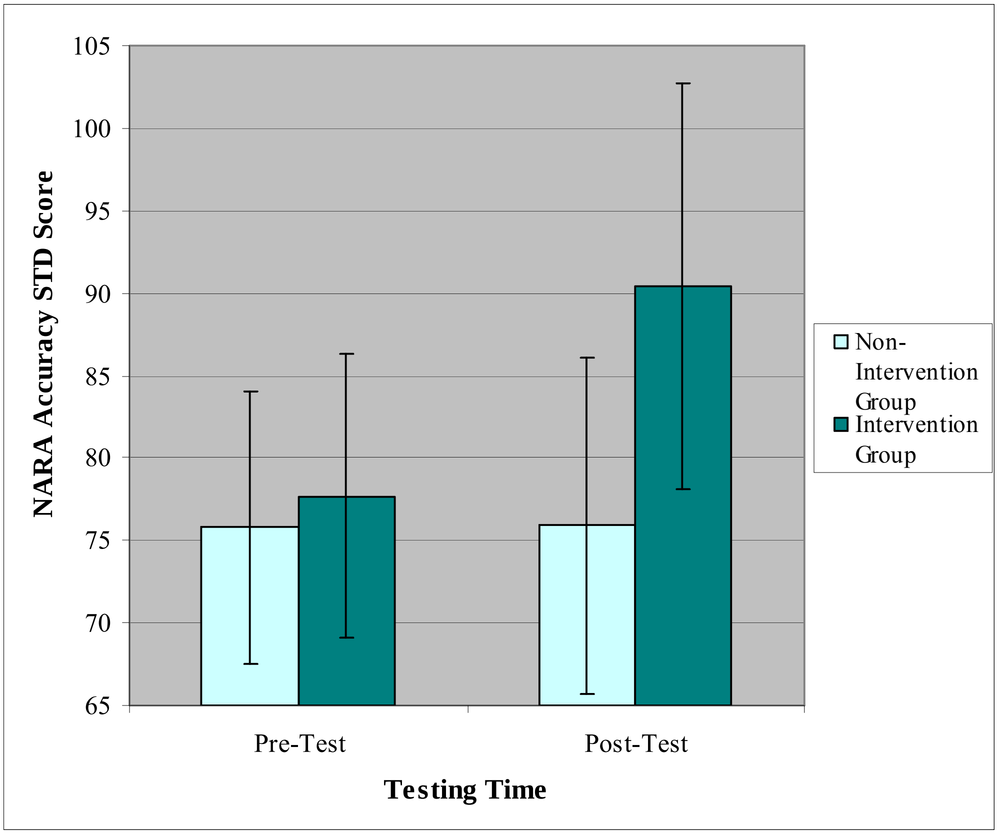


Figure 15: *Pre-and post-test NARA Accuracy STD scores for non-intervention and intervention groups*

NARA Rate

A mixed ANOVA showed that there was no significant main effect of testing time on the performance on the NARA Rate STD Score ($F(1, 18) = 1.48$, $p>0.05$), no significant main effect of groups ($F(1,18) = <1$, and there was no significant interaction between testing time and group ($F(1, 18) = <1$). (See Table 21 for means and standard deviations).

NARA Comprehension

A mixed ANOVA showed that there was a significant main effect of testing time on the performance on the NARA Comprehension STD Score ($F(1, 18) = 29.06$, $p<0.001$). There was no significant main effect of groups ($F(1,18) = 3.33$, $p>0.05$.

There was, however, a significant interaction between testing time and group, ($F(1, 18) = 24.51$, $p<0.001$), with the non-intervention group making modest gains across testing time (77.4 at pre-test, 78.0 at post-test) and the intervention group gaining more across testing time (77.6 at pre-test, 91.7 at post-test). Newman-Keuls post hoc tests showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$) Newman-Keuls post hoc tests also showed a significant difference between comprehension scores at pre-test and post-test for the intervention group ($p<0.01$) but not the non-intervention group ($p>0.05$). (See Table 21 for means and standard deviations; see Figure 16 for mean comprehension scores at pre-test and post-test with standard deviation bars).

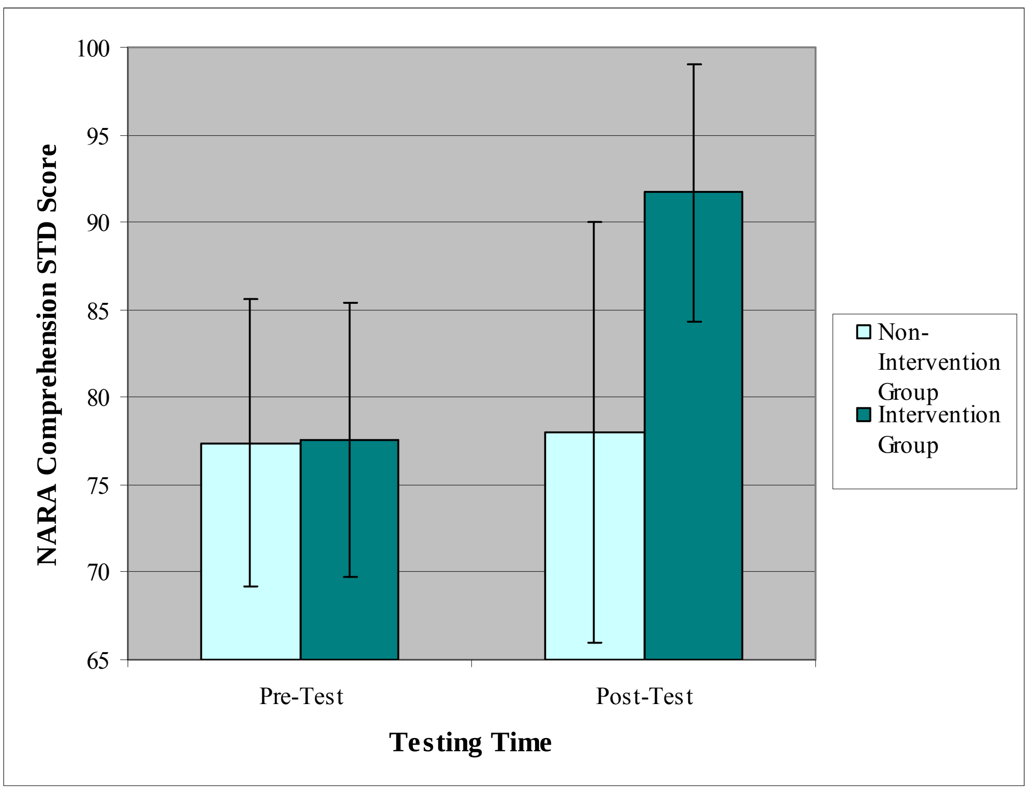


Figure 16: *Pre-and post-test NARA Comprehension STD scores for non-intervention and intervention groups*

NARA Listening Comprehension

A mixed ANOVA showed that there was no significant main effect of testing time on the performance on the NARA Listening Comprehension STD Score ($F(1, 18) = <1$), no main effect of groups ($F(1,18) = 2.19$, $p>0.05$, and no significant interaction between testing time and group ($F(1, 18) = 1.73$, $p>0.05$). (See Table 21 for means and standard deviations).

Graded Non-Word Task Accuracy

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of Graded Non-Word Task Accuracy ($F(1, 18) = 17.24$, $p<0.001$. There was no significant main effect of groups, ($F(1,18) = <1$. There was however, a significant interaction between testing time and group ($F(1, 18) = 10.12$, $p<0.05$), with the non-intervention group showing mild gains in performance across testing time (56.5 at pre-test, 60 at post-test) and the intervention group gaining more across testing time (51.5 at pre-test, 77.9 at post-test). Newman-Keuls post hoc tests on Graded Non-Word Task Accuracy showed a significant difference in performance between the intervention group and the non-intervention group at pre-test ($p<0.01$), with the non-intervention group achieving higher accuracy levels than the intervention group (although the previously reported one-way ANOVA showed this difference to be non-significant), and at post-test ($p<0.01$) with the intervention group scoring higher than the non-intervention group. Newman-Keuls post hoc tests also showed a significant difference between accuracy scores at pre-test and post-test for the intervention group ($p<0.01$) and the non-intervention group ($p<0.05$). (See Table 21 for means and standard deviations; see Figure 17 for mean accuracy scores at pre-test and post-test with standard deviation bars).

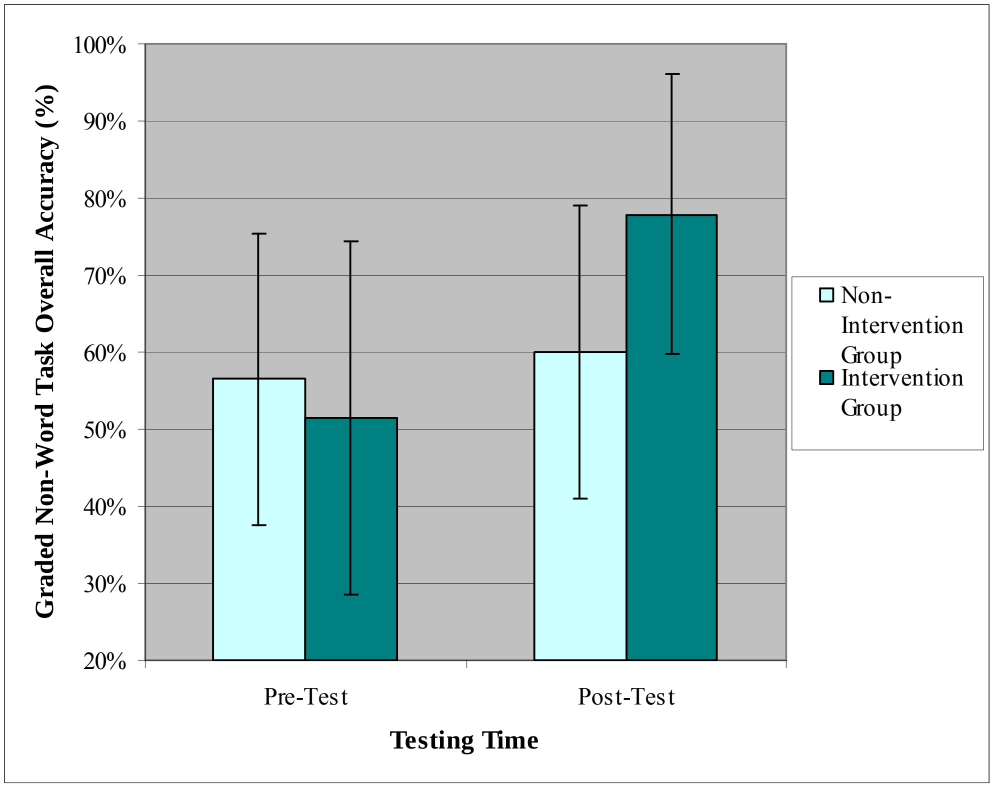


Figure 17: *Pre-and post-test Graded Non-Word Task overall accuracy levels for non-intervention and intervention groups*

Graded Non-Word Task Reaction Times

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of Graded Non-Word Task Reaction Times ($F(1, 18) = 10.49$, $p<0.05$). There was no significant main effect of groups ($F(1,18) = 1.16$, $p>0.05$. There was no significant interaction between testing time and groups ($F(1, 18) = 1.06$, $p>0.05$). (See Table 21 for means and standard deviations).

Phoneme Awareness

CTOPP Elision STD Score

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of CTOPP Elision STD Score ($F(1, 18) = 7.59$, $p<0.05$). There was no significant main effect of groups ($F(1,18) = >1$. There was, however, a significant interaction between testing time and group ($F(1, 18) = 5.64$, $p<0.05$), with the non-intervention group maintaining the same performance across testing time (71 at pre-test, 73 at post-test) and the intervention group gaining across testing time (65 at pre-test, 92 at post-test). (See Figure 18 for mean STD scores for testing time and group with standard deviation bars). Newman-Keuls post hoc tests on CTOPP Elision scores showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$). Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post test ($p<0.01$), but the non-intervention group did not ($p>0.05$). (See Table 22 for means and standard deviations; see Figure 18 for mean Elision scores at pre-test and post-test with standard deviation bars).

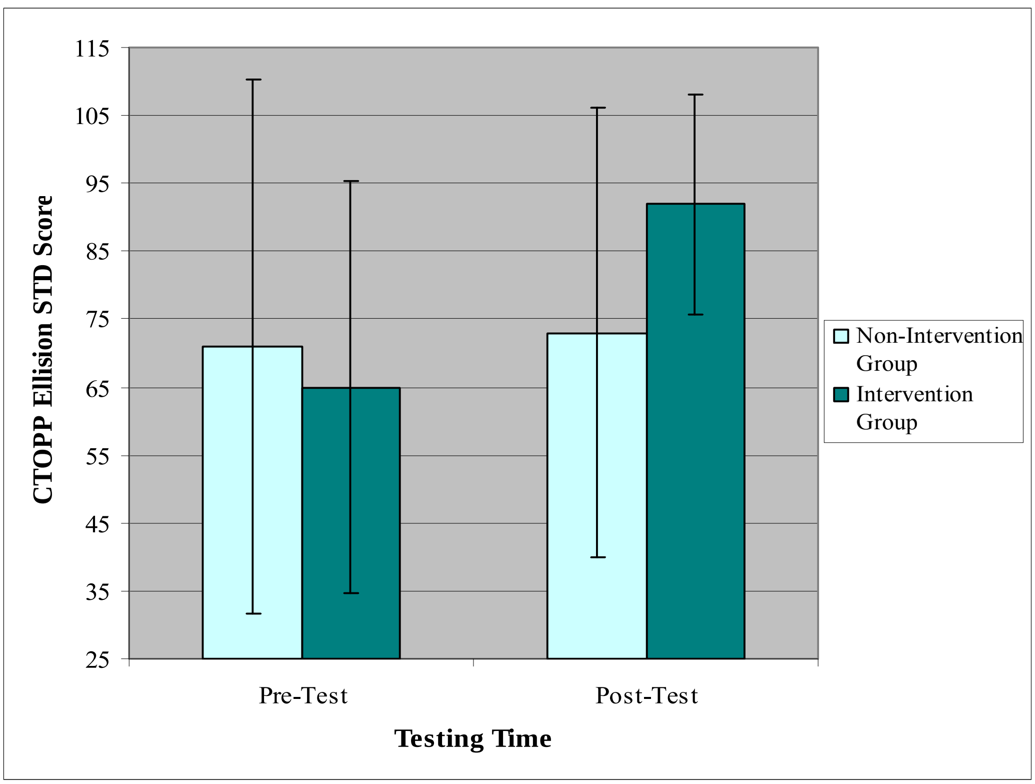


Figure 18: *Pre-and post-test CTOPP Elision STD scores for non-intervention and intervention groups*

Table 22

Post-tests of Phoneme Awareness for non-intervention and intervention groups

	Intervention group		Non-intervention group	
	Mean	SD	Mean	SD
CTOPP Elision STD Score	92.00	16.19	73.00	33.02
CTOPP Blending Words STD Score	122.00	21.50	98.00	19.89
CTOPP Non-Word Repetition STD Score	106.00	17.76	75.00	20.14
CTOPP Phoneme Reversal STD Score	98.00	26.58	87.00	24.52
CTOPP Blending Non-Words STD Score	119.00	24.70	90.00	15.63
CTOPP Segmenting Words STD Score	119.00	20.79	89.00	19.12
CTOPP Segmenting Non-Words STD Score	113.00	18.29	80.00	25.39

CTOPP Blending Words STD Score

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of CTOPP Blending Words STD Score ($F(1, 18) = 10.88$, $p<0.01$). There was no significant main effect of groups ($F(1,18) = 1.06$, $p>0.05$. There was, however, a significant interaction between testing time and group ($F(1, 18) = 16.73$, $p<0.001$), with the non-intervention group maintaining performance across testing time (101 at pre-test, 98 at post-test) and the intervention group gaining across testing time (94 at pre-test, 122 at post-test). Newman-Keuls post hoc tests on CTOPP Blending Words scores showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$), with the intervention group performing significantly more accurately than the non-intervention group. Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post-test ($p<0.01$), but the non-intervention group did not (p>0.05). (See Table 22 for means and standard deviations; see Figure 19 for mean blending words scores at pre-test and post-test with standard deviation bars).

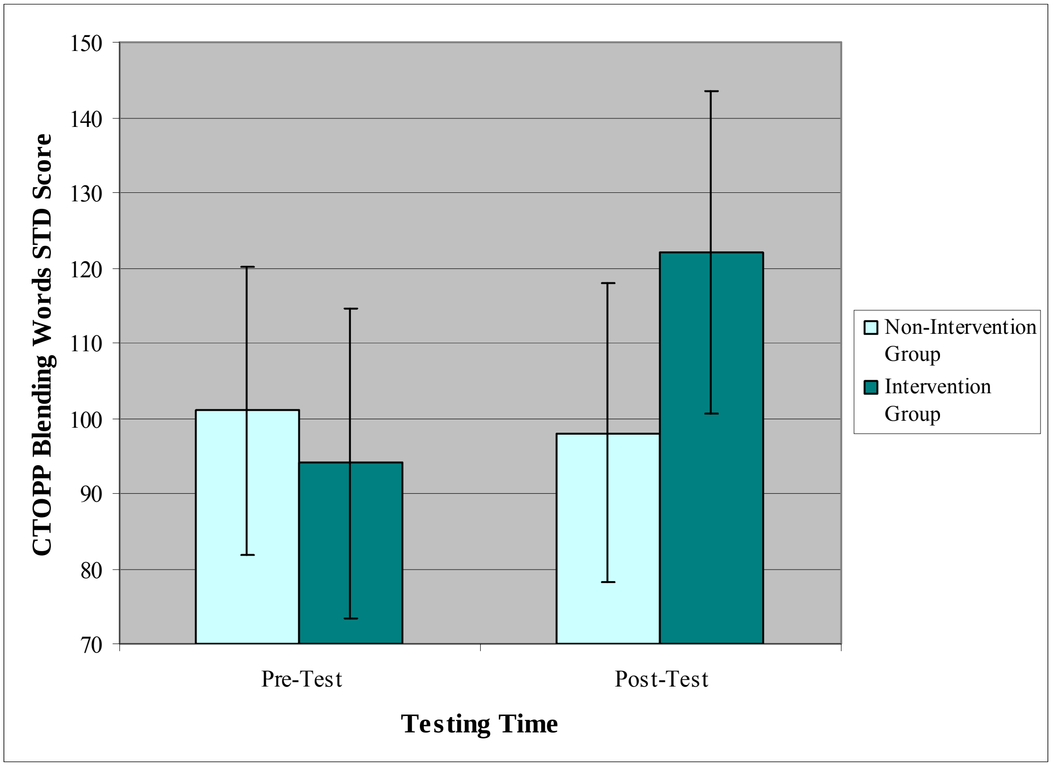


Figure 19: *Pre-and post-test CTOPP Blending Words STD scores for non-intervention and intervention groups*

CTOPP Non-Word Repetition STD Score

A mixed ANOVA showed that there was no significant main effect of testing time on the performance of CTOPP Non-Word Repetition STD Score ($F(1, 18) = 1.44$, $p>0.05$), and there was no significant main effect of groups ($F(1,18) = 4.00$, $p>0.05$. There was, however, a significant interaction between testing time and group interaction ($F(1, 18) = 12.99$, $p<0.01$), with the non-intervention group showing a drop in performance across testing time (93 at pre-test, 75 at post-test) and the intervention group gaining across testing time (97 at pre-test, 106 at post-test). Newman-Keuls post hoc tests on CTOPP Non-Word Repetition scores showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$). Newman-Keuls post hoc tests also showed a significant difference between accuracy scores at pre-test and post-test for the intervention group ($p<0.05$) and the non-intervention group ($p<0.01$), with the non-intervention group showing a significant drop in performance from pre-testing to post-testing, and the intervention group showing a significant increase. (See Table 22 for means and standard deviations; see Figure 20 for mean non-word repetition scores at pre-test and post-test with standard deviation bars).

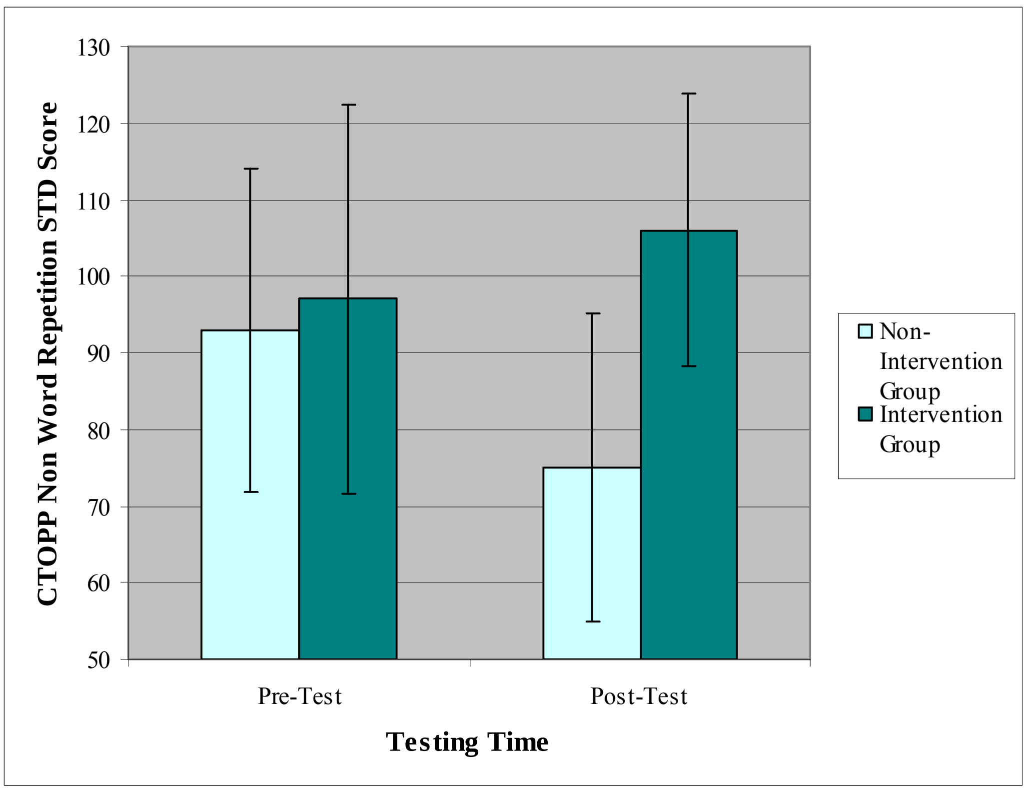


Figure 20: *Pre- and post-test CTOPP Non-Word Repetition STD scores for non-intervention and intervention groups*

CTOPP Phoneme Reversal STD Score

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of CTOPP Phoneme Reversal STD Score ($F(1, 18) = 7.55$, $p<0.05$), with scores improving overall from an average of 83.5 (SD = 21.59) at pre-test, to 92.5 (SD = 25.52) at post-test. There was no significant main effect of groups ($F(1,18) = <1$. There was no significant interaction between testing time and groups ($F(1, 18) = 1.49$, $p>0.05$). (See Table 22 for means and standard deviations).

CTOPP Blending Non-Words STD Score

A mixed ANOVA showed that there was a significant main effect of testing time on the performance of CTOPP Blending Non-Words STD Score ($F(1, 18) = 8.54$, $p<0.01$). There was no significant main effect of groups ($F(1,18) = 3.21$, $p>0.05$. There was however, a significant interaction between testing time and group ($F(1, 18) = 16.75$, $p<0.001$) with the non-intervention group maintaining performance across testing time (93 at pre-test, 90 at post-test) and the intervention group gaining across testing time (101 at pre-test, 119 at post-test). Newman-Keuls post hoc tests on CTOPP Blending Non-Words scores showed a significant difference in performance between the intervention group and the non-intervention group at pre-test ($p<0.01$) with the intervention group achieving higher scores than the non-intervention group (although a one-way ANOVA did not show this to be significant) and at post-test ($p<0.01$) with the intervention group still achieving higher scores than the non-intervention group. Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post-test ($p<0.01$), but the non-intervention group did not ($p>0.05$). (See Table 22 for means and standard deviations; see Figure 21 for mean blending non-word scores at pre-test and post-test with standard deviation bars).

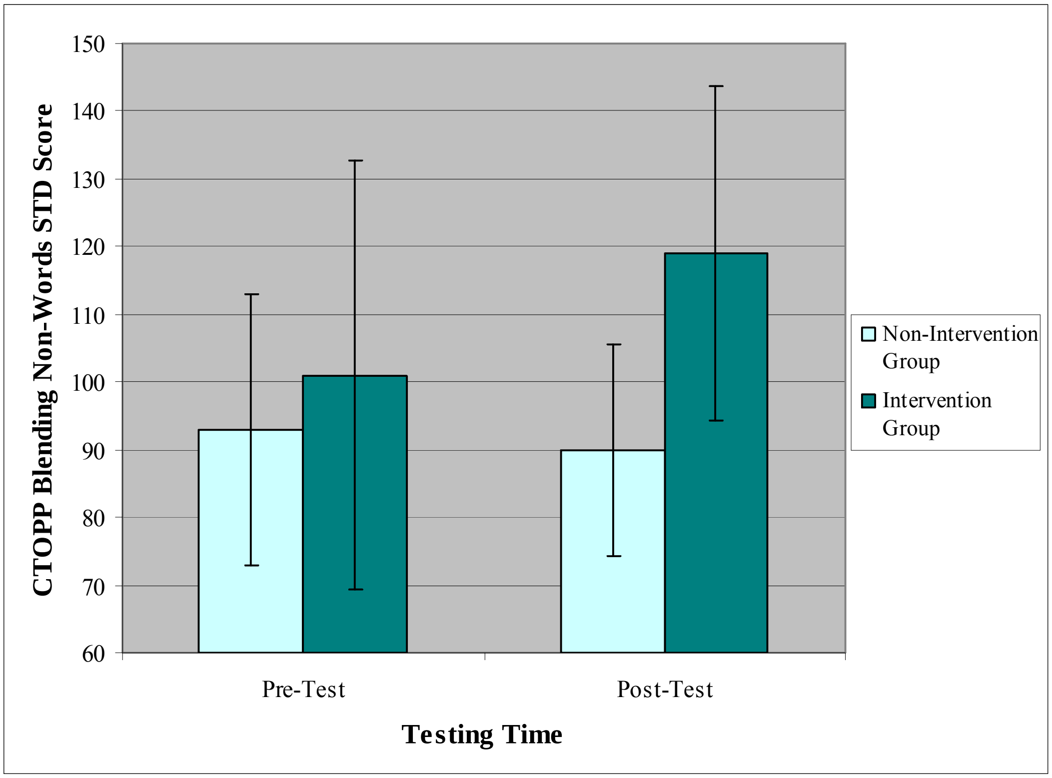


Figure 21: *Pre- and post-test CTOPP blending non-words STD scores for non-intervention and intervention groups*

CTOPP Segmenting Words STD Score

A mixed ANOVA showed that there was no significant main effect of testing time on the performance of CTOPP Segmenting Words STD Score ($F(1, 18) = 3.72$, $P>0.05$), and there was no significant main effect of groups ($F(1,18) = <1$. There was, however, a significant interaction between testing time and group ($F(1, 18) = 13.14$, $p<0.01$), with the non-intervention group showing a drop in performance across testing time (100 at pre-test, 89 at post-test) and the intervention group gaining across testing time (83 at pre-test, 119 at post-test). Newman-Keuls post hoc tests on CTOPP Segmenting Words scores showed a significant difference in performance between the intervention group and the non-intervention group at pre-test ($p<0.05$), with the non-intervention group scoring higher than the intervention group (although a one-way ANOVA did not show this to be significant) and at post-test ($p<0.01$), with the intervention group scoring higher than the non-intervention group. Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post-test ($p<0.01$), but the non-intervention group did not ($p>0.05$). (See Table 22 for means and standard deviations; see Figure 22 for mean segmenting word scores at pre-test and post-test with standard deviation bars).

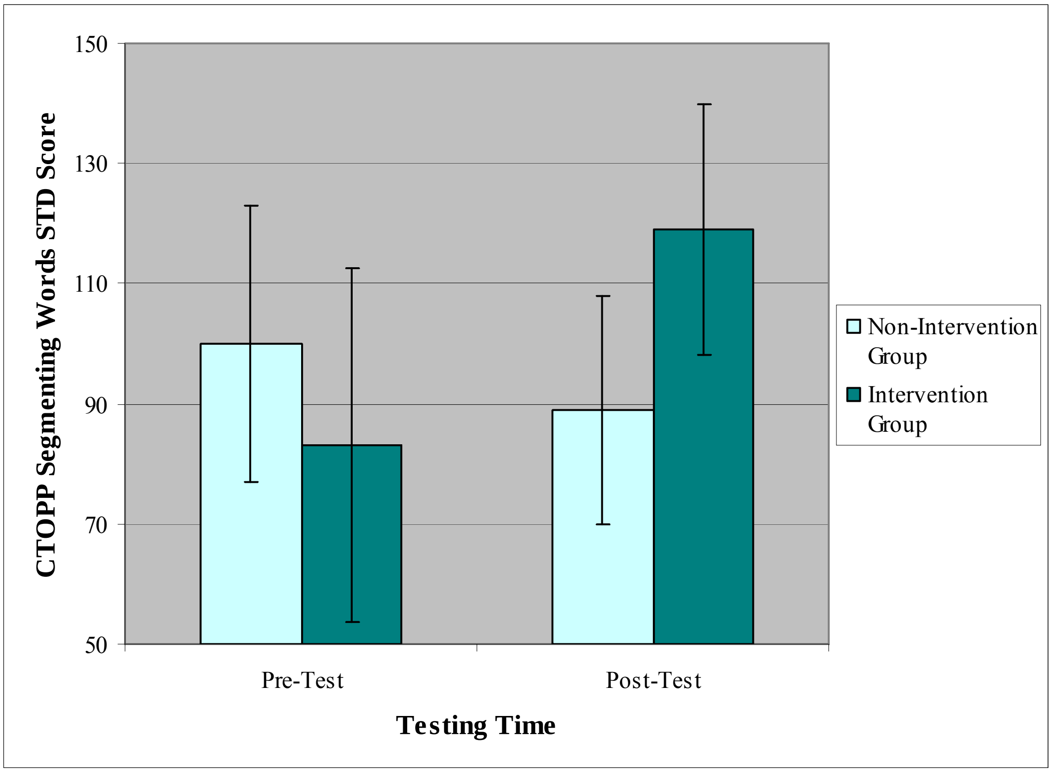


Figure 22: *Pre- and post-test CTOPP Segmenting Words STD scores for non-intervention and intervention groups*

CTOPP Segmenting Non-Words STD Score

A mixed ANOVA showed that there was no significant main effect of testing time on the performance of CTOPP Segmenting Non-Words STD Score ($F(1, 18) = <1$). There was no significant main effect of groups ($F(1,18) = 1.60$, $p>0.05$. There was, however, a significant interaction between testing time and group ($F(1, 18) = 15.40$, $p<0.001$), with the non-intervention group showing a drop in performance across testing time (99 at pre-test, 80 at post-test) and the intervention group gaining across testing time (93 at pre-test, 113 at post-test). Newman-Keuls post hoc tests on CTOPP Segmenting Non-Words scores showed no significant difference in performance between the intervention group and the non-intervention group at pre-test, but a significant difference at post-test ($p<0.01$). Newman-Keuls post hoc tests also showed the intervention group significantly improved in performance from pre-test to post-test ($p<0.01$), but the non-intervention group significantly dropped in performance from pre-test to post test ($p<0.01$). (See Table 22 for means and standard deviations; see Figure 23 for mean segmenting non-words scores at pre-test and post-test with standard deviation bars).

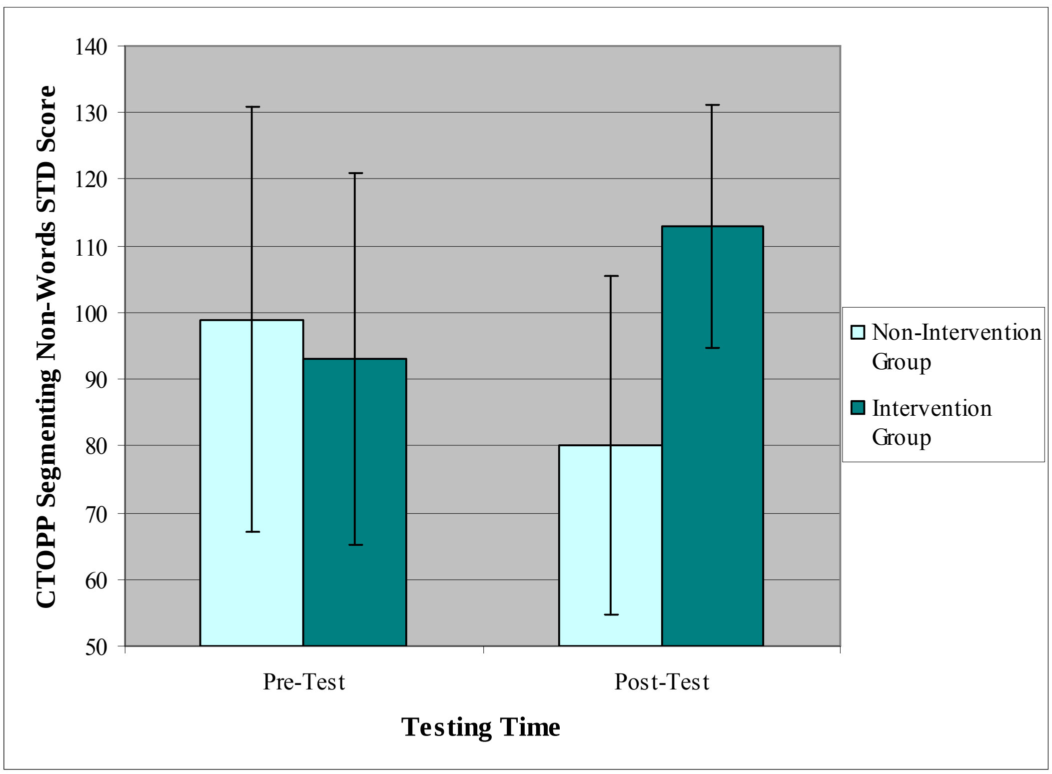


Figure 23: *Pre- and post-test CTOPP Segmenting Non-Words STD scores for non-intervention and intervention groups*

Discussion

Before the intervention programme was carried out, both the non-intervention group and the intervention group showed no significant difference in performance on any of the tests or their sub-tests. After the intervention programme, however, the intervention group had made significant improvements in terms of reading age and reading comprehension, performing well above that of the non-intervention group. They had also made significant improvements in terms of phonemic awareness, with scores on the elision task rising from the below-average range to the low-average range, scores on the blending words task improving from the low-average range to the above-average range, and scores on the segmenting words task improving from the below-average range to the above-average range. In terms of spelling ability, however, both groups made improvements at a similar rate.

The non-intervention group, who had received minimal extra remedial help within the school, had made few significant improvements, with their performance in some areas getting substantially worse over time. It must be noted, however, that the lack of any remedial treatment, or placebo, raises several issues. Firstly, it could be considered unethical to identify an educational need for this group of pupils, and yet not to attempt to rectify this. These pupils' literacy difficulties, however, were brought to the notice of the school, and would therefore be given the usual amount of assistance available, including classroom assistant support etc. Secondly, it is difficult to attribute the success of the intervention specifically to *Toe-by-Toe*, as opposed to using any intervention *per* se, as the non-intervention group received no intervention or treatment whatsoever. The reasons for the lack of treatment for the control group were purely due to time-constraints and a lack of resources available.

The non-intervention group did, however, make significant improvements in their single-word reading age, improving by five months. The intervention group, however, had improved by 24 months in the same period. Surprisingly, there was some improvement overall in terms of spelling ability, but neither group improved more than the other. This was unexpected, as the intervention group had been taught to pay particular attention to the letters and letter-order within words in order to sound and blend when reading and this might have been expected to have had some influence on spelling ability. Spelling, however, is very hard to improve in this age group, particularly for children with reading difficulties, even with direct spelling teaching (Ehri, 2001). The present study suggests that improving word recognition ability may have little impact on the development of spelling skills. The author of *Toe-by-Toe* has also published a dedicated spelling programme entitled '*Stairway to Spelling*'. The need to publish a second programme supports the idea that the remediation of word reading ability does not necessarily lead to improvements in spelling.

Toe-by-Toe introduces the pupils to individual grapheme to phoneme correspondences, practised first with the use of non-words, then with words. In this current study, an added element of synthetic phonics was employed, requiring pupils to elongate sounds of phonemes to blend them together to form the correct pronunciation. All words and non-words must be read correctly in three consecutive lessons before they are deemed to be 'learned' and any persistent errors are intensively worked upon by the pupil writing the word and saying it aloud repeatedly several times per lesson, until they can correctly read it first time in three consecutive lessons.

Stairway to Spelling ensures that pupils can read the words before being taught to spell them, and then teaches the 300 most commonly used words. The letter sounds are not emphasised, and letter-sound correspondence rules are not explicitly taught. Spelling errors are worked on until correctly written five consecutive times, with persistent errors being read aloud, analysed, written and spelled verbally until correctly written first time, in five consecutive lessons. Word pairs which are commonly known to cause problems i.e. '*of/off*', '*how/who*' are worked on separately with the use of mnemonics and repeated reading and writing of the words until correctly spelled five consecutive times. This approach may be necessary with older children; however, it is possible that the approach to synthetic phonics used by Johnston and Watson (2004), where beginning readers were also taught to segment spoken words for spelling, may also be beneficial. Further research will be needed to examine this issue.

On the NARA, speed of reading remained unchanged for both groups across testing times, however, reading accuracy and reading comprehension improved significantly for the intervention group, whereas the non-intervention group maintained their initial levels of ability. The latter finding is particularly interesting as reading comprehension was not trained in the *Toe-by-Toe* programme. It has, however, been demonstrated that adequate reading comprehension depends upon a person already knowing between 90 and 95 percent of the words within a text (Nagy & Scott, 2000). The children in the intervention group did have better word recognition ability, as measured by both the NARA and the BAS-II. The gain in reading comprehension may thus have arisen because the intervention group knew more of the words, and also because they had a technique for decoding any unfamiliar words they encountered.

Another explanation for the increase in reading comprehension is that having a greater working memory capacity is positively related to increased reading comprehension and drawing inferences from text (Baddeley, 1992). In the current study, the intervention group did show a significant increase in working memory from pre-test to post-test as shown by the CTOPP digit span task. The improvement seen in reading comprehension could also be explained as a consequence of the improvement of non-word decoding skills. It has been suggested that once a child can sound-out nonsense words quickly and accurately, they can be said to have mastered the decoding process necessary for reading and therefore are able to free-up their working memory in order to concentrate on the comprehension of meaning (Kintsch, 1998). Kintsch further suggests that it typically takes several years of decoding practice before a child can comprehend a printed text as rapidly as they can process the same text when listening to it, signifying that fluent decoding is a necessary pre-requisite to comprehension. In support of this, the current study showed that the intervention group made significant improvements in accuracy levels from pre-test to post-test on the graded non-word reading test, showing improvements in their decoding skills, however, there were no significant improvements made on the listening comprehension task. This lack of improvement in listening comprehension, despite improvements in reading comprehension can be explained in that oral comprehension typically places an upper limit on reading comprehension; if a word is not recognised upon hearing it, it is unlikely that it will be recognised or comprehended upon reading it (Sticht, 1984; as cited in Hirsch, 2003). Therefore, in the current study, it is likely that the children were not reading to their full potential in terms of their ability to comprehend words at pre-test, but after the reading intervention, their reading ability, and therefore their reading comprehension, was closer to their full potential of comprehension ability.

As far as phoneme awareness is concerned, scores on the Elision task, the blending words task, and the blending non-words task revealed that the intervention group had made significant improvements across testing time, but the non-intervention group did not. The non-intervention group's performances on the non-word repetition task, and the segmenting non-words task, dropped significantly across testing time, whereas the intervention group significantly improved. The latter was predicted, as there is a close relationship between word reading skill and phoneme awareness (e.g. Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001; Hulme, Hatcher, Nation, Brown, Adams, & Stuart, 2002; Caravolas, Volín, & Hulme, 2005). However, the drop in performance by the non-intervention group could be due to a lack of attention and motivation within this group at post-testing on these rather difficult and abstract tasks. It is possible to suggest that during pre-testing both groups were motivated to perform to their best ability on these unusual tasks, and they were receiving one-to-one attention, a highly unusual and prized commodity within most secondary schools. It was noted by the researcher, however, that some members of the non-intervention group, were slightly more reluctant to complete these post-tests. This was partly due to the stigma associated with being seen by a reading teacher, and partly because they were often found out of class, roaming the corridors playing with their friends, and did not want to stop that to do tests.

Performances on the segmenting words task showed the non-intervention group to have maintained their ability scores, but that the intervention group made significant improvements across testing time. There were significant differences between the groups at pre-test, with the intervention group showing poorer segmentation ability than the non-intervention group, however, at post-test the intervention group significantly out-performed the non-intervention group. Both groups showed significant improvements in their non-word reading on the graded non-word reading task (Snowling, Stothard & McLean, 1996).

The intervention group significantly improved on the digit span test, although their scores remained in the low-average ability range. The non-intervention group, however, maintained scores within the average ability range. Researchers suggest that there is evidence of a close correspondence between reading and memory span (e.g. Johnston, Rugg & Scott, 1987). As digit span is a stable measure used in calculating IQs, however, it has always been supposed that memory span is a skill outside the reading system which can be used to support processes in reading. However, the synthetic phonic method used in this study asked children to sound and blend phonemes in order to find out how words were pronounced. In order to blend sounds together, the child would need to recall the phonemes in serial order (albeit with the letters representing the sounds in front of them). It seems quite possible that intensive practice of sounding and blending led to better performance in the auditory memory span task.

Overall it can be concluded that the intervention group showed dramatic improvements in their word reading accuracy, reading comprehension and phonemic awareness skills when compared to the non-intervention group. This can be explained as a direct consequence of the intervention programme, although it is not certain whether this is due to the synthetic phonics element of the training, or whether these skills would have improved by *Toe-by-Toe* alone. This training also led to an improved performance on the digit span task possibly due to regularly engaging in recalling phonemes in serial order during their reading instruction. It could therefore be suggested that before the intervention programme, the children in this group were not reading words via this approach, but that they were at post-test. It was therefore thought necessary to examine more closely the approach to reading taken by the children in both groups, via a regularity task and a non-word reading task (i.e. reading regular and irregular words), and to compare these findings with that of reading-age matched controls. This issue will be addressed in Chapter 6.

CHAPTER SIX

An Examination of Reading Strategies of Poor Readers: Do They Change With Synthetic Phonics Teaching?

In this chapter, an examination is made of the approach to reading taken by the poor readers in the previous chapter, to see whether synthetic phonics teaching led to any changes in the approach taken by the intervention group.

According to the dual-route theory, words are read in one of two ways, either by phonological recoding or by sight (Coltheart, 1978, in Underwood, 1978). Phonological recoding involves translating letters into sounds by applying letter-sound rules and then identifying the word from its pronunciation. Sight-word reading involves creating direct connections between the visual appearance of the printed words and their particular meanings in memory, obtained through repeated readings of these words. This theory assumes that these two processes are separate, but that sight-word reading of a word occurs after it has been recoded several times. Regular words are said to be read faster and more accurately than irregular words, as regular words can be read by either route, whereas irregular words can only be read by the visual route.

Most studies find poor readers to show a normal regularity effect (e.g. Waters, Seidenberg & Bruck, 1984; Metsala, Stanovich & Brown, 1998), however, a few studies have found that poor readers read high frequency irregular words as well as or better than high frequency regular words (e.g. Johnston, Perrett, Anderson & Holligan, 1990; Johnston and Morrison, 2007). The typical pattern of showing normal regularity effects suggests that poor readers generally use the same approach to reading as reading age controls, that is, that they take a phonological route to reading (at least with low frequency words). The prevailing view, however, is that poor readers show phonological deficits in reading, as evidenced by the non-word reading problems often found in comparison with controls (Rack, Snowling & Olsen, 1992). Indeed, one of the most direct ways of measuring an individual's decoding ability is by asking them to read non-words that they have not encountered before, e.g. 'tegwop' or 'balras'. Researchers have found that poor readers often show non-word reading deficits when compared to both chronological age (CA) matched controls, and reading age (RA) matched controls (e.g. Elbro, Nielson, & Petersen, 1994; Snowling, Goulandris & Defty, 1996), indicative of impaired phonological decoding skills. Research by Van Ijzendoorn and Bus (1994) supports this statement with findings from a meta-analysis of 16 studies, showing that poor readers had a significant non-word reading deficit. They also found that variation in IQ between poor and normal readers in different studies was a factor in determining whether a group difference was found, with those with more severe reading difficulties having greater difficulty with non-word reading. Johnston and Morrison (2007) have now shown only high IQ poor readers to have a non-word reading deficit, and low IQ poor readers to have reading age appropriate levels of non-word reading ability. This would imply that poor readers might adopt different reading strategies with non-words depending upon their IQ level.

This poses problems for the dual route model, as the regularity effect and non-word reading ability are both said to be a product of the phonological route. Thus, finding that poor readers show normal regularity effects but impaired non-word naming is problematical for this model.

Connectionist theorists disagree with the dual-route theory and have instead produced models that are based on a 'triangle' model of reading (e.g. Seidenberg & McClelland, 1989). Similar to the dual-route theory, the triangle model also suggests two pathways between the written and spoken word; the first being a pathway mapping directly from orthography to phonology, and the second being a pathway which maps from orthography to phonology via semantics. There are, however, differences between the two. In the triangle model there is no explicit distinction between the types of words processed by each pathway, whereas the dual-route theory suggests that irregular items can only be read by the lexical route and non-words by the phonological route. The dual-route theory also suggests that different types of processing underlie the lexical and sub-lexical routes, whereas it is suggested that a single mechanism underlies all processing in the triangle model. Connectionist models suggest that the connection strengths for patterns that occur frequently are greater than for those that occur less frequently, therefore giving regularity and consistency effects in reading.

In terms of the developing reader, however, two stage theorists have produced models that adopt features of the dual-route model (Frith, 1985) or the connectionist models (Ehri, 2004), and both models state that reading skills are learned via a developmental series of steps, with new strategies introduced at different points in the sequence. Both models hold that their individual phases are completed in order and that no phase can be omitted.

The first stage of Frith's (1985) theory is the 'logographic' stage, where readers can instantly recognise familiar words, but where letter sounds are ignored and all connections between the printed word and its meaning are arbitrarily related. The second stage of Frith's theory is the 'alphabetic' stage, where readers apply their knowledge of letter-sound correspondences in order to read words. The first three stages of Ehri's (2004) model can be closely matched with the first two stages of Frith's theory. The first stage of Ehri's model is the 'pre-alphabetic' phase, where readers use visual cues to recognise words. Ehri's second phase of reading development is the 'partial alphabetic' phase (which does not occur in Frith's theory), where readers begin to relate letter-sounds to printed words, although they do not look at all the letters in a printed word in serial order. Readers in this stage often pay close attention to the initial and end letters of words only. The third stage of Ehri's model is the 'full alphabetic stage' where readers are able to make connections between letters and sounds all through the word. Thus children having difficulty in non-word reading would be showing arrest at Stage 1 in Frith's model, but might be showing arrest at Stage 2 of Ehri's model. The latter view is a better fit to the data, as children in Stage 1 of Frith's model should show no non-word reading ability, but would show rudimentary skills if at Stage 2 of Ehri's model.

The final stages of these two theories are very different in their views of how fluent readers can sight-read familiar words. Frith's (1985) final stage is the 'orthographic' stage, where readers can instantly analyse words into orthographic units without phonological conversion. Ehri's (2004) final stage, the 'consolidated alphabetic' phase, proposes that whilst in the second and third stages, the systematic visual-phonological connections between letters seen in words and their pronunciations were only partial; these connections are completed in this final stage. Readers in this phase also fully analyse spellings to determine phonemic features in pronunciation. Therefore, according to Ehri, sight-word reading involves remembering systematic connections between spellings and pronunciations of words.

If, however, readers have difficulty in reading novel words, or non-words, dual-route theorists would infer that this would slow down the creation of a sight vocabulary, thus compromising their ability to develop word recognition skills. This would therefore force these readers to learn all words via the visual route, thus eliminating any advantage for reading regular words, which could otherwise have been processed by either route. Connectionist theorists, however, would assume that a non-word reading deficit is a consequence of other limitations within the system, thus inhibiting the generalisation of spelling-sound correspondence rules. For example, knowing how to pronounce the 't' in '*tap*', will not necessarily help the poor reader to pronounce the '*t*' in '*cat*', where the '*t*' appears in a different position.

The underlying assumption of Ehri's (2004) theory is that systematic visual-phonological connections exist between spellings and their pronunciations. This implies that readers use all available systematic relations, and do not just memorise the entire form of the word. Frith's (1985) theory directly contrasts this, and assumes that the printed form of the word instantly accesses its meaning in memory, via internal representations of letter-by-letter detail, implying that the word form in its entirety is rote-memorized. Frith's theory would therefore propose that any evidence of poor readers reading irregular words better than regular words would be indicative of an arrest at stage one, possibly indicating that children displaying this pattern of reading are able to use logographic skills, but perhaps not alphabetic or orthographic skills. Ehri's (2004) model would imply, however, that this type of reading pattern is indicative of an arrest at stage two, suggesting that they are unable to fully phonologically recode words, thus also precluding them from reading words by sight using the consolidated alphabetic strategy. Thus a synthetic phonics intervention should increase the ability to read by Stage 2 processes in the Frith model, or Stage 3 processes in the Ehri model, reducing any tendency to read irregular words better than regular words, and might also increase non-word reading ability.

In the current study, the two groups of poor readers, and matched reading age controls, carried out regularity and non-word reading tasks. It was predicted that, according to the prevailing phonological deficit view, the poor readers at pre-test would show evidence of phonological reading problems by being impaired for reading age in non-word reading ability; additionally some poor readers might show better reading of irregular than regular words. It was alternatively predicted, however, that as the poor readers in this study were of low IQ, that their non-word reading might be found to be reading age appropriate. It was predicted that the children in the synthetic phonics intervention at post-test would show increased non-word reading ability compared to the non-intervention children. Furthermore, poor readers showing better reading of irregular than regular words might also show a more normal pattern of performance after the synthetic phonics intervention.

Participants

Poor Readers

These were as in Experiment 5. Ten children were in the non-intervention group and ten children were in the intervention group as previously described (Chapter 5). After the pre-tests were conducted (reported below), one case, XP, was found to be of particular interest, and his progress will be charted individually. He read irregular words much better than regular words (e.g. reading regular words such as 'pest' as 'pets', 'kept' as 'keept', and 'strong' as 'storing') and his scores on the WISC block design and the WISC vocabulary sub-tests were very disparate (WISC block design, standardised score = 80 and WISC Vocabulary, standardised score = 55). It was thought possible that this pattern of reading performance was due to the way in which he was taught to read; he confirmed that when he came across an unfamiliar word in junior school, he was first asked to guess the word, and if this failed, he was asked to sound the word out and then to see what it sounded like. XP would guess unfamiliar words before attempting to sound and blend them; he also required prompting before sounding and blending, as he would not do so spontaneously. Where possible he would opt for reading irregular words before reading regular words, and complained when asked to read nonsense words. In July 2004, XP was registered as SEN status P (School Action Plus) meaning that he was entitled to additional support or teaching at school, however, this status was removed in October 2006. He is now classed as having a moderate learning difficulty, but no details are available.

Reading Age Matched Controls

Forty-four junior school children took part in this study as controls for the intervention and non-intervention groups. They were split into four groups, with eleven pupils in each, to form two pre-test control groups and two post-test control groups, each matched in terms of reading age and WISC Block Design Scaled Scores to the intervention group or to the non-intervention group.

Procedure

The intervention and non-intervention groups were tested both before and after the *Toe-by-Toe* programme had been conducted. The four reading age control groups, however, were tested only once.

Materials

Reading strategies were tested with a word reading task, and a non-word reading task. Seventy-two unrelated mono-syllabic words of similar length (see appendix 12) (adapted from Seidenberg, Waters, Barnes & Tanenhaus, 1984) were presented to participants. Eighteen were high frequency regular words, 18 were high frequency irregular words, 18 were low frequency regular words, and 18 were low frequency irregular words (see Table 23 for mean word length and frequencies). A repeated measures ANOVA was carried out. There was a main effect of frequency (as taken from Kuĉera and Francis (1967), cited in MRC Psycholinguistic Database 2008) ($F(1,17) = 22.354$, $p<0.001$) but no main effect of word type, ($F(1, 17) = <1$) and no interaction, ($F(1, 17) = <1$) showing that the matching had been effective. The computer used was a Dell laptop with a 60hz refresh rate. The screen was 14 inches, with a resolution of 1024x768. The program used was MemWord version 1.4. Each word appeared individually in the centre of the screen. Participants were asked to pronounce each word as quickly and as accurately as possible. Once pronunciation began, the item disappeared from the screen. Accuracy was recorded, as was the actual pronunciation given; the reaction times for each correct response were also recorded.

Table 23:

Mean Word Length Frequencies for Stimuli Used in the Regularity Task

	High Frequency		Low Frequency	
	Regular	Irregular	Regular	Irregular
Mean Frequency	358.22	504.5	26.61	31.94
SD	251.13	714.64	23.60	29.45
Mean Length	4.28	4.17	4.22	4.44
SD	0.73	0.5	0.53	0.60

A non-word reading task consisting of 42 one and two-syllable non-words (see appendix 13), (taken from Johnston & Morrison, 2007) was administered. The computer used was a Dell laptop with a 60hz refresh rate. The screen was 14 inches, with a resolution of 1024x768. The program used was MemWord version 1.4. Each non-word appeared individually in the centre of the screen, and disappeared the moment the participant began to speak into the microphone. Responses were carefully recorded phonetically by an experimenter. Accuracy was recorded, as was the actual pronunciation given; the reaction times for each correct response were also recorded.

Results:

Pre-Test:

Intervention Group: Pre-test Reading Age Matched Controls

Eleven analytic-phonics taught junior school children aged between 6.9 and 8.8 years (Mean = 8.02, SD = 0.60), reading no more than six months behind their chronological age, and no more than ten months ahead (reading age ranged between 7.6 and 9.2 years, Mean = 8.28, SD = 0.53), took part in this experiment as controls for the intervention group. They were tested and matched at pre-test on scaled scores on the WISC Block design task (Mean = 8.64, SD = 2.09) (see Table 24 for means and standard deviations). WISC vocabulary scores were not used as a basis for matching due to the poor scores achieved by the poor readers, as their poor performance might be due to a lack of exposure to an extensive vocabulary, and not due to a lack of ability to learn.

Non-Intervention Group: Pre-test Reading Age Matched Controls

Eleven analytic phonics taught junior school children aged between 7.1and 9.3 years (Mean = 8.15, SD = 0.63), reading no more than six months behind their chronological age (reading age ranged between 7.6 and 9.3 years, Mean = 8.45, SD = 0.46), and matched to the poor readers' non-intervention group's reading ages at pre-test, took part in this experiment. They were tested and matched to the poor-readers at pre-test on scaled scores on the WISC Block design task (Mean = 8.18, SD = 2.32) (see Table 24 for means and standard deviations). WISC vocabulary scores were not used as a basis for matching due to the poor scores achieved by the poor readers, as their poor performance might be due to a lack of exposure to an extensive vocabulary, and not due to a lack of ability to learn.

Table 24:

Chronological Age, Reading Age and Block Design Scaled Scores for Poor Readers (XP, Intervention Group and Non-Intervention Group) and Reading Age Matched Controls, at Pre-Test

	XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Chronological Age	12.58	12.10	0.38	8.02	0.60	12.16	0.31	8.15	0.63
Reading Age	7.50	8.58	1.08	8.28	0.53	8.36	0.96	8.45	0.46
WISC Block Design Scaled Score	6.00	7.1	2.51	8.64	2.09	6.30	1.95	8.18	2.32

One way ANOVAs were carried out on all test scores to compare the intervention group, the intervention reading age matched controls, the non-intervention group and the non-intervention reading age matched controls to ensure there were no significant differences in ability prior to the intervention. There were no significant differences between the groups in terms of WISC Block Design scaled scores ($F(3, 38) = 2.53$, $p>0.05$), or reading age ($F(3, 41) = <1$).

Regularity Task Accuracy

Newman Keuls post hoc tests, where reported, are significant at the .05 level.

A three-way repeated measures ANOVA was carried out on the word reading task accuracy levels, with frequency (high and low) and regularity (regular and irregular) as the within subjects factors and groups (intervention group, intervention reading age matched controls, non-intervention group and non-intervention group reading age matched controls) as the between subjects factor. There was no significant main effect of group, ($F(1, 19 = <1)$. There was a significant main effect of frequency ($F(1, 38) = 229.92$, $p<0.001$) with accuracy levels on high frequency words (87.12%, SD = 11.15%) being significantly higher than accuracy levels on low frequency words (63.39%, SD = 15.73%). There was no significant interaction between frequency and group ($F(3, 38) = <1$). There was a significant main effect of word type ($F(1, 38) = 77.62$, $p<0.001$) with accuracy levels being significantly higher on regular words (83.13%, SD = 13.88%) than irregular words (67.36%, SD = 13.00%). There was no interaction between word type and group ($F(3, 38) =1.07$, $p>0.05$). There was a significant interaction between frequency and word type $F(1, 38) = 41.28$, $p<0.001$). Newman Keuls post-hoc tests showed a significant regularity effect for low frequency words ($p<0.05$). There was no significant interaction between frequency, word type and group ($F(3, 38) = 1.323$, $p>0.05$). See Figure 24 for mean accuracy levels split by word type and word frequencies with standard deviation bars; see Table 25 for means and standard deviations.

Table 25

Mean accuracy levels for regularity task at pre-test

			XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Testing Time	Word Frequency	Word Type	Mean	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pre-Test	High	Regular	61.11%	83.89%	13.21%	92.19%	07.45%	91.67%	10.23%	91.39%	07.99%
		Irregular	88.89%	86.11%	16.41%	86.35%	11.27%	81.11%	11.17%	83.65%	10.54%
	Low	Regular	61.11%	77.22%	19.32%	74.66%	22.67%	71.67%	15.37%	81.65%	12.58%
		Irregular	72.22%	50.56%	17.06%	54.14%	13.35%	47.22%	14.41%	49.34%	11.27%

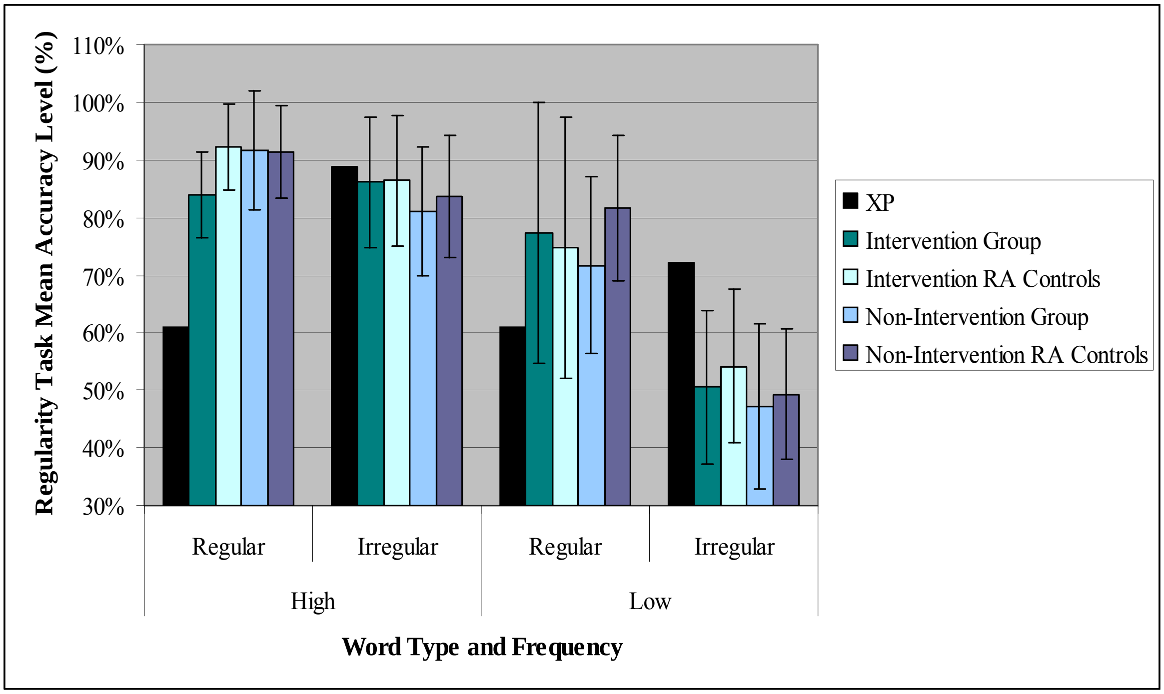


Figure 24: *Mean accuracy levels of regularity task split by word type and word frequency at pre-test*

Regularity Task Reaction Times

A three-way repeated measures ANOVA was carried out on the word reading task median reaction times to correct responses, with frequency (high and low) and regularity (regular and irregular) as the within subjects factors and groups (intervention group and intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was no significant main effect of group (F(1, 38 = <1). There was a significant main effect of frequency (F(1, 38) = 15.50, p<0.001) with reaction times on high frequency words (1007.52ms, SD = 424.88ms) being significantly faster than reaction times on low frequency words (1280.48ms, SD = 838.87ms). There was no significant interaction between frequency and group (F(3, 38) = <1). There was no significant main effect of word type (F(1, 38) = <1). There was no interaction between word type and group (F(3, 38) = <1). There was no significant interaction between frequency and word type F(1,

38) = <1). There was no significant interaction between frequency, word type and group (F(3, 38) = <1). See Table 26 for means and standard deviations.

Table 26

Mean reaction times for regularity task at pre-test

		XP		Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Word Frequency	Word Type	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
High	Regular	0870.0	124.24	1115.65	576.89	944.68	305.04	1030.75	479.35	882.23	310.64
	Irregular	0966.0	285.83	1086.05	518.72	1008.09	362.87	1094.85	540.38	924.86	308.95
Low	Regular	1017.0	580.92	1478.05	965.73	1184.77	658.52	1221.40	731.55	1218.73	848.45
	Irregular	1326.0	240.00	1536.55	1344.91	1202.59	704.79	1336.60	891.40	1106.09	560.45

Non-Word Task Accuracy: Short Non-Words and Long Non-Words

A two-way repeated measures ANOVA was carried out on the non-word reading task accuracy levels with non-word length (one syllable and two syllable) as the within subjects factor and groups (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was no significant main effect of group, ($F(3,38) = <1$). There was a significant main effect of non-word length ($F(1, 38) = 86.39$, $p<0.001$), with accuracy levels on short non-words (76.26%, SD = 16.66%), being significantly higher than on long non-words (47.33%, SD = 25.36%). There was no significant interaction between word length and group, ($F(3, 38) = 2.76$, $p>0.05$). See Table 27 for means and standard deviations.

Table 27:

Non-word accuracy at pre-test

	XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Non-Word Length	Mean	Mean	SD	Mean	SD	Mean	SD	Mean	SD
One Syllable	40.0%	78.00%	18.44%	72.99%	18.64%	76.00%	14.30%	78.18%	16.77%
Two Syllable	20.0%	55.50%	29.10%	41.04%	21.25%	57.50%	25.63%	36.94%	22.33%

Non-Word Task Reaction Times: One Syllable Non-Words and Two Syllable Non-Words

A two-way repeated measures ANOVA was carried out on the non-word reading task median reaction times to correct responses, with non-word length (one syllable and two syllable) as the within subjects factor and group (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was no significant main effect of group ($F(3, 38) = <1$). There was a significant main effect of word length on performance ($F(1, 38) = 32.39$, $p<0.001$), with reaction times for short non-words (1492.38ms, SD = 759.46ms) being significantly faster than reaction times to long non-words (2938.60ms, SD = 2215.14ms). There was no significant interaction between word length and group ($F(3, 38) = <1$). See Table 28 for means and standard deviations.

Table 28:

Non-word reaction times pre-test

	XP		Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Non-Word Length	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
One Syllable	1263.50	650.59	1494.80	892.45	1418.32	651.59	1511.90	720.44	1546.50	864.07
Two Syllable	2394.00	1068.31	2826.80	2237.07	3076.73	2499.99	3132.80	2236.40	2725.55	2177.52

Post-Test**:**

Intervention Group: Post-test Reading Age Matched Controls

Eleven analytic-phonics taught junior school children aged between 9.4 and 11.3 years (Mean = 10.73, SD = 0.62), reading no more than six months behind their chronological age, and no more than ten months ahead (reading age ranged between 9.8 and 11.8 years, Mean = 10.91, SD = 0.77), took part in this experiment as controls for the intervention group. They were tested and matched at post-test on scaled scores on the WISC Block design task (Mean = 8.18, SD = 0.98) (see Table 29 for means and standard deviations). WISC vocabulary scores were not used as a basis for matching due to the poor scores achieved by the poor readers, as their poor performance might be due to a lack of exposure to an extensive vocabulary, and not due to a lack of ability to learn.

Non-Intervention Group: Post-test Reading Age Matched Controls

Eleven analytic phonics taught junior school children aged between 7.1and 9.3 years (Mean = 8.25, SD = 0.59), reading no more than six months behind their chronological age (reading age ranged between 7.6 and 9.3 years, Mean = 8.41, SD = 0.98), and matched to the poor readers' non-intervention group's reading ages at post-test, took part in this experiment. They were tested and matched to the poor-readers at post-test on scaled scores on the WISC Block design task (Mean = 8.36, SD = 1.96) (see Table 29 for means and standard deviations). WISC vocabulary scores were not used as a basis for matching due to the poor scores achieved by the poor readers, as their poor performance might be due to a lack of exposure to an extensive vocabulary, and not due to a lack of ability to learn.

Table 29:

Chronological Age, Reading Age and Block Design Scaled Scores for Poor Readers (XP, Intervention Group and Non-Intervention Group) and Reading Age Matched Controls, at Post-Test

	XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Chronological Age	13.75	13.14	0.43	10.73	0.62	13.32	.30	8.25	.059
Reading Age	8.33	10.57	2.00	10.91	0.77	8.86	1.48	8.41	0.56
WISC Block Design Scaled Score	7.00	7.60	1.71	8.18	0.98	6.50	2.32	8.36	1.96

One way ANOVAs were carried out on all test scores to compare the intervention group to the intervention reading age matched controls, to ensure there were no significant differences in ability after the intervention. There were no significant differences between the groups in terms of WISC Block Design scaled scores ($F(1, 20) = <1$), or reading age ($F(1, 20) = <1$).

One way ANOVAs were carried out on all test scores to compare the non-intervention group to the non-intervention reading age matched controls, to ensure there were no significant differences in ability after the intervention. There were no significant differences between the groups in terms of WISC Block Design scaled scores ($F(1, 20) = 3.971$, $p>0.05$), or reading age ($F(1, 20) = <1$).

Regularity Task Accuracy

Newman Keuls post hoc tests, where reported, are significant at the .05 level.

A three-way repeated measures ANOVA was carried out on the word reading task accuracy levels, with frequency (high and low) and regularity (regular and irregular) as the within subjects factors and groups (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was a significant main effect of group, ($F(3, 38 = 9.90$, $p<0.001$). Newman Keuls post-hoc tests showed that the intervention group reading age controls performed significantly more accurately than the non-intervention group ($p<0.05$) and the non-intervention group reading age matched controls ($p<0.05$). The intervention group also performed more accurately than the non-intervention group ($p<0.05$) and the non-intervention group reading aged matched controls ($p<0.05$). There was a significant main effect of frequency ($F(1, 38) = 112.18$, $p<0.001$) with accuracy levels on high frequency words (91.26%, SD = 10.36%) being significantly higher than accuracy levels on low frequency words (74.85%, SD = 18.55%). There was a significant interaction between frequency and group ($F(3, 38) = 4.055$, $p<0.05$). Newman Keuls post-hoc tests showed that on high frequency words, the intervention group reading aged matched controls performed significantly more accurately then the non-intervention group ($p<0.05$) and the non-intervention group reading aged matched controls ($p<0.05$). On low frequency words, the intervention group reading aged matched controls performed significantly more accurately then the intervention group($p<0.05$), the non-intervention group ($p<0.05$)and the non-intervention group reading g aged matched controls($p<0.05$). There was a significant main effect of word type ($F(1, 38) = 92.98$, $p<0.001$) with accuracy levels being significantly higher on regular words (90.41%, SD = 12.62%) than irregular words (75.70%, SD = 16.28%). There was no interaction between word type and group ($F(3, 38) = 2.60$, $p>0.05$). There was a significant interaction between frequency and word type $F(1, 38) = 26.50$, $p<0.001$). Newman Keuls tests showed a significant regularity effect for both high and low frequency words, but there was clearly a much greater advantage for regular words over irregular words with the latter. There was no significant interaction between frequency, word type and group ($F(3, 38) = <1$). See Figure 25 for mean accuracy levels split by word type and word frequencies with standard deviation bars; see Table 30 for means and standard deviations.

Table 30

Mean accuracy levels for regularity task at post-test

		XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Word Frequency	Word Type	Mean	Mean	SD	Mean	SD	Mean	SD	Mean	SD
High	Regular	94.44%	94.44%	10.14%	100.0%	00.00%	90.00%	08.20%	91.41%	09.18%
	Irregular	83.33%	92.78%	10.49%	96.10%	04.92%	86.67%	14.39%	78.56%	10.10%
Low	Regular	72.22%	90.00%	09.73%	98.71%	02.87%	79.44%	20.96%	78.56%	18.91%
	Irregular	66.67%	68.89%	18.37%	79.86%	18.58%	54.44%	16.73%	48.30%	12.17%

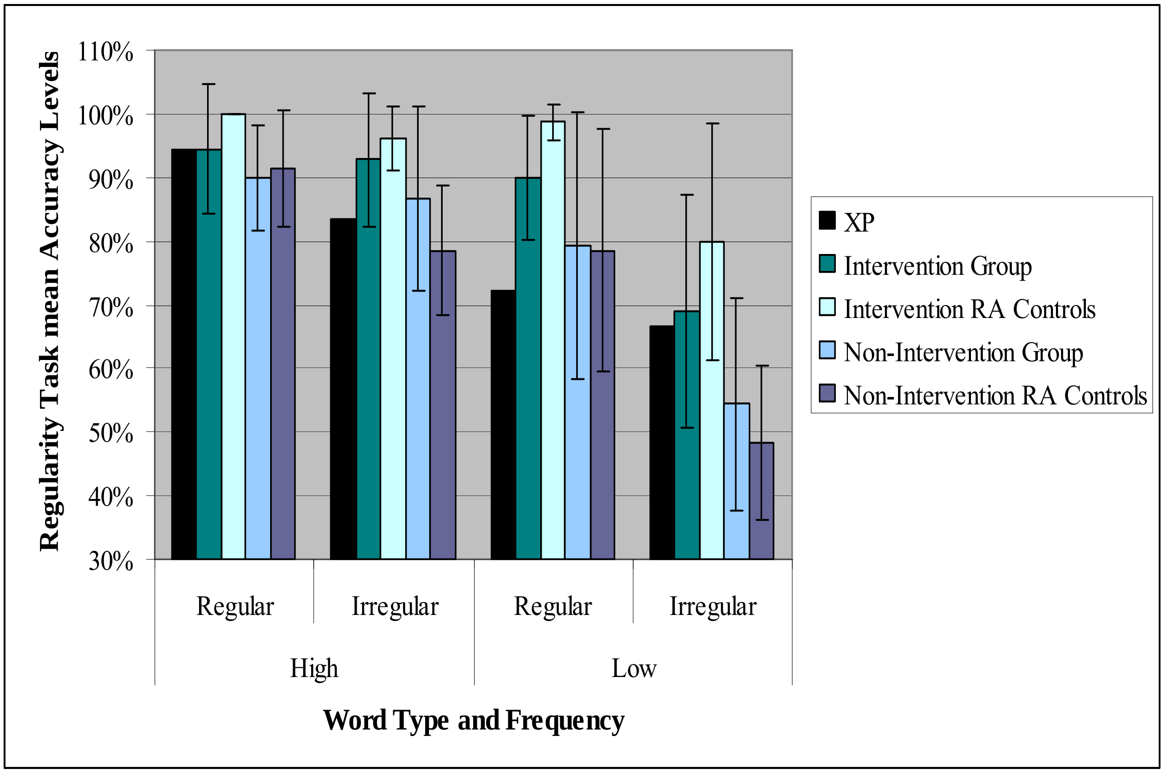


Figure 25: *Mean accuracy levels of regularity task split by word type and word frequency at post-test*

Regularity Task Reaction Times

A three-way repeated measures ANOVA was carried out on the word reading task median reaction times to correct responses, with frequency (high and low) and regularity (regular and irregular) as the within subjects factors and groups (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was no significant main effect of group ($F(3, 38 = 2.14, p>0.05)$. There was a significant main effect of frequency ($F(1, 38) = 14.15$, $p<0.001$) with reaction times on high frequency words (895.96ms, SD = 340.29ms) being significantly faster than reaction times on low frequency words (1052.05ms, SD = 580.56ms). There was no significant interaction between frequency and group ($F(3, 38) = 1.60$, $p>0.05$). There was no significant main effect of word type ($F(1, 38) = <1$). There was no interaction between word type and group ($F(3, 38) = 1.581$, $p>0.05$). There was no significant interaction between frequency and word type $F(1, 38) = 2.602$, $p>0.05$). There was no significant interaction between frequency, word type and group ($F(3, 38) = 1.20$, $p>0.05$). See Table 31 for means and standard deviations.

Table 31

Mean reaction times for regularity task at post-test

		XP		Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Word Frequency	Word Type	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
High	Regular	1006.0	392.69	889.50	344.30	679.50	139.42	906.30	255.10	998.82	366.50
	Irregular	1000.0	611.75	1023.20	628.17	729.18	138.09	934.25	260.28	1022.36	289.81
Low	Regular	1194.0	563.88	1137.15	677.70	690.27	127.75	1080.35	374.33	1347.09	879.45
	Irregular	1097.50	957.34	1236.05	853.15	744.36	124.62	1059.30	336.65	1149.55	528.47

Non-Word Task Accuracy: Short Non-Words and Long Non-Words

A two-way repeated measures ANOVA was carried out on the non-word reading task accuracy levels non-word length (one syllable and two syllable) as the within subjects factor and groups (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was a significant main effect of group, ($F(3,38) = 11.57$, $p<0.001$). Newman Keuls tests showed that the non-intervention group reading age controls performed significantly less accurately than the non-intervention group ($p<0.05$), the intervention group ($p<0.05$) and the intervention group reading aged matched controls ($p<0.05$). There was a significant main effect of non-word length ($F(1, 38) = 32.85$, $p<0.001$), with accuracy levels on short non-words (85.19%, SD = 16.20%), being significantly higher than on long non-words (69.79%, SD = 25.11%). There was a significant interaction between word length and group, ($F(3, 38) = 6.58$, $p<0.001$). Newman Keuls tests showed that the non-intervention group reading age controls performed significantly less accurately on long non-words than the non-intervention group ($p<0.05$), the intervention group ($p<0.05$) and the intervention group reading aged matched controls ($p<0.05$). See Table 32 for means and standard deviations, see Figure 26 for mean accuracy scores at pre-test and post-test with standard deviation bars.

Table 32:

Non-word accuracy at post-test

	XP	Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Non-Word Length	Mean	Mean	SD	Mean	SD	Mean	SD	Mean	SD
One Syllable	90.0%	94.50%	03.69%	90.91%	09.70%	80.50%	21.66%	75.26%	17.15%
Two Syllable	70.0%	85.50%	11.41%	82.08%	11.47%	73.00%	22.26%	40.29%	22.56%

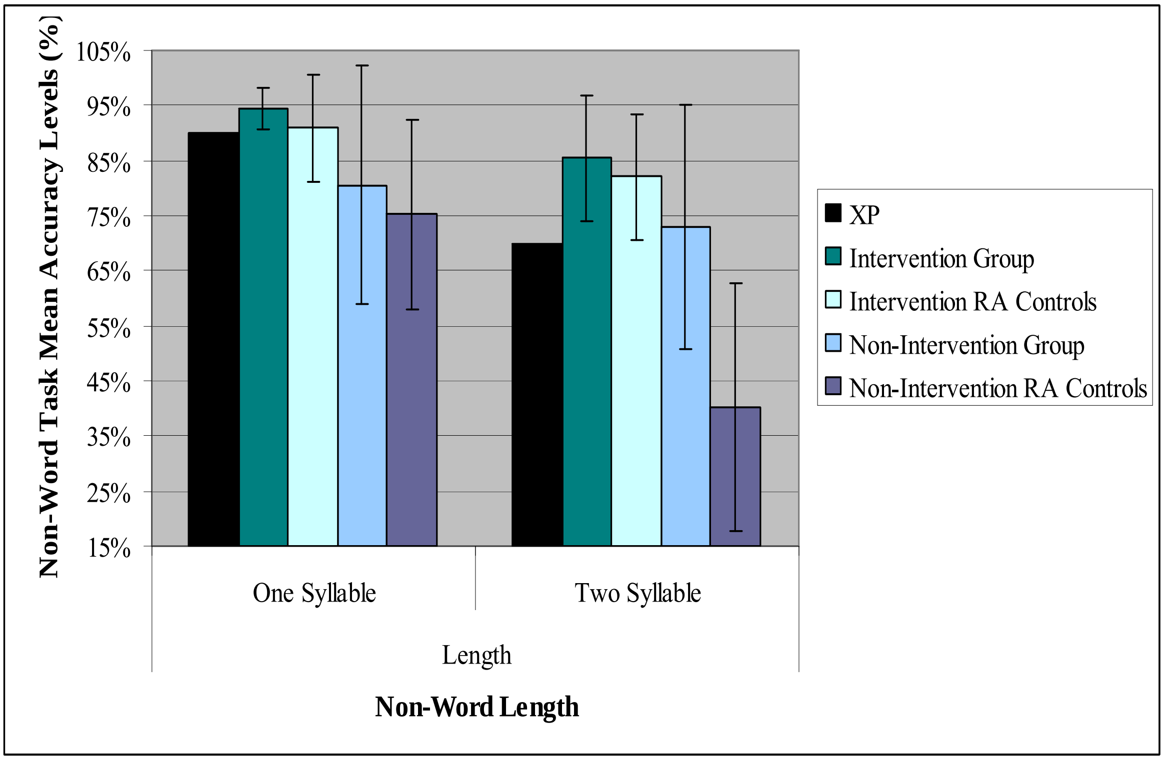


Figure 26: *Post-test non-word task overall accuracy levels by non-word length for intervention and non-intervention groups*

Non-Word Task Reaction Times: One Syllable Non-Words and Two Syllable Non-Words

A two-way repeated measures ANOVA was carried out on the non-word reading task median reaction times to correct responses, with non-word length (one syllable and two syllable) as the within subjects factor and group (intervention group, intervention reading age matched controls, non-intervention group and non-intervention reading age matched controls) as the between subjects factor. There was a significant main effect of group ($F(3, 38) = 3.53$, $p>0.05$). Newman Keuls post-hoc tests showed that the non-intervention reading age controls responded significantly slower then the non-intervention group ($p<0.05$), the intervention group ($p<0.05$) and the intervention group reading age controls ($p<0.05$). There was a significant main effect of word length on performance ($F(1, 38) = 37.30$, $p<0.001$), with reaction times for short non-words (1160.01ms, SD = 597.78ms) being significantly faster than reaction times to long non-words (1917.19ms, SD = 1348.38ms). There was no significant interaction between word length and group ($F(3, 38) = 2.63$, $p>0.05$). See Table 33 for means and standard deviations.

Table 33:

Non-word reaction times at post-test

	XP		Intervention Group (Including XP) (N = 10)		Intervention Group Reading Age Controls (N = 11)		Non-Intervention Group (N = 10)		Non-Intervention Group Reading Age Controls (N = 11)	
Non-Word Length	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
One Syllable	1095.0	1224.11	1089.60	446.42	885.36	235.48	1050.90	454.77	1597.86	0856.41
Two Syllable	2498.0	1151.8	1959.45	967.52	1236.14	656.92	1600.10	805.39	2848.09	2007.67

Discussion

The reading strategies of poor readers were examined using a regularity task and compared to reading age matched controls.

At pre-test, all four groups showed significantly higher accuracy levels to high frequency words than low frequency words, and all groups showed a regularity effect, however, this regularity effect was stronger to low frequency words. Reaction times were significantly faster to high frequency than low frequency words. There were no significant differences between the groups in terms of accuracy or reaction times.

At post-test, however, both the intervention group and the intervention group reading age controls performed significantly more accurately overall than the non-intervention group and the non-intervention group reading age matched controls. For low frequency words the intervention group reading aged matched controls out-performed all other groups, whereas on high frequency words the intervention group reading aged matched controls did not perform significantly differently to the intervention group, but did perform significantly more accurately than the non-intervention group and the non-intervention group reading age matched controls.

The non-word reading abilities of poor readers were also compared to reading age matched controls. At pre-test all four groups were performing at a similar level of non-word reading accuracy, however this was not the case at post-test, where the non-intervention reading age matched controls performed slower and less accurately than all other groups, and less accurately than all other groups on long non-words.

At pre-test, XP showed a similar pattern of performance on the non-word task to the rest of his group, in that he responded more accurately to short non-words than to long non-words; however, his non-word reading was the worst in his group. At post-test however, his accuracy levels were more similar to the others in his group, with one child scoring lower than him on long non-word accuracy.

In terms of the regularity task, at pre-test, the poor reader intervention group read high frequency irregular words a little better than regular words, however, most children from both poor and normal reader samples showed a strong regularity effect, reading regular words with higher accuracy than irregular words. One child from the non-intervention reading age matched controls showed a tendency to read high frequency irregular words better than high frequency regular words (7.2% difference). Two children from the intervention reading age matched controls also showed a tendency to read high frequency irregular words better than high frequency regular words (7.2% difference), with one of these two children also showing this trend for low frequency words (14.3% difference).

Of the poor readers, five children tended to read high frequency irregular words better than high frequency regular words (16.67% difference). And one of these five children, XP, also read low frequency irregular words with 10% better accuracy than low frequency regular words. Purely due to chance, all of the five children were in the intervention group. Overall, however, there was no evidence that poor readers were more prone to better reading of irregular words than controls.

At post-test, the children from all four samples (intervention group, intervention group controls, non-intervention group and non-intervention group controls) still showed a strong regularity effect, however, some individuals did not. Two children from the intervention group showed a slight tendency to read high frequency irregular words better than high frequency regular words, and another child showed this trend for low frequency words, with more than a ten percent difference in accuracy between low frequency irregular words and low frequency regular words. At pre-test there had been five children showing this pattern, however at post-test there was only three, all of which were in the original five. XP, however, showed a normal regularity effect in his accuracy levels, both for words of high and low frequency.

Only one child from the intervention group's reading aged matched controls showed a slight tendency to read low frequency irregular words better than low frequency regular words, however, there was none at pre-test showing this pattern. Two children from the non-intervention group showed a slight tendency to read high frequency irregular words better than high frequency regular words, and one of these children, plus one other child showed this trend for low frequency words. No children from the non-intervention group reading aged matched controls showed this trend

The majority of researchers have also found that poor readers tend to show normal regularity effects in their word reading accuracy on low frequency items, with smaller or no regularity effects found for high frequency items (e.g. Waters, Seidenberg & Bruck, 1984; Metsala, Stanovich & Brown, 1998). Although in the current study some poor readers did have an unusual pattern to begin with, the intervention seems to have led to a change of pattern of performance, especially for XP.

In contrast, however, there is a vast amount of research showing reading age matched controls to have better non-word reading than poor readers (e.g. Snowling, 1981; Snowling, Goulandris, & Defty, 1996; Van Ijzendoorn & Bus, 1994). It could therefore be considered highly unusual to find poor readers reading as well as reading age matched controls, perhaps indicating that the poor readers in the current study have no phonological reading deficits. Research by Johnston and Morrison (2007) however, found that only high IQ poor readers had a non-word reading deficit, and low IQ poor readers had reading age appropriate levels of non-word reading ability, thus implying that poor readers might adopt different reading strategies with non-words depending upon their IQ level. The current study supports these findings, as neither the intervention nor the non-intervention groups showed a non-word reading deficit.

Further, the non-intervention reading age matched control groups had similar reading ages at both pre- and post-testing, and there were no differences between the groups in their abilities to read words or non-words. The post-test intervention group reading aged matched controls however, were much more accurate at non-word reading than the pre-test group and had a reading age of about 31 months higher, showing that the higher levels of word reading skill are reflected in much better non-word reading.

Summary

Overall, the reading strategies of poor readers were examined using a regularity task and a non-word reading task, and compared to reading age matched controls. At pre-test, there was a slight, non-significant, tendency for the poor reader intervention group to read high frequency irregular words better than high frequency regular words, indicating, perhaps, a tendency for these children to read via the visual route. At post-test, however, this group showed a strong regularity effect for high and low frequency words. At pre-test, there was no significant difference in accuracy levels, or reaction times, between the poor readers and their reading age matched controls. At post-test, however, whilst both the intervention group and the intervention reading age matched controls showed large improvements the other two groups did not. XP showed a marked pattern of better reading of irregular than regular word reading, and very low non-word reading skills prior to the intervention. At post-test, however, XP was displaying a robust regularity effect.

In terms of non-word reading accuracy, there was no significant difference in accuracy levels, or reaction times, between the poor readers and their reading age matched controls at pre-test, but at post test, whilst the intervention group were still similar in accuracy to their reading aged matched controls, the non-intervention group, were significantly better than their reading aged matched controls. XP showed very low non-word reading abilities at pre-testing, but had improved dramatically at post-testing.

CHAPTER SEVEN

A Discussion of Poor Readers' Approaches to Recognising Printed Words

Throughout this thesis the reading strategies of poor readers have been investigated and examined. Despite the prevailing phonological deficit hypothesis of dyslexia, the poor readers partaking in this research have shown only mild phonological awareness problems. The poor readers from Chapters 4 and 5, however, were of low-average IQ, and it is possible that IQ plays an important role in the reading strategies used by, and impairments found in, those with reading difficulties. This, however, does not explain the reading problems faced by DB, the adult dyslexic case-study from Chapters 1, 2, and 3; although it could be argued that, due to his age, he had already found ways to compensate for his phonological deficits.

Phonological deficit hypothesis

Most recent definitions of dyslexia propose that it is a difficulty in reading caused by a core phonological deficit. The Orton Dyslexia Society of the USA (now the International Dyslexia Association) (1994; cited in Snowling, 2000, p24-25) states: "Dyslexia is … characterised by difficulties in single-word decoding, usually reflecting insufficient phonological processing abilities..."

Phoneme awareness is the ability to identify the smallest meaningful unit of sounds within spoken words, and is one aspect of phonological awareness. A deficit in phoneme awareness would lead to difficulties in learning phoneme-grapheme conversion rules, and thus would limit an individual's ability to read via a phonological route.

According to the dual-route model of reading, individuals with phonological reading problems should exhibit a smaller or non-existent regularity effect, and impaired non-word reading, compared to controls. Metsala, Stanovich and Brown (1998) conducted a meta-analysis of regularity effects to explore the inconvenient finding that while many studies show poor readers to have impaired non-word reading, regularity effects are often found (e.g. Olsen, Kliegel, Davidson & Foltz, 1985; Szeszulski & Manis, 1987). This, at first glance, is inconsistent with the theory of a deficit in phonological processing. Metsala *et al.* argue, however, that this is due to a difference between a reader being able to use a phonological strategy and the level of skill they possess in using that strategy. Rack, Snowling and Olsen (1992) also concluded that a dyslexic reader is more likely to have poorer phonological skills than a non-dyslexic, and that there will be differing contributions of phonological processes required depending upon the particular word or non-word presented, with irregular words requiring less phonological processing than regular words, and non-words requiring the highest level of phonological skill.

Other definitions, however, look at phonological deficits as only one possible symptom of dyslexia, and not necessarily as a causal factor, e.g. the British Dyslexia Association, who imply that other difficulties are evident in dyslexia, and all should be treated with equal importance in both diagnosis and treatment (Crisfield, 2002, p. 67). Additionally, different research suggests that there are neurological differences between dyslexic and non-dyslexic readers, with dyslexics showing less activation across the left-hemisphere (e.g. Shaywitz *et al.*, 1998).

In the current study, however, the adult dyslexic case-study DB (see Chapters 1, 2, and 3) did not show signs of a distinct phonological deficit. He did not present the pattern of performance expected from a phonological dyslexic, as although his non-word reading was mildly impaired he showed little sign of a phonemic awareness deficit on a phoneme deletion task or a graphophonemic awareness test when compared to university undergraduate controls, although he did read irregular words equally as well as regular words. It is possible, however, that DB had somehow learned to compensate for any phonological impairment he may have had, and that the tests used were not sufficiently probing enough to detect them. Indeed, Snowling (2000, p137) suggests that "it is rare to find a dyslexic child who does not have some kind of phonological problem if they are tested using sensitive enough measures." On a sensitive test of hemispheric processing, however, DB showed no signs of under-activation in the LH when compared to non-dyslexic controls, nor did he show over-reliance on RH processing. This is in direct contrast to the patterns shown by a developmental dyslexic in Lavidor, Johnston and Snowling's (2006) study.

The children with reading problems who took part in the reading intervention programme showed varying deficits in their skills relating to reading. Overall, before the intervention, the children showed below average scores on tests of general abilities, including vocabulary IQ, performance IQ and memory span. The pupils also displayed below average literacy skills, including poor reading and spelling abilities. On tests of phoneme awareness, the pupils also showed some difficulties with the Elision sub-test, and were in the low-average range on the phoneme reversal and segmenting words sub-tests. Elision tasks, however, may be cognitively difficult and their low performance on this task could be explained in terms of the pupil's generally low ability, as shown by their IQ scores. It could also be argued that because reading develops phoneme awareness, their elision skills might be appropriate for their reading age, if not for their chronological age. In support of this, their scores on other tests of phoneme awareness, such as blending the sounds in spoken words and non-words, and non-word repetition were within the average ability range for their chronological age. This therefore suggests that the pupils in the current study did not have marked phonological awareness deficits. The deficits found in their short-term memory spans, however, could be the underlying cause of their reading and spelling difficulties.

Researchers are divided over how to differentiate between a dyslexic reader and any other poor reader. It has been suggested that in order to do this, the relationship between reading skill and intelligence needs to be examined. It is generally accepted that there is a positive correlation between an individual's IQ score and their reading level (Snowling, 2000). A child is considered dyslexic if their reading ability is significantly below their expected reading ability, however, if a child has a reading age significantly below their chronological age, but not out of line with expectation due to their IQ being relatively low, (s)he is considered to have a general reading difficulty. Due to the reliance on the construct of IQ, a vague and imprecise term, this classification approach is not without problems. As IQ is a general measure of intelligence, testing both verbal IQ and performance IQ, it can be argued that if low verbal IQ can be due to reading disability (because of a lack of exposure to print, or due to lowered comprehension of written texts) it could also mask the specificity of a child's reading problem.

Stanovich (1991) proposed an alternative method of differentiating between general poor readers and those with specific reading difficulties. He suggested that it is a reasonable expectation that a child's listening comprehension will be at a similar level to their reading comprehension (providing that they can read). If this is true, then a child whose reading comprehension is below their listening comprehension could be described as having a specific reading difficulty.

Support for using IQ to distinguish between a dyslexic and a non-dyslexic poor reader, however, has been put forward by Johnston and Morrison (2007). They found that poor readers with high IQs (above 110) had difficulty using a phonological approach to reading, showing no advantage for high frequency regular words over high frequency irregular words, whereas low IQ poor readers (with IQ scores below 90) showed a more phonological approach. The high IQ poor readers had impaired non-word reading abilities, whereas the low IQ poor readers showed slower reading of non-words but similar accuracy levels when compared to reading-age controls. They therefore conclude that there is a continuum of severe to mild deficits in taking a phonological approach to reading that is associated with IQ levels, with those of higher IQ showing more severe phonological reading deficits.

In this thesis, a variance in the IQ scores amongst the poor readers who were studied has been shown, with DB showing a fairly high IQ and the children from the reading intervention study showing fairly low IQ scores. Additionally, the poor readers from the reading intervention study showed similar listening comprehension abilities to reading comprehension abilities at pre-test, but better reading comprehension than listening comprehension at post-test. DB was not tested on his listening comprehension; however, his reading comprehension was in the average range when compared to other university students, suggesting that his reading comprehension would be at a similar level to his listening comprehension. This therefore suggests that overall the poor readers involved in the current reading intervention study did not have a specific reading disability, and they do not fit the profile of 'classic developmental dyslexics'. They were, however, slightly slower and slightly poorer than reading age matched controls on non-word reading accuracy at pre-test, although neither speed nor accuracy differences reached statistical significance, thus providing some support for the findings of Johnston and Morrison (2007).

There were, however, four children in the intervention group, one of whom was XP, who showed better listening comprehension than reading comprehension at pre-test. At post-test, however, only two children from the intervention group, one of whom was XP, were showing this trend with the difference between their reading and listening comprehension standard scores being reduced.

It must be noted, however, that the methodology of the reading intervention study was not wholly robust, in that only a small sample size was employed, meaning that any conclusions drawn must be treated with caution. Also, although there was a group of unseen controls, there was no placebo control group participating. One future methodological improvement would be to include a control group who receive the *Toe-by-Toe* reading programme without any synthetic phonics teaching, so that the benefits of the synthetic phonics element can be better evaluated. It may also be prudent to include a group of controls who receive an equal amount of time, but who spend it being read to, without seeing the text. Taken together these two additional control groups would help to eliminate any placebo or 'Hawthorne' effects, i.e. the phenomena of behavioural changes due solely to behaviour being measured, or due to receiving additional attention. In the current study, however, it is unlikely that the cognitive improvements made by the intervention group, e.g. their increased memory span, were due to a Hawthorne effect. It is also unlikely that such a change would be predicted to occur with an unmodified version of *Toe-by-Toe*, as this would not emphasise serial processing skills.

Unfortunately, the very existence of a control group raises ethical issues, particularly in identifying a need but not addressing it. In the current study, the school was made aware of the children in the control group, and consequently they would therefore have been given the usual amount of assistance available, including classroom assistant support. In future studies, the methodological design should be altered to include a subsequent reading intervention for any control group participants.

Unfortunately, this was not included in the current study due to strict deadlines and a lack of resources.

Reading strategies of poor readers

The majority of the poor readers from Chapters 4 and 5 showed strong regularity effects in their reading, suggesting that they do not have phonological deficits, but that they tend to read printed words according to their letter-sound rules. Such a pattern of performance is not necessarily indicative of normal phonological reading skills, however, as normal regularity effects have also been found in studies of poor readers where deficits in non-word reading were found (Metsala, Stanovich & Brown, 1998).

It is interesting, however, that some of the poor readers in this study did not show strong regularity effects, in fact they read irregular words better than regular words, a pattern also shown by DB, suggesting that these readers utilised a visual approach to printed words. This strategy seemed to alter somewhat after a synthetic phonics intervention programme, resulting in some of these children showing a stronger regularity effect, with less reliance on a visual approach.

All five 'visual' readers showed improvements between their pre- and post-test accuracy levels for high frequency regular words, with XP showing a marked increase. On low frequency regular words, however, two children showed a drop in their accuracy levels, and the remaining three showed lower levels of improvement compared to high frequency words. Overall, four of the children showed normal regularity effects on low frequency words at post-testing, and only two showed a visual approach to high frequency words. XP was now showing normal regularity effects at post-testing on both high and low frequency words. These improvements in performance levels and the establishment of regularity effects for these children are likely to be a direct consequence of the use of synthetic phonics alongside the *Toe-by-Toe* reading programme.

The effectiveness of phonics teaching in developing normal patterns of performance for some poor readers can be accounted for by both dual route and connectionists models, and indeed by developmental models. Dual-route theory holds that regular words are found to be at an advantage over irregular words, as regular words can be read by either a phonological or a visual route, whereas irregular words can be read only via the visual route. Connectionist models, however, reject the notion that regular and irregular words are processed by individual sub-systems, and therefore would find it more difficult to explain the reading pattern displayed by the poor readers in this study.

The dual-route theory infers that if a child has difficulty learning and generalising letter-sound rules, this would negatively impact upon the development of word recognition skills. If these difficulties were severe, it could result in the children being forced to learn all words through the visual route, and consequently, these children would show no advantage for reading regular words over irregular words because the phonological route would remain unused (Snowling, 2000). Dual-route theory might suggest that such children should be taught to read via a phonics instructional programme, in order to develop the phonological route. Teaching children phonological recoding skills by phonically reading words which are of regular spelling and contain letter-sound relations may lead to the appearance of the normal advantage in reading regular versus irregular words, if the habit of reading visually is not completely entrenched.

Connectionist theorists do not subscribe to the belief, held by dual-route theorists, that regular and irregular words are processed by individual sub-systems. Instead, they argue that these models are sensitive to the statistical regularities intrinsic to the orthography of the words on which they are trained. Therefore, the connection strengths for patterns that occur frequently are greater than for those occurring less frequently, and this is the cause of regularity and consistency effects in reading. Systematic phonics teaching methods encourage children to convert letters (graphemes) into sounds (phonemes) and then to blend the sounds to form recognisable words, as is the case in synthetic phonics; or teaching children whole-words then encouraging them to analyse these into their component parts, as in analytic phonics. Both phonics methods therefore emphasise the relationships between printed letters and their sounds with an aim to facilitate the development of mappings between orthography and phonology. These methods, in terms of a connectionist model, therefore allow regularly occurring patterns to be 'stronger' than patterns occurring less frequently. As analytic phonics starts with establishing sight word reading, however, this early training using a visual approach not connected to letter sounds may lead to some children developing a rather visual approach to reading, as was found for some of the poor readers in the present study. Thus in connectionist terms, the synthetic phonics approach might be better at developing the mappings between orthography and phonology, as this connection is established early on.

The two stage-theories of Frith (1985) and Ehri (1992) also offer alternative explanations to the dual-route theory, although there are some similarities between them. Dual-route theorists suggest that the process used for reading familiar words, i.e. sight-word reading, is separate from the process used for reading novel words i.e. phonological recoding. Frith's (1985) theory argues that there are connections between the various phases of acquisition, occurring via a union of skills between the strategies, however, she does agree with dual-route theory that sight-reading of familiar words does not involve phonological recoding. Ehri's (1992) theory agrees that letter-sound relations are used initially to read an unfamiliar word; however, she argues that this process would also be used when sight-reading familiar words. Both models imply that systematic phonics approaches would be effective as they would develop the phonological route for reading, thus enabling the beginning reader to read novel words.

Summary and future directions

The findings in this thesis would appear to disagree with the prevailing phonological deficit hypothesis of dyslexia, which suggests that individuals with reading difficulties should show poor non-word reading and an absence of regularity effects. That view was always problematical, as so many studies find normal regularity effects in poor readers. This thesis has further shown that not all poor readers display marked phonological deficits, and yet show signs of a rather visual approach to reading. These findings do not support the view that a visual approach develops as a compensation for phonological problems (Snowling, 2000). The problem may lie with the way that contradictory findings have been ignored or explained away. For example, a meta-analysis by Metsala, Stanovich and Brown (1998), based upon 1,116 participants from 17 studies, 536 of whom had reading disabilities, found strong evidence against the absence of regularity effects in children with reading disabilities, instead finding robust word regularity effects both in those with, and those without, reading difficulties. This evidence against the phonological deficit hypothesis is then explained in six different ways, five of which are discarded by the authors as being most unlikely. The sixth account was accepted, with the regularity effects found explained away as a consequence of varying phonological demands placed upon the individual by different tasks, assuming that non-word reading demands more phonological skill than reading of regular and irregular words. This was explained in terms of connectionist models of reading whereby impaired phonological representations stopped the models being able to generalise from the letter ‘t’ in ‘tap’ to the letter ‘t’ in ‘pat’. This led to impaired non-word reading by the models, but preserved the regularity effects of word reading.

The findings of this thesis, however, show that some poor readers read visually, despite displaying no obvious signs of phonological impairment in skills directly relevant to the reading process. These children knew how to sound and blend, but did not apply these skills to word reading. It can therefore be argued that had they been taught to read, in the first instance, by a synthetic phonics method, they might not have become poor readers. The current data, therefore, do not support the idea of there being a clear-cut phonological disorder, with a visual approach to reading arising as a compensatory approach. One possible explanation for this disparity, also put forward by Rack, Snowling and Olsen (1992, p40), is that findings are published only when they fit the current ideology. The current data show, however, that a simplistic view of phonological disorders does not fit with reality.

Further research is needed to detect, or confirm, what the underlying cause of dyslexia is, as there are indications of it not being due to a phonological deficit in all cases. Any additional research undertaken should also consider expanding upon different age-groups, to include younger children, and utilising larger sample sizes to help clarify findings. It would also be beneficial to examine in more detail the benefits of interventions such as *Toe-by-Toe*, and how much additional remediation is added by synthetic phonics. Perhaps more investigation of neurological differences is also necessary, as although it is now generally accepted that the right (RH), as well as the left (LH), hemisphere contributes to the lexical processing in the normal brain (e.g. Chiarello, Shears, Liu, & Kacinik, 2005), it is less clear whether this is still the case for developmental dyslexics.

Lavidor, Johnston and Snowling (2006) found that a case-study dyslexic’s performance on a lexical decision task (LDT) revealed a greater reliance on RH orthographic processing strategies than normal readers, and poorer LH performance, suggesting less efficient reading processes in the LH. Several other researchers have also suggested that dyslexics show less activation across the left-hemisphere, for example, during rhyme processing and short-term memory tasks (Palesu *et al.*, 1996), reading aloud (Brunswick, McCrory, Price, Frith and Frith, 1999), and other tasks involving various levels of phonological demands, including single letter and non-word rhyme judgements (Shaywitz *et al.*, 1998). These findings are suggestive of dyslexia, for some individuals, being a specific form of language impairment affecting the way in which the brain encodes written and phonological information.

This thesis has shown that phonological awareness deficits may not be the only cause of specific reading difficulties, and that despite adequate phonological skills, some individuals still read visually. It has also shown synthetic phonics to be beneficial, not only to beginning readers in terms of reducing the amount of children who later develop reading difficulties (Johnston and Watson, 2005), but also as remediation for older readers who have already developed difficulties. This type of intervention may also help to break visual readers out of this strategy, and may encourage them to take a more phonological approach to reading, thus enabling easier reading of unfamiliar words.

REFERENCES

Australian Government. Department of Education, Science and Training (2005). *Teaching Reading .Report and Recommendations. National Enquiry into the Teaching of Literacy.* Barton: Australia: Department of Education, Science and Training.

Baddeley, A. (1992). Working memory. *Science*, **255**, pp. 556-559.

Baddeley, A.D., Emslie, H., & Nimmo-Smith, I. (1994). *Doors and People: A Test of Visual and Verbal Recall and Recognition*. Bury St. Edmunds, England: Thames Valley Test Co.

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, **26**, pp.49-66.

Berninger, V.W., Vermeulen, K., Abott, R.D., McCutchen, D., Cotton, S., Cude, J., Dorn, S., & Sharon, T. (2003). Comparison of three approaches to supplementary reading instruction for the low-achieving second-grade readers. *Language, Speech, and Hearing Services in Schools,* **34**, pp.101-116.

Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. *Developmental Medicine and Child Neurology,* **15**, pp.663-687.

Bradley, L., & Byrant, P.E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, **271**, pp.746-747.

Bradley, L., & Byrant, P.E. (1983). Categorising sounds and learning to read: A causal connection. *Nature*, **301**, pp. 419-421.

Bruck, M. (1990). Word recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology*, **26**, pp.439-454.

Brunswick, N., McCrory, E., Price, C.J., Frith, C.D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain*, **122**, pp.1901-1917.

Bryant, P.E. (1982). The role of conflict and of agreement between intellectual strategies in children's ideas about measurement. *British Journal of Psychology,* **73**, pp.243-251

Byrne, B. (1998). The foundation of literacy: The child's acquisition of the alphabetic principle. Hove: Psychology Press.

Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, **92**, 107-139

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, **91**, pp.77-111.

Chiarello, C., Shears, C., Liu, S., & Kacinik, N.A. (2005). Influence of word class proportion on cerebral asymmetries for high- and low- imagery words. *Brain and Cognition.* **57**, pp. 35-38.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.) *Strategies of Information Processing,* London: Academic Press, pp. 151-216.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. *Attention and Performance VI: The Psychology of Reading*. S. Dornic (Ed.). 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**, pp.204-256.

Crisfield, J. (2002). *The dyslexia handbook*. Reading: British Dyslexia Association.

Critchley, M. (1970). *The Dyslexic Child*. London: Heinemann Medical Books.

Cunningham, A. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, **50**, pp. 429-444.

Duncan, L.G., Colé, P., Seymour, P.H.K., & Magnan, A. (2006). Differing sequences of metaphonological development in French and English. *Journal of Child Language*, **33**, pp. 369-399.

Daryn, E. (2000). Lessons from the follow-up of developmental dyslexics. *Medical Hypotheses*, **54**, pp.434-437.

Department for Education and Employment (DfEE), (1998). *The National Literacy Strategy Framework for Teaching*. Sudbury: DfEE Publications.

Dixon, M., & Stuart, M. (2002). Phonological awareness improves word recall. *Literacy Today*, December 2002, p20.

Dombey, H. (2006). How should we teach children to read? *Books for Keeps*, **156**, pp. 6-7.

Ehri, L.C. (1992). Reconceptualising the development of sight word reading and its relationship to recoding. In Gough, P.B., Ehri, L.C., & Treiman, R., (Eds.) *Reading Acquisition*, London: Erlbaum.

Ehri, L.M. (2004). 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., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, **36**, pp.250-287.

Elbro, C., Nielson, I., & Petersen, D. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia* **44**, pp. 205-226.

Elliott, C.D., Smith, P., & McCulloch, K. (1996). *British Ability Scales II (BAS II)*. Windsor, England: NFER-Nelson.

Ellis, A.W. (2001). *Reading, Writing and Dyslexia: A Cognitive Analysis (2nd Ed).* East Sussex: Psychology Press.

Ellis, A.W., Young, A.W., & Anderson, C. (1988). Modes of visual word recognition in the left and right cerebral hemispheres. *Brain and Language*, **35**, pp. 254–273.

Foorman, B.R., Francis, D.J., Winikates, D., Metha, P., Schatschneider, C., & Fletcher, J.M. (1997). Early interventions for children with reading disabilities. *Scientific Studies of* Reading, **1**, pp.255-276.

Francis, D. J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational psychology*, **88**, pp.3-17.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.) *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading* (pp 301-331). London: Erlbaum.

Gustafson, S., & Samuelsson, S. (1999). Intelligence and dyslexia: Implications for diagnosis and intervention. *Scandinavian Journal of Psychology*, **40**, pp.127-134.

Hanley, J.R. (1997). Reading and spelling impairments in undergraduate students with developmental dyslexia. *Journal of Research in Reading*, **20**, pp.22-30.

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

Harm, M.W., & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review,* **106**, pp.491-528.

Hirsch, E.D. (2003). Reading comprehension requires knowledge – of words and the world. *American Educator*, **27**, pp. 10-22, 28.

Hulme, C., Hatcher, P.J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset–rime awareness. *Journal of Experimental Child Psychology*, **82,** 2–28

Hulme, C., Maughan, S., & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language,* **30**, pp.685-701.

Hulme, C., Newton, P., Cowan, N., Stuart, G., & Brown, G.D.A. (1999). Think before you speak: Pauses, memory search and trace redintegration processes in verbal memory span. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **25**, pp.447-463.

Hulme, C., & Snowling, M. J. (1997). *Dyslexia: Biology, cognition and intervention*. London:Whurr.

Hutzler, F., Ziegler, J.C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, **91**, pp.273-296.

Jastak, J. F., Bijou, S.W., & Jastak, D.E. (1976). *The Wide Range Achievement Test.* Jastak Associates.

Johnston, R. S., & Morrison, M. (2007). Towards a resolution of inconsistencies in the phonological deficit theory of reading disorders: phonological reading difficulties are more severe in high IQ poor readers*. Journal of Learning Disabilities*, **40**, pp. 66-79.

Johnston, R S, Perrett, D I, Anderson, M and Holligan, C. (1990). Perceptual dysfunction in poor readers: Evidence for visual as well as auditory segmentation problems in a sub-group of poor readers. *British Journal of Educational Psychology,* **60**, pp. 212- 219.

Johnston, R.S., Rugg, M.D., & Scott, T. (1987). Phonological similarity effects, memory span and developmental reading disorders: The nature of the relationship. *British Journal of Psychology*, **78**, pp.205-211.

Johnston, R. S., & Watson, J. (2004). Accelerating the development of reading, spelling, and phonemic awareness skills in initial readers. *Reading and Writing: An Interdisciplinary Journal,***17**, pp.327-357.

Johnston, R. S., & Watson, J. (2005). *The Effects of Synthetic Phonics Teaching of Reading and Spelling Attainment: A Seven-Year Longitudinal Study*. Accessed 04/04/2008, from http://www.scotland.gov.uk/resource/doc/36496/0023582.pdf

Karlsen, P.J., & Snodgrass, J.G. (2004). The word-frequency paradox for recall/recognition occurs for pictures. *Psychological Research*, **68**, pp. 271-276.

Karminloff-Smith, A. (1979) Micro- and macro developmental changes in language acquisition and other representational systems. *Cognitive Science,* **3,** pp. 91-118.

Karminloff-Smith, A. (1984). Children's problem solving. In M.E. Lamb, A.L. Brown, & B. Rogoff (Eds.), *Advances in Developmental Psychology, Vol. III.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Kerr, J. (1986). School hygiene in its mental, moral and physical aspects. *Journal of the Royal Statistical Society,* **60**, pp.613-680.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition.* N.Y.: Cambridge University Press.

Kuĉera, H., & Francis, W.N. (1967). Computational analysis of present-day American English. Providence: Brown University Press. Cited in MRC Psycholinguistic Database (2004). Last accessed 22/09/2007: http://www.psych.rl.ac.uk/MRC_Psych_Db.html

Lavidor, M., & Ellis, A.W. (2001). Orthographic neighbourhood effects in the right but not the left cerebral hemisphere. *Brain and Language.* **80,** pp. 63-76.

Lavidor, M., Hayes, A., Shilcock, R., & Ellis, A.W. (2004). Evaluating a split processing model of visual word recognition: effects of orthographic neighbourhood size. *Brain and Language* **88**, pp. 312-320.

Lavidor, M., Johnston, R.S, & Snowling, M.J. (2006). When phonology fails: Orthographic neighbourhood effects in dyslexia. *Brain and Language,* **96**, pp. 318-329.

Lavidor, M., & Walsh, V. (2003). A magnetic stimulation examination of orthographic neighbourhood effects in visual word recognition. *Journal of Cognitive Neuroscience*. **15,** *3*, pp.354-363.

Marshall, J.C. & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research*, **2**, pp.175-199.

McCaughey, M.W., Juola, J.F., Schadler, M & Ward, N.J. (1980). Whole word units are used before orthographic knowledge in perceptual development. *Journal of Experimental Child Psychology,* **30**, pp.411-421.

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

McDougall, S., Hulme, C., Ellis, A.W., & Monk, A. (1994). Learning to read: The role of short-term memory and phonological skills. *Journal of Experimental Child Psychology,* **58**, pp.112-133.

Mesulam, M. M. (1985). Verbal and non-verbal cancellation tests. *Principles of Behavioural Neurobiology.* F. A. Davis. Philadelphia.

Metsala, J.L., Stanovich, K.E., & Brown, G.D.A. (1998). Regularity effects and the phonological deficit model of reading disabilities: A meta-analytic review. *Journal of Experimental Psychology*, **90**, pp.279-293.

Miles, T.R. (1983). *Dyslexia: The pattern of difficulties*. London: Granada.

Morgan, W.P. (1896). A cause of congenital word blindness. *British Medical Journal,* **2**, pp.1564-1570

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive Processes in Spelling.* London: Academic Press.

MRC Psycholinguistic Database (2004). Last accessed 22/04/2008: http://www.psych.rl.ac.uk/MRC_Psych_Db.html

Nagy, W.E., and Scott, J. (2000). Vocabulary processes. in M.L Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.) *Handbook of Reading Research*, vol. III. Mahwah, N.J.: Lawrence Erlbaum Associates.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH publication number 00-4754).* Washington, DC: US Government Printing Office.

National Literacy Trust. (2007) *Literacy Changes Lives.* Accessed 09/10/2007, from http://www.literacytrust.org.uk/Database/Primary/phonicsresources.html

Nicholson, R. (2005). Dyslexia beyond the myth. *The Psychologist.* **18,** *11,* pp.658-659.

Nicholson, R.I., & Fawcett, A.J. (1990). Automaticity: A new framework for dyslexia research? *Cognition*, **35**, pp.159-182.

Nicholson, R.I., Fawcett, A.J., & Dean, P. (2001). Dyslexia, development and the cerebellum. *Trends in Neurosciences*, **24**, pp.515-516.

Neale, M.D. (1989). *Neale Analysis of Reading Ability Revised British Edition (NARA).* Windsor, England: NFER-Nelson.

Olson, R. K., Kliegel, R., Davidson, B. J., Foltz, G. (1985). Individual and developmental differences in reading disability. In MacKinnon, G. E. and Waller, T. G. (eds.),*Reading Research: Advances in Theory and Practice* (Vol. 4), Academic Press: New York, pp. 1–64.++

Palesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, F.S.J., & Frith, C.D. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain*, **119**, pp.143-157.

Perin, D. (1983). Phonemic segmentation and spelling. *British Journal of Psychology*, **74**, pp.129-144.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review,***103**, pp.56-115.

Powell, D., Plaut, D., & Funnell, E. (2006). Does the PMSP connectionist model of single word reading learn to read in the same way as a child? *Journal of Research in Reading,* **29**, pp.229-250.

Rack, J.P., Snowling, M.J., & Olson, R.K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, **27**, pp.28-53.

Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, **126**, pp.841-865.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance,* **25**, pp. 482-503.

Richek, M.A. (1977-1978). Readiness skills that predict initial word learning using two different methods of instruction. *Reading Research Quarterly,* **13**, pp.200-222.

Rose, J. (2006). *Independent Review of the Teaching of Early Reading*. Accessed on 04/04/2008, from http://www.standards.dfes.gov.uk/phonics/report.pdf

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

Ryan, M. (1994). *The Other Sixteen Hours: The Social and Emotional Problems of Dyslexia*. Baltimore: The International Dyslexia Association.

Savage, R., Stuart, M., & Hill, V. (2001). The role of scaffolding errors in reading development: Evidence from a longitudinal and a correlational study. *British Journal of Educational Psychology,* **71**, pp.1-13.

Scarborough, H., Ehri, L., Olson, R., & Fowler, A. (1998). The fate of phonemic awareness beyond the elementary school years. *Scientific Studies of Reading*, **2**, pp. 115–142.

Schonell, F.J. (1932). *Essentials in Teaching and Testing Spelling*. London: McMillan.

Seidenberg, M.S., & McClelland, J. (1989). A distributed, developmental model of word recognition. *Psychological Review,* **96**, pp.523-568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour,* **23,** pp. 383-404.

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

Shaywitz, S. E. (1996). Dyslexia. *Scientific American* **275**, pp. 98-104.

Shaywitz, S.E., Shaywitz, B.A., Pugh, K.R., Fulbright, R.K., Constable, R.T., Mencl, W.E., Shankweiler, D.P., Liberman, A.M., Skudlarski, P., Fletcher, J.M., Katz, L., Marchione, K.E., Lacadie, C., Gatenby, C., & Gore, J.C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of The National Academy of Sciences USA*, 95, pp.2636-2641.

Snowling, M.J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research*, **43**, pp. 219-234.

Snowling, M.J. (2000).*Dyslexia (2nd Ed).* Oxford: Blackwell Publishing.

Snowling, M.J., Goulandris, N. & Defty, N. (1996). A longitudinal study of reading development in dyslexic children. *Journal of Educational Psychology*, **88**, pp.653-669.

Snowling, M.J., Stothard, S.E., & McLean, J. (1996). *Graded Non-Word Reading Test.* Harcourt Assessment: England

Stanovich, K.E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly,* **26**, pp.7-29.

Stanovich, K.E., & Siegel, L.S. (1994). Phenotypic performance profile of reading-disabled children: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, **86**, pp.24-53.

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, **20**, pp.147-152.

Sticht, T.G., & James, J.H. (1984). Listening and reading. In Pearson, D. (Ed.) *Handbook of Research on Reading,* **1**, pp. 293-317. Cited in Hirsch, E.D. (2003). Reading comprehension requires knowledge – of words and the world. *American Educator*, **27**, pp. 10-22, 28.

Stuart, M., & Masterson, J. (1992). Patterns of reading and spelling in 10-year-old children related to pre-reading phonological abilities. *Journal of Experimental Child Psychology,* **54**, pp.168-187.

Swanson, H.L., & Berninger, V. (1995). The role of working memory in skilled and less skilled readers' comprehension. *Intelligence*, **21**, pp.83-108.

Szeszulski, P.A., & Manis, F.R. (1987). A comparison of word recognition processes in dyslexic and normal readers at two different reading-age levels. *Journal of Experimental Child Psychology*, **44**, pp364-376.

Talcott, J.B., Hansen, P.C., Willis-Owen, C., McKinnell, I.W., Richardson, A.J., & Stein, J.F. (1998). Visual magnocellular impairment in adult developmental dyslexics. *Neuro-Opthalmology*, **20**, pp. 187-201.

Temple, C., & Marshall, J.C. (1983). A case study of developmental phonological dyslexia. *British Journal of Psychology*, **38**, pp.77-110.

Torgerson, C.J., Brooks, G., & Hall, J. (2006). *A Systematic Review of the Research Literature on the Use of Phonics in the Teaching of Reading and Spelling.* London: Department for Education and Skills (DfES).

Torgesen, J.K. (2002). Lessons learned from intervention research in reading: A way to go before we rest. *Learning and Teaching Reading*, **1**, pp.89-103.

Torgesen, J.K., Morgan, S., & Davis, C. (1992). The effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology*, **84**, pp.364-370.

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, **27**, pp.276-286.

Torgesen, J.K., Wagner, R.K., Rashotte, C.A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, **91**, pp.579-593.

Treiman, R. (1993). Beginning to spell: A study of first-grade children. New York: Oxford University Press.

Valdois, S., Bosse, M., & Tainturier, M. (2004). The cognitive deficits responsible for developmental dyslexia: review of the evidence for a selective visual attention disorder. *Dyslexia*, **10**, pp.339-363

Van Ijzendoorn, M.H., & Bus, A.G. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly*, **29**, pp.267-275.

Wagner, R.K., & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, **101**, pp.192-212.

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1999). *Comprehensive Test of Phonological Processing (CTOPP).* Austin, Texas: Pro-Ed Inc.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory and Cognition,* **12,** *3*, pp. 293-305.

Wechsler, D. (1997). *Wechsler Abbreviated Scale of Intelligence (WAIS-III).* The Psychological Corporation, San Antonio, Texas.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children Fourth UK Edition (WISC-IV UK)*. San Antonio, Texas: Pearson Education Inc.

Witkin, H. A., Dyk, R. B., Faterson, H.F., Goodenough, D.R. & Karp, S.A. (1962). *Psychological Differentiation: Studies of Development*. New York, John Wiley & sons Inc.

Wyse, D., & Styles, M. (2007). Synthetic phonics and the teaching of reading: The debate surrounding England's 'Rose report'. *Literacy*, 41, pp.35-42.

Zorzi, M., Houghton, G., & Butterwoth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology, Human Perception and Performance,* **24**, pp.1131-1161.

APPENDICIES

Appendix 1:
Phonological distinctivness task

Appendix 2:
E-Deletion Task, and ‘e’s analysed.

Appendix 3:
Words used in experiment 3 with word frequencies

Appendix 4:
DB’s errors for Experiment 1

Appendix 5:
Questions on DB’s errors for Experiment 1

Appendix 6:
Non-word errors for Experiment 2

Appendix 7:
Word stimuli for Experiment 3

Appendix 8:
Non-word stimuli for Experiment 3

Appendix 9:
Word stimuli for Experiment 4

Appendix 10:
Non-word stimuli for Experiment 4

Appendix 11:
Graded non-word reading task stimuli for Experiment 5

Appendix 12:
Regularity task stimuli for Experiment 6 with word frequencies

Appendix 13:
Johnston and Morrison’s non-word reading task stimuli for Experiment 6

Appendix 1. Phonological Distinctiveness Task

Please read the following sentences aloud, into the microphone. Please read the sentences as if you were reading to a small child – i.e. slowly and with careful pronunciation. Please read through the sentences once, and when all have been read, please read them through for a second time in the same manner. The experimenter will demonstrate using the first sentence as an example. Thank you.

1. Using your best *pronunciation*, please read this article aloud.
2. Freud was very interested in *neuroticism*.
3. It is not *necessarily* a bad plan; it just needs a little tweaking.
4. That book doesn't fit in with the others it should go in the *miscellaneous* category.
5. The whole affair will end in *catastrophe*, you'll see.
6. The children played musical instruments in a *rhythmical* manner.
7. *Parliament* was re-opened by the Queen in accordance with tradition.
8. Alcohol *licensing* has appeared in the news very frequently recently.
9. I went to visit my *psychiatrist* today, she is very understanding.
10. An *Isosceles* triangle has only two equal sides.
11. The music was very *melancholy*.
12. The *Chrysanthemum* in my Nan's garden is beautiful.
13. The end of term always brings out the children's sense of *euphoria*.
14. Exam periods can be very *overwhelming* for students.
15. The boy bumped into the man, but it was *inadvertent*.
16. A *retrospective* view is helpful to understand past events.
17. I now know how to *resuscitate* someone, and when to use this skill.
18. We can use this example to *exemplify* Beethoven's genius.
19. The adolescents today are far from *naive*.
20. My laptop arrived neatly packaged in *polystyrene*.
21. The big black *locomotive* has broken down.
22. They needed a *thermometer* to measure the temperature of the water.
23. He always had a sandwich with *liver pâté* in his lunch packet.
24. Following the last *repair*, the car had caused nothing but trouble.
25. She found a *photograph* of a large lion in a trunk in the loft.
26. Her *handkerchief* had become dirty and needed a wash.
27. They saw an *elephant* in a film from Africa.
28. Thursday, they are going to a *birthday party* at Steven's.

The words in italics are the key words used for scoring, but did not appear in italics for the participants.

Appendix 2. E-Deletion task, and ‘e’s analysed.

Please read the following passage, crossing out every ‘e’ you come across. Please complete this task as quickly and as accurately as possible. Thank you.

Some of the Attractions in Hull

Hull City, located in the north-east of England, is home to numerous museums, and the world’s first submarium. The old-town hosts a museum quarter, with seven museums within walking distance of each other. The latest of these is the Hull and East Riding Museum (HERM), which recreates life from the Saxons and Romans to the Normans and Roundheads. Next door to this is the Streetlife Museum of Transport, which covers two hundred years of transport history; employing an 18^{th} Century stagecoach ride, a selection of cars and bicycles and an old bus. The Arctic Corsair is located at the rear of the Streetlife Museum and is Hull’s last sidewinder trawler, whose skippers were awarded the Silver Cod, a prize for the greatest catch of the season, and is moored on the river.

Hands on History is a museum dedicated to the treasures of the Egyptians, Victorian inventions and the story of Hull. Hull Maritime Museum, in the city centre, depicts life as a mariner, incorporating marine paintings, whale songs and a detailed history of fishing. Similar to this, Spurn Lightship illustrates the history of the Humber, and shows how the crew would have lived and worked. The seventh museum in the quarter is Ferens Art Gallery, which hosts an internationally renowned collection and has frequent visiting exhibitions.

The city also has many other attractions and attractive features, including a vast indoor market, cheap housing, the largest village in Europe, a large, friendly, university, a quickly-improving football team, and a Super League Rugby team.

Word	Times used	How many ‘e’s in the passage from this word	Stressed/Unstressed/Silent for each ‘e’ in the word
the	27	27	stressed
museum	7	7	stressed
other	2	2	unstressed
streetlife	2	6	stressed/stressed/unstressed
quarter	2	2	unstressed
located	2	2	unstressed
east	2	2	stressed
team	2	2	stressed
museums	2	2	stressed
life	2	2	silent
humber	1	1	unstressed
illustrates	1	1	silent
lived	1	1	silent
worked	1	1	unstressed
seventh	1	2	stressed/unstressed
have	1	1	silent
crew	1	1	stressed
mariner	1	1	unstressed
maritime	1	1	silent
centre	1	2	stressed/silent
depicts	1	1	stressed
inventions	1	1	stressed
egyptians	1	1	stressed
ferens	1	2	stressed/unstressed
detailed	1	2	stressed/silent
whale	1	1	silent
marine	1	1	silent
exhibitions	1	1	stressed
europe	1	2	stressed/silent
large	1	1	silent
village	1	1	silent
largest	1	1	unstressed
cheap	1	1	stressed
friendly	1	1	stressed
university	1	1	stressed
league	1	2	stressed/silent
super	1	1	unstressed
market	1	1	unstressed
frequent	1	2	stressed/unstressed
collection	1	1	stressed
renowned	1	2	stressed/unstressed
gallery	1	1	unstressed
internationally	1	1	unstressed
treasures	1	2	stressed/silent
features	1	2	stressed/silent
attractive	1	1	silent

were	1	2	stressed/silent
recreates	1	3	stressed/stressed/silent
HERM	1	1	stressed
roundheads	1	1	stressed
hundred	1	1	unstressed
years	1	1	stressed
covers	1	1	unstressed
next	1	1	stressed
these	1	2	stressed/silent
latest	1	1	unstressed
numerous	1	1	unstressed
home	1	1	silent
England	1	1	stressed
distance	1	1	silent
each	1	1	stressed
seven	1	2	stressed/unstressed
employing	1	1	stressed
silver	1	1	unstressed
awarded	1	1	unstressed
some	1	1	silent
whose	1	1	silent
skippers	1	1	unstressed
prize	1	1	silent
river	1	1	unstressed
moored	1	1	unstressed
season	1	1	stressed
greatest	1	2	stressed/unstressed
trawler	1	1	unstressed
sidewinder	1	2	silent/unstressed
bicycles	1	1	silent
selection	1	2	unstressed/stressed
ride	1	1	silent
century	1	1	stressed
stagecoach	1	1	silent
rear	1	1	stressed
dedicated	1	2	stressed/unstressed

Appendix 3: Words Used in Experiment 1 With Frequencies (Per Million) Taken From Kuĉera and Francis, MRC Psycholinguistic Database (2004).

Key: 1 = High Frequency Regular; 2 = High Frequency Exception 3 = High Frequency Strange; 4 = Low Frequency Regular, 5 = Low Frequency Exception, 6 = Low Frequency Strange

Word	Word Type	Word Freq	Word	Word Type	Word Freq
made	1	1125	deep	1	109
sew	6	6	chic	5	7
aire	6	0	age	1	227
find	1	399	thin	1	92
best	1	351	fete	6	3
plank	4	7	sour	6	3
hail	4	10	park	1	94
kick	4	16	warm	3	67
comb	6	6	said	3	1961
bass	6	16	post	3	84
lump	4	7	caught	3	98
flock	4	10	child	1	213
held	1	264	little	1	831
meet	1	148	weight	3	91
rock	1	75	gross	3	66
why	2	404	waif	5	0
sign	2	94	crux	5	2
line	1	298	death	3	277
life	1	715	done	3	320
friend	2	133	wand	6	1
eye	2	122	steak	6	10
bark	4	14	knight	5	18
smile	1	58	some	3	1617
blend	4	9	knew	2	395
loop	4	21	edge	2	78
weird	5	10	eight	2	104
gouge	5	1	touch	3	87
cliff	4	11	watch	3	81
scare	4	3	raid	6	10
knife	2	76	realm	6	19
view	2	186	isle	5	5
whole	2	309	yacht	5	4
rest	4	163	awe	5	5
wrong	2	129	doll	6	10
voice	2	226	break	3	88
lose	3	58	heard	3	247
give	3	391	puck	4	0
back	1	967	scab	4	0

Word	Word Type	Word Freq	Word	Word Type	Word Freq
two	2	1412	ache	5	4
earth	2	150	still	1	782
maths	4	0	beach	1	61
tusk	4	0	aunt	5	22
gig	4	1	debt	5	13
plant	1	125	word	3	274
play	1	200	sleigh	5	0
morgue	5	1	rhyme	5	3
glove	6	9	head	3	424
zest	4	5	worth	3	94
scene	2	106	suede	5	0
corps	2	109	tooth	6	20
pear	6	6	pint	6	13
bowl	6	23	gnat	5	0
once	2	499	hymn	5	9
piece	2	129	bier	6	0
rent	4	21	deaf	6	12
rip	4	6	own	3	772
doubt	2	114	bear	3	57
whose	2	252	worm	6	4
sock	4	4	soot	6	1
crypt	5	1	reign	5	7

Appendix 4. DB's errors for Experiment 1.

Word	Response	Word	Response	Word	Response
made	made	eye	none	age	age
sew	so	bark	bark	thin	thin
aire	air	smile	smile	fete	feet
find	find	blend	blend	sour	sour
best	bes	loop	loop	park	park
plank	plan	weird	weird	warm	warm
hail	hail	gouge	gouge	said	said
kick	kik	cliff	cliff	post	post
comb	com	scare	scare	caught	caught
bass	bass	knife	knife	child	child
lump	lum	view	view	little	little
flock	flo	whole	whole	weight	weight
held	held	rest	rest	gross	gross
meet	meet	wrong	wrong	waif	waif
rock	rock	voice	voice	crux	crux
why	why	lose	lose	death	death
sign	sign	give	give	done	done
line	line	back	back	wand	wand
life	life	deep	deep	steak	streek
friend	friend	chic	chic	knight	knight
some	some	knew	knew	edge	edge
eight	eight	touch	touch	watch	watch
raid	raid	realm	realm	isle	isle
yacht	yacht	awe	awe	doll	doll
break	break	heard	heard	puck	puck
scab	scab	two	none	earth	earth
maths	mass	tusk	tusk	gig	gig
plant	plant	play	play	morgue	morgue
glove	glove	zest	zest	scene	scene
corps	corpse	pear	pear	bowl	bowl
once	once	piece	piece	rent	rent
rip	rip	doubt	doubt	whose	whose
sock	sock	crypt	crypt	ache	ache
still	still	beach	beach	aunt	aunt
debt	debt	does	does	sleigh	sleigh
rhyme	rhyme	head	head	worth	worth
suede	suede	tooth	tooth	pint	pint
gnat	gnat	hymn	hymn	bier	bier
deaf	deaf	own	own	bear	bear
worm	worm	soot	soot	reign	reign

Appendix 5: Questions on DB's errors for Experiment 1.

Best
a. The most inferior thing in a group
b. The arrangement of certain items
c. The only item of a particular group
d. The most excellent of a particular group
DB answered correctly with d.

Plank
a. A length of sawn timber
b. A word for 'no'
c. A woody stem arising from a trunk of a tree
d. An alcoholic drink consisting of spirit distilled from a grape vine
DB answered correctly with a.

Comb
a. A strip of plastic with no teeth
b. A strip of bone or plastic with teeth
c. To make comfortable
d. To move towards something
DB answered correctly with b.

Lump
a. A dish of Scottish origin made from oats
b. A small solid mass without definite shape
c. A sound that is loud and/or shape
d. A child at the earliest stage of its life
DB answered correctly with b.

Flock
a. A group of animals of one kind
b. A group of porpoises or similar aquatic mammals that swim together
c. A case or covering for the blade of a knife
d. An ordered pile or heap
DB answered correctly with a.

Eye
a. No longer existing
b. Organ or faculty of sight
c. Reaching high or highest degree
d. Enlightening experience
DB answered correctly with b.

Two
a. To a greater extent than is desirable or permissible
b. One more than one
c. Small branch of trees or shrubs
d. Acute infectious fever
DB answered correctly with b.

Fete
a. An event to conclude a funeral
b. A small high-pitched flute
c. What determines a course of events
d. A gala, bazaar or similar entertainment
DB answered correctly with d.

Steak

a. A thick slice of meat
b. A stick or post sharpened at one end
c. An orderly pile or heap
d. A portion or share

DB answered correctly with a.

Maths

a. A group of sciences including algebra and geometry
b. A group of stars in the solar system
c. A group of sciences including biology and physics
d. An error or blunder in action

DB answered correctly with a.

Corps

a. A dead body
b. Charge made by a restaurant
c. Curved like the outside of a sphere or circle
d. Military unit with particular function.

DB answered incorrectly with a.

Appendix 6: Non-Word Errors

Non-Words as they are meant to be	**Non-word responses**
cad ney pol	cad ney pol
lin dap shig	lin dap shig
pruk min das	pruk min das
sul grim toab	slur grim too ab
kred lum dig	kred lum dig
ros tel crop	ros tel crop
brem tad lum	brim tu lum
go am strak	groam strak
mun tel klin	mun tek lin
tul phon kep	tul phon kep
jig lum tem	none
swa bla nap	none
wim trep fag	wim trep fag
sul wab clig	sul wab clig
drep cal nog	drep cal nog
rup nim kas	rup nim kas
nup clik das	nup clik das
fair krum dup	none
heg sim fap	heg sim fap
yoad un lim	yo dun lim

Appendix 7. Word Stimuli for Experiment 3.

Key	Word Type	Word	Key	Word Type	Word
1=regular	1	puck	1=regular	2	chic
2=strange	1	scab	2=strange	2	gnat
3=exception	1	maths	3=exception	2	hymn
	1	tusk		2	morgue
	1	gig		2	isle
	1	plank		2	waif
	1	hail		2	crux
	1	kick		2	gauge
	1	lump		2	weird
	1	flock		2	awe
	1	zest		3	sour
	1	bark		3	tooth
	1	blend		3	comb
	1	loop		3	bass
	1	cliff		3	pear
	1	sock		3	bowl
	1	rip		3	pint
	1	rent		3	glove
	1	rest		3	bier
	1	scare		3	deaf
	2	crypt		3	worm
	2	ache		3	soot
	2	aunt		3	wand
	2	debt		3	steak
	2	yacht		3	sew
	2	rhyme		3	raid
	2	suede		3	realm
	2	knight		3	doll
	2	sleigh		3	aire
	2	reign		3	fete

Appendix 8. Non-Word Stimuli for Experiment 3.

crupt	pisk	smur
acke	scib	torth
lunt	miths	camb
dest	tisk	basp
racht	gug	plar
rhame	plenk	bowm
srede	cail	pont
unight	louk	flove
slergh	lomp	blar
teign	fluck	draf
chit	zist	worb
grat	bara	sowt
hymt	blund	yand
morgup	lorp	stelk
iske	cleff	bew
waff	seck	rald
crun	rop	eralm
gluge	rint	dolt
weirm	rast	gire
aze	scire	lete

Appendix 9. Word Stimuli for Experiment 4.

Key	Word Type	Word
1=High lead High end	1	assent
2=High lead Low end	1	banker
3=Low lead High end	1	barbed
4=Low lead Low end	1	barman
	1	barred
	1	batten
	1	boring
	1	brandy
	1	burial
	1	castor
	1	cherry
	1	chilly
	1	clever
	1	corpse
	1	cradle
	1	denial
	1	forage
	1	frigid
	1	granny
	1	greedy
	1	insect
	1	manage
	1	manure
	1	marvel
	1	misuse
	1	seaman
	1	shovel
	1	sponge
	1	squire
	1	strand
	1	thrill
	1	timber

Key	Word Type	Word
1=High lead High end	2	ballot
2=High lead Low end	2	barrow
3=Low lead High end	2	breeze
4=Low lead Low end	2	cannon
	2	carpet
	2	carrot
	2	cheese
	2	craven
	2	deceit
	2	decree
	2	forbid
	2	forgot
	2	format
	2	heresy
	2	impair
	3	bitten
	3	bubble
	3	buffer
	3	dipper
	3	fiddle
	3	hamlet
	3	hammer
	3	hamper
	3	hasten
	3	hidden
	3	hunger
	3	keeper
	3	kettle
	3	kitten
	3	legion
	3	lessen
	3	piping
	3	rotate

Key
1=High lead High end
2=High lead Low end
3=Low lead High end
4=Low lead Low end

Word Type	Word
3	seller
3	violet
3	wallet
4	anthem
4	censor
4	chrome
4	coffin
4	donkey
4	errand
4	feudal
4	freeze
4	furrow
4	genera
4	grudge
4	hoarse
4	jagged
4	jockey
4	lappet
4	muzzle
4	oyster
4	pantry
4	puzzle
4	racket
4	refuse
4	rocket
4	scream
4	ticket
4	upkeep
4	vanity
4	vulgar
4	zenith

Appendix 10. Non-Word Stimuli for Experiment 4.

astent	ballat	ratate
balker	barraw	saller
brabed	bretze	vimlet
balman	cannin	waslet
barted	cerpet	anshem
balten	carret	celsor
bosing	chesle	catome
trandy	cravin	coffen
burtal	dechit	dunkey
castir	decred	errind
chesry	forbod	ferdal
chissy	frogot	fretze
clovor	formit	furrew
cortse	herest	genira
crable	imphir	gridge
densal	botten	hotuse
forate	bibble	jaggid
frimid	buffur	juckey
grinny	dippor	leppet
gresty	faddle	mazzle
insact	hamlit	soyter
manate	hammor	puntry
mabure	humper	pizzle
morvel	hastin	recket
misske	hodden	rofuse
setman	hungar	rockit
shivel	kesper	screlm
spinge	kottle	tocket
squore	bitsen	upkelp
strond	legson	vanuty
thrull	messen	volgar
timbar	puping	zanith

Appendix 11. Graded Non-Word Reading Task Stimuli (Taken from Snowling, Stothard & McLean, 1996) for Experiment 5.

hast
kisp
mosp
drant
prab
sted
gromp
trolb
snid
twesk
tegwop
balras
molsmit
nolcrid
twamket
stansert
hinshink
chamgalp
kipthirm
sloskon

Appendix 12. Regularity Task Stimuli (taken from Seidenberg, Waters, Barnes & Tanenhaus, 1984) for Experiment 6, With Frequencies (Per Million) Taken From Kuĉera and Francis, MRC Psycholinguistic Database (2004).

KEY: 1 = High Frequency Irregular; 2 = High Frequency Regular; 3 = Low Frequency Irregular; 4 = Low Frequency Regular.

WORD	WORD TYPE	WORD FREQUENCY
dance	2	90
foot	1	70
treat	4	26
prove	3	53
spook	3	0
bear	1	57
bowl	1	23
glove	3	9
great	1	665
bush	3	14
pest	4	4
lent	4	5
day	2	686
good	1	807
spade	4	10
down	2	895
broad	3	84
were	1	3284
pint	3	13
doll	3	10
rink	4	2
disk	4	25
soul	3	47
help	2	311
bough	3	2
come	1	630
deaf	3	12
aunt	3	22
kept	2	186
strong	2	202
wool	3	10
wake	4	23
lose	3	58
love	1	232
gross	3	66
base	4	91
touch	3	87
says	1	200
heard	1	247
phase	3	72
steak	3	10
luck	4	47
best	2	351
gate	4	37
gone	1	195
dust	4	70
bring	2	158
put	1	437
mile	4	48
gang	4	22
heat	2	97
green	2	116
slate	4	10
bread	1	41
take	2	611
turn	2	233
give	1	391
shall	1	267
still	2	782
does	1	485
rub	4	6
got	2	482
spear	4	7
stick	2	39
part	2	500
hard	2	202
went	2	507
both	1	730
stuck	4	23
dive	4	23
sew	3	6
done	1	320

Appendix 13. Johnston and Morrison's (2007) Non-Word Reading Task Stimuli for Experiment 6.

hig
nal
muntal
kug
renbok
gantok
bis
gok
minlan
dep
ritney
sanlud
nurdal
kun
ged
daspog
lar
ludpon
jek
foy
lan
culgin
yomter
mip
fambey
pos
kesdal
ruk
libnol
bosdin
dal
ped
lemfid
fik
mitson
lom
sul
goklup
bantik
puklon