VISIBLE ORTHOGRAPHY TRAINING AND BEGINNING READING SKILLS

WORD READING SKILLS OF BEGINNING NON-READERS: EFFECTS OF TRAINING WITH A VISIBLE ORTHOGRAPHY

By

HEATHER LOUISE POOLE, B.Sc.H.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2009) McMaster

University

(Department of Psychology, Neuroscience, & Behaviour) Hamilton, ON

TITLE: Word Reading Skills of Beginning Non-Readers: Effects of Training with a Visible Orthography

AUTHOR: Heather Louise Poole, B.Sc.H. (Queen's University)

SUPERVISOR: Professor B. A. Levy

NUMBER OF PAGES: vii, 100

ABSTRACT

The experiments presented here investigated the effects of manipulating the visibility of spelling patterns in English print, without concurrent oral segmentation, on the word identification skills of beginning non-readers. Visibility of the orthographic patterns was manipulated by presenting material organized into rime families (blocked) or with rime families distributed throughout training (unblocked), as well as through highlighting common rimes in the same colour of print. Experiment 1 demonstrated that while a program emphasizing the orthographic patterns in the English writing system (without concurrent phonological segmentation) led to rapid improvements in beginning non-readers' online word identification, the benefits of such training did not persist beyond the training context. Experiments 2A and 2B investigated whether the failure to transfer word reading skills beyond the blocked training context was mitigated by training programs that required increased focus on the letter patterns (2A) and the letter-sound relations (2B). These manipulations did not influence performance; children continued to demonstrate poor transfer beyond the training context. Experiment 3 focused on determining the mechanisms underlying the poor transfer following blocked training. To evaluate performance, this final experiment used a novel measure comprising word identification as well as onset and rime identification. Training materials were blocked either by rimes or onsets. The question of interest was whether training on material that blocks by orthographic units allows children to identify the blocked units during training without actively decoding their letter-sound relations, thus decreasing the probability of forming connections between the graphemes and phonemes comprising them. Results indicated that this is the case when children were trained on material blocked by rimes, but not that blocked by onsets.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Betty Ann Levy for all the support and guidance she has provided me over the years. I think there were times we both questioned whether I would ever get this project done! Your ability to demand results and provide support at the appropriate times has helped me get to this point.

Secondly, my committee members, Dr. Bruce Milliken and Dr. Karin Humphreys, have been a huge source of support, knowledge, and guidance throughout my years at McMaster. Thank you both for everything you've done. I've been very fortunate to have you both on my committee.

I would not have gotten through this experience without the support of good friends and wonderful family. My parents and siblings will have my everlasting appreciation for supporting me emotionally and often financially as I completed this degree. I have been fortunate to have many good friends who have given me input on my research, but more importantly, moral support over the years. Most of all I wish to thank a few special friends who I first met in my years as an undergraduate at Queen's University. Mayu Nishimura, we've now been through two sets of degrees together, and have known each other for over 10 years. Thank you for always being such an awesome housemate, a thoughtful friend, and an enthusiastic and skilled procrastinator! I may have finished earlier if you hadn't been around, but it sure wouldn't have been so much fun! Danny Krupp, thank you for all the coffee breaks where we discussed the academic and social challenges of the degree; your moral support and suggestions were irreplaceable.

There are many people who have supported me in ways that often don't get noticed. Nancy Riddell, thank you for your friendship and for guiding me through the administrative aspects of the degree. Gary Weatherill, thank you for keeping my (several) computers working over the years! Thanks also to you and Dave Thompson for your hard work on programming several of my experiments. Thanks to the research assistants who so willingly worked with the children in these experiments. Willow Fuchs, you were always a reliable, enthusiastic assistant. Thank you for all your hard work, and for your friendship. I've missed having you around as I've finished up at McMaster. Niki Brophy, thank you for your hard work in my final set of experiments. Melissa Bebee, thank you for the beautiful artwork in the books used in this dissertation.

Finally, I would like to thank my husband, Kevin Abbott. You know what you've done to help me through the last several years. I'm so excited for what we have ahead of us!

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS, CHARTS, DIAGRAMS

Introduction

Fluent reading relies on at least two interacting processes: decoding individual words and processing the sequence of words into a representation of meaning. While the ultimate goal of teaching children to read is to allow them to extract meaning from a text, they must first be able to decode individual words (Adams, 1990; Chall, 1989). This argument is present in several discussions and models of reading development. For example, Ehri (1999; 2005) has proposed that children are most efficient at extracting meaning from text once they have developed a large set of sight words. These are words that can be recognized automatically, so that attention can be fully directed to comprehension of the text. Fitting with Ehri's model, Adams (1990) has suggested that a child should be able to recognize at least 80% of the words on a page before being asked to read a text.

1. Chall's Stage Model of Reading Development

Jeanne Chall's (1983) model has major implications for such an argument. Well before publishing her model, Chall (1967) published a book summarizing the findings and implications of a massive three-year study undertaken to determine the best ways to teach beginning readers to read. The study was completed at a time when the majority of beginning reading materials used in the education system approached reading instruction through an emphasis on meaning, instead of an emphasis on deciphering the written code. The argument behind the use of meaning-based materials was that children needed to have the correspondence between written words and meaning emphasized if they were to gain comprehension skills when reading (Chall, 1983; Goodman, 1967). However, another school of thought supported the use of a code-based approach with beginning readers, arguing that children must be taught the code used in written language before they are asked to read a page of print (Chall, 1983). Chall's work in the 1960s compared these two approaches through analyses of published reading programs, visits to classrooms, interviews with experts in the field, and approaches used in laboratory and clinical settings. It was hoped that this breadth of resources, combined with the length of the study, would provide a clear answer to the best way to teach beginning reading skills.

The 1967 study, as well as a follow-up report (Chall, 1983), indicated that code-based instructional approaches were generally superior to meaning-based approaches for developing reading skills of children up to approximately Grade 3. That is, most work that directly compared phonics (code-based) approaches with whole language (meaning-based) approaches with young children revealed either no difference between the two, or a benefit of the former. However, this pattern did not persist with older readers. Indeed, the benefits of code-based instruction were most reliable in readers in Grades 1 and 2; the reading or reading-related skills of younger (pre-school) and older readers did not show nearly as strong a benefit of approaches that employed phonics. Thus, while Chall's (1967) conclusion that code-based approaches were the best method for young readers led to a massive shift from meaning-based to code-based emphasis in early education, research on early reading, published materials for beginning readers, and teacher education, it also led to some questions regarding the source of the age-related differences of the benefits of such an approach. Chall (1983) addressed the root of these differences in her model of stages of reading development.

The model comprises six stages of development and describes reading and its related language skills from birth through to adulthood. It provides approximate typical age ranges for each of the stages and explains qualitative differences in reading at each stage. These qualitative differences provide a possible explanation for the age-related shifts in the benefits of code-emphasis in reading instruction. Essentially, the first three stages can be considered as a time when students are 'learning to read', while in the last three stages, they are 'reading to learn' (Chall, 1983). While learning to read, an emphasis on the code used in written language is important. It is only once readers have developed skills through experience with that code that they become less dependent on active decoding of words and develop the more fluent reading that is typical of skilled readers, and which allows them to extract more complex ideas from the material.

In Stage 0, the prereading stage (birth to approximately 6 years), children are acquiring language skills that will later help with reading, such as forming the concept of using words to represent meaning and developing a vocabulary. By the time they enter school, many children are able to name most letters of the alphabet, and may even be able to write their names. By the time they are four years old, many children are able to discriminate writing from other symbols or pictures on a page (Levy, Gong, Hessels, Evans, & Jared, 2005). Fitting with Chall's proposal that the stages of her model progress in a hierarchical manner, the pre-reading skills demonstrated in Stage 0 have been successfully used to predict later reading skills (e.g. Levy et al., 2005). Chall argues that in Stage 0, children make use of the context provided through semantic or pictorial clues to identify words. In this stage, they have not yet formed associations between letters and their sounds, and so depend on clues provided by the material to derive its meaning.

In Stage 1 (age 6-7 years), children come to understand that letters can be used to represent sounds and combined to form words. They learn to decode words using the grapheme-phoneme correspondence rules; they also learn to recognize familiar words. Their ability to do this has been demonstrated by work on early readers' errors. Biemiller (1970) found that at the beginning of Grade 1, the majority of children's reading errors were consistent with the meaning of the passage. Children then progressed through a phase where they often failed to attempt to identify the words in the passage; this was followed by a final phase wherein their errors became consistent with the letters in the words, although they were still constrained by the context of the passage (Biemiller, 1970). Biemiller argued that children were progressing from an essentially pre-reading phase, where they did not understand the association between letters and sounds, to a phase where they were developing an understanding that letters could provide information on meaning but were not yet confident or skilled enough to attempt many guesses at the sounds. In this way, they eventually reached a phase where they were able to understand and use the associations between print and sounds to reach the meaning of the material. He believed that this developing understanding of the association between letters and their sounds was crucial to the transition to fluent reading, which is the next stage in Chall's model.

It is in Chall's Stage 2 (age 7-8 years) that children are able to demonstrate fluent reading skills, which allow them to rapidly recognize words on the page without having to work to identify the sound associated with each of the letters comprising them. Now that children's word decoding skills have developed somewhat (in Stage 1) they begin to efficiently process the meaning of text. Chall (1983) points out the importance of having familiar material available in this stage, so that children are able to use their knowledge of the subject matter to practice their reading skills and become more confident, fluent readers. Thus, children in Stage 2 are still 'learning to read'; they have not yet progressed to a level of expertise where they are able to 'read to learn'.

When they enter Stage 3 (age 9-14 years), children progress beyond relating print to sound and become able to relate print to ideas. They have developed relatively fluent reading skills which allow them to recognize and learn new words and process new ideas. The material read in this stage exposes children to new information and builds their knowledge about a variety of subjects, although it must do so in a manner that is simple enough for their cognitive abilities. For example, children in Stage 3 may be quite challenged by reading material that presents a variety of viewpoints, but would be able to read to learn a variety of facts and ideas about simple concepts.

In Stage 4 (age 14-18 years), readers progress to a level where they are able to process material that presents complex concepts and multiple viewpoints. The ability to build on existing world knowledge is crucial in this stage. Finally, in Stage 5 (beyond 18 years of age), readers are able to read constructively. They can actively interpret the material they read and use it to form new ideas of their own. They make use of their existing knowledge and their reading experience to know not only what to read, but also what not to read.

From the developmental stages that Chall (1983) presents, it follows that reading instruction should vary in kind according the age of the child. That is, presenting new ideas in books may be an appropriate way to teach a 13-year-old new information, but a 7-year-old would likely need to focus too much on decoding individual words to be able to extract much meaning from the passage. In contrast, asking a 6-year-old to read a book that presents repeated opportunities to rehearse simple three-letter words will likely improve his ability to recognize those words, while a 10-year-old would already be able to fluently identify them, and so would gain little from such an experience.

Crucially, according to Chall's (1983) model, children will not be able to 'read to learn' until they have developed fluent reading skills, and they will not be able to read fluently until they have developed strong decoding and word recognition skills. Thus, according to Chall's model, children should begin to read by learning the associations between letters and their sounds, which will allow them to gradually learn to recognize common words 'by sight', and to decode a wide variety of words they encounter in print. Asking them to read without first allowing them to develop strong decoding skills would miss a stage in their developmental trajectory.

It is important to note once again that such arguments are common in the literature. Recall that Adams (1990) has suggested that a child should be able to recognize at least 80% of the words on a page before being asked to read a text and Ehri (2005) has proposed that children are most efficient at extracting meaning from text once they have developed a large set of sight words. Empirical work indicating the importance of early teaching of the written code is also common. For example, Vellutino's (1991) review of research comparing whole language and phonics approaches in early reading instruction revealed a clear benefit of the latter (although it did not preclude the benefits of combining of the two). Similarly, a classroom-based study demonstrated that an early focus on phonics aids children who are just developing reading skills, but once they have developed the ability to read independently, more complex materials that allow discussions and presentations of various ideas and a broad vocabulary become effective (Juel & Minden-Cupp, 2000). This latter study affords external validity for Chall's (1983) model.

2. Phonological awareness as a predictor of reading ability

Chall's work in the 1960s led to a major shift in the focus of early reading instruction and research. Prior to her work, instruction had emphasized a whole-word approach, with the idea that this would lead children to focus on (and thereby extract) the meaning of the text (Chall, 1983). Her work, however, indicated that programs focusing on teaching children to link sounds to letters, known as decoding, reliably led to better reading skills, both in terms of word identification and comprehension.

This shift led to a heavy focus in the scientific literature on skills that would aid children in developing decoding skills that would allow them to identify a broad set of words. One skill that has been shown to have a robust relation to children's early word reading is phonological awareness (e.g. Adams, 1990; Christenson, 1997; Parrila, Kirby, & McQuarrie, 2004; Torppa et al., 2007), which is defined as the skills involved in detecting, manipulating, and distinguishing between the sound components within words. Phonological awareness can be measured at the level of syllables (Oakhill & Kyle, 2000) by asking children to detect or manipulate the syllables in a multi-syllabic word. For example, children could be asked to count the number of syllables in the word *doghouse*, or to name the syllables in the word *table*. Phonological awareness can also be measured in monosyllabic words by measuring it at the level of onsets and rimes or phonemes (Oakhill & Kyle, 2000). As an example, in the former case, children may be asked to orally segment the beginning sound from the word *lot* (l-ot), while in the latter case, children may be asked to orally segment each sound in the word (l-o-t). Initial hypotheses regarding the benefits of teaching children to read words by emphasizing their sound segments arose from the consistent evidence of a strong relationship between word reading ability and the performance on tests of phonological awareness (e.g. Adams, 1990; Goswami & Bryant, 1990; Parrila et al., 2004; Torppa et al., 2007).

Phonological awareness develops with age (Cardoso-Martins & Pennington, 2004; Liberman, Shankweiler, Fischer, & Carter, 1974); while young children's ability to understand language indicates that they can obviously hear the sounds in spoken language, they have difficulty segmenting or manipulating the sound units until they are near the age at which they begin to read. This pattern is especially apparent with onset-rime (Ehri & Robbins, 1992) and phoneme distinctions (Liberman et al., 1974). This developmental pattern is not surprising. In children's early interactions with language, the signal is oral, and the goal is to extract meaning from a string of sounds. It is not until children must learn to read and write that they must explicitly focus on the sound components of words. In alphabetic scripts such as English, these sound components correspond to individual letters or letter pairs. Thus, to extract meaning from written English, it is necessary to learn these correspondences; to extract meaning from spoken language, no such understanding of these correspondences is necessary (Goswami & Bryant, 1990).

It has been suggested that learning these correspondences is fundamental to reading an alphabetic language (e.g. Ehri et al., 2001). If this is true, we would expect to see a correlation between phonological awareness (specifically phonemic awareness) and word reading ability. As mentioned above, such a correlation has been consistently observed, and in a variety of populations. An early demonstration of the correlation between phonological awareness and reading ability came from illiterate adults in Portugal (Morais, Cary, Alegria, & Bertelson, 1979). When compared to literate adults, illiterate adults performed significantly worse on a phoneme addition and a phoneme deletion task. In a related study, illiterate adults did not demonstrate nearly such an extreme disability in syllable segmentation (Morais, Bertelson, Cary, & Alegria, 1986). These findings suggest that the skills that are necessary for learning to read an alphabetic language (phonemic awareness) are specifically poor in adult non-readers, despite the fact that they are able to understand and complete a segmentation task that involves larger sound units.

Additional data regarding this correlation comes from the finding that dyslexic and poor readers commonly have impaired phonological processing skills, as measured by tests of phonological awareness (Boada & Pennington, 2006) which have been shown to contribute to their reading disabilities (Wolf & Bowers, 1999). Cardoso-Martins and Pennington (2004) demonstrated that the phonemic awareness performance of children at a high familial risk of dyslexia was poorer than that of children at a low risk. This pattern was apparent longitudinally from Kindergarten through Grade 2. This detriment has even been demonstrated in terms of poor readers' ability to detect or segment onsets and rimes (Bradley & Bryant, 1978), a task which has been proposed by some to be less dependent on reading experience than phonemic awareness (Goswami, 2002).

Finally, early performance on phonological awareness tasks is correlated with current and later reading success, up to approximately Grade 2 (Ehri, 1999). Boscardin, Muthen, Francis, and Baker (2008) recently demonstrated that deficits in phonological awareness (measured using phoneme and onset-rime tasks) in Kindergarten are strongly predictive of word reading deficits in Grade 2. An additionally interesting effect was that the children with the poorest phonological awareness skills in Kindergarten showed the slowest developmental trajectory in word reading ability over the next two years. Thus, their early delays were associated with a Matthew effect (Stanovich, 1986), in that they fell further and further behind their peers over the first few years of school. McDougall, Hulme, Ellis, and Monk (1994) also demonstrated that performance on tests of rhyme awareness and phoneme deletion were positively correlated with reading ability in 7-to-9-year-olds, even with IQ controlled. Their work revealed that rhyme awareness was a somewhat less powerful predictor of reading ability than phoneme awareness, although these two variables contributed independent variance to the prediction of reading ability. Fitting with this finding, Hulme, Goetz, Gooch, Adams, and Snowling (2007) demonstrated that 7-to-12-year-old average readers' performance on a phoneme deletion task was predictive of their performance on word reading tasks. Indeed, performance on that task was the only unique predictor in their model that explained children's ability to read non-words, which requires a phonological approach to word reading since children have never seen the words before. Performance on phoneme deletion explained a massive 59% of the variance in non-word reading. Similarly, phonemic awareness has been shown to explain 50% of variance in 7-year-olds' word reading accuracy, even when differences in working memory are controlled (Leather & Henry, 1994).

In sum, the effects of the well-established correlation between performance on phonological awareness tasks and reading tasks have led to the declaration that phonological awareness is the single best predictor of word reading ability (Elbro, 1996; Parrila et al., 2004). The relationship between reading and phonological awareness is sensible when one considers the skills required to learn to read. For example, learning to read an alphabetic language requires processing of words at the level of phonemes, since each letter in an alphabetic language corresponds to a given phoneme. Thus, compared to children with weak phonological processing skills, those who are skilled at manipulating and discriminating different sounds will likely have an easier time learning how those sounds correspond to letters. A heavy emphasis on using this skill as an aid to reading development has thus emerged in the literature.

Despite this, it is important to recognize that in past years debate has continued between members of the research and educational communities with regards to the manner in which words should be presented to children who are learning to read (Ehri, Nunes, Stahl, & Willows, 2001). Some have continued to propose that children should be taught to read using a whole word manner of presentation, which does not encourage internal analysis of words, but uses repeated exposure so that readers come to recognize each word as a visual whole (Brown & Felton, 1990).

However, reading skill subsequent to whole word learning is typically surpassed by reading skill subsequent to word segmentation learning (Christenson & Bowey, 2005), although exposure to a whole word teaching paradigm does help children learn how to read words. For example, Levy and Lysynchuk (1997) demonstrated substantial learning in beginning nonreaders when they were taught to read words through a whole word program, wherein children simply read the training words as wholes, with no oral segmentation. After 15 days of training, the children were able to read, on average, 20 words that they had previously been unable to read. However, children trained to orally segment words as they were reading them acquired an average of approximately 28 new words over the same training period, thus learning approximately 25% more words than the group trained using a whole word method. Such contrasts between the relative benefits of whole word and segmentation training programs are common in the literature (Ehri et al., 2001). Thus, despite evidence that children are able to learn new words using a whole word method, and although this method has often been advocated in school systems (Ehri et al., 2001; Levy & Lysynchuk, 1997), reading skill subsequent to whole word learning is typically surpassed by reading skill subsequent to word segmentation learning. This finding will be further discussed below.

3. Phonological training programs

Although the link between phonological awareness and early reading skills is clearly robust, most of the work investigating the link has been limited to correlational designs (Ehri et al., 2001). Thus, the issue of causation has remained unclear. In a start to addressing this issue, the two skills have been shown to interact reciprocally (Adams, Treiman, & Pressley, 1998; Perfetti, Beck, Bell, & Hughes, 1987), in that strong phonological awareness skills are linked with strong reading skills, which enable better phonological awareness. Given the existence of this reciprocal relationship, researchers have proposed that phonological awareness is causally related to early reading ability, and thus that improvements in phonological awareness will lead to improvements in word reading ability (e.g. Bradley & Bryant, 1983; Bradley & Bryant, 1985; Goswami & Bryant, 1990; Hulme, Snowling, Caravolas, & Carroll, 2005).

3.1. Issues of segmentation unit

The logical way to improve phonological awareness is to teach children about the sound segments in language, and this has been a commonly used approach in the literature. However, even among the supporters of segmentation approaches, there exists considerable debate regarding the optimal unit to segment in reading programs for children. Some researchers advocate teaching children to read through the segmentation of phonemes, claiming that they are able to hear these and that, when they are not instructed on how to segment words, children choose these (Brown & Deavers, 1999).

Proponents of the phoneme segmentation view are supported by Ehri's (1999; 2005) theory, wherein she claims that the first representations children develop in their lexicon are those of the correspondence between each letter and its sound, and only later in reading development do they combine these representations into larger letter-sound chunks. Specifically, Ehri proposes that children progress through four phases as they develop the ability to read words by sight. The ultimate outcome of these phases (reading words by sight) is that words are read automatically and as units, instead of letter-by-letter.

Ehri's first phase is the pre-alphabetic phase. In this phase, children are non-readers and have not yet formed knowledge of the letter-sound connections inherent to the alphabetic principle. They may be able to identify certain words that are common in the environment (e.g. Coke), but this ability is not driven by decoding of the letter-sound correspondences. In the second phase, partial alphabetic, children know some of the correspondences between letters and their sounds. They use this partial knowledge of the alphabetic system to make educated 'guesses' at words, based on the letters they recognize and can identify with a sound. However, these guesses are not based on the entire sequence of letters in the word, so that children in the partial alphabetic phase might confuse the words *house* and *horse*. It is not until the third phase, the full alphabetic phase, that children know almost all of the correspondences between letters and their phonemes and can segment entire words into sounds letter-by-letter. By this point, representations of words in memory contain all of the letter-sound connections, instead of connections to a few letters in the word (as was the case in the partial alphabetic phase). Finally, in the fourth phase, the consolidated alphabetic phase, children become able to consolidate the sounds associated with individual letters into sounds associated with common groups of letters. Children in this phase have a large collection of sight words which are read automatically and in a unitized fashion, as well as a collection of sight units, such as syllables, rimes, and morphemes that are read in a similar way to sight words.

Clearly, according to Ehri's theory, programs that expose beginning readers to phonological training at the level of rimes or syllables would be inappropriate. There are data to support this view. For example, children have a hard time using rime-based analogies (i.e. using their knowledge of a given rime to determine the pronunciation of a new word containing that rime) until they have developed some decoding skill (Ehri & Robbins, 1992). Programs for beginning readers should instead, according to Ehri, present emphasis of phonemes.

However, others have countered this with the argument that children are not, in fact, able to segment individual phonemes until they have acquired some rudimentary reading skill (at the very least, letter knowledge) (Goswami & Bryant, 1990; Ziegler & Goswami, 2005). Supporters of this view propose that early reading programs should instead emphasize the segmentation of larger phonological units, such as syllables, or onsets and rimes, which are easier for nonreaders to handle (Goswami, 2002). This argument stems from research demonstrating that very young children who cannot yet read are able to identify words using analogies to words with the same rime (e.g. Goswami, 1986; Goswami & East, 2000; Wood & Farrington-Flint, 2002). Further, advocates of large-unit segmentation point out that such units have a more consistent letter-sound relationship than do phonemes (Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995), and are thus easier to consolidate into a representation.

3.2. Effects of phonological training on reading success

Regardless of the approach used to train children, there is, however, evidence to indicate that training children on phonological skills alone is not the best means to improve their reading. In a classic set of studies, Bradley and Bryant (1983; 1985) selected beginning readers who had poor phonological awareness skills, as measured by a rhyme oddity task. The children were divided into two phonological training groups, one of which received training on phoneme categorization, and the other of which received the same training paired with instruction using plastic letters to teach which letters were associated with each sound unit. Two other groups were also included, one of which received training in conceptual categorization (as a control for the phonological categorization group), and one of which was a classroom control that received no training. The experiment was longitudinal in design and posttests given after two years revealed that although children trained on phonological categorization demonstrated a trend of better reading ability than children trained on conceptual categorization, this trend was not statistically significant. However, the children who received phonological categorization training paired with letter-sound training demonstrated the best reading skills of any of the groups.

More recently, Hatcher, Hulme, and Ellis (1994) exposed poor readers to reading intervention programs that differentially applied phonological segmentation. One group was trained on phonological tasks only, without relating these activities to written words, while another group was trained on similar phonological tasks, but also on linking the sound units to letters and words that they read and wrote. When compared to a classroom control, there were clear benefits on all of the immediate measures of reading ability for the latter program, but not for the former. When the children were tested nine months later, this pattern persisted. Apparently, relating phonology to actual reading tasks is necessary for children's reading skills to improve; simply training them on phonological analysis alone will not have a notable influence on reading ability.

Similar findings were demonstrated with beginning readers who were performing below the level of their peers (Berninger et al., 2000).

Work by Lovett and colleagues also made this point through a comparison of training programs that required poor readers to read words at the whole-word level (no phonological segmentation training) or to orally segment the words into phonemes and then blend the phonemes together. In an early study, Lovett, Warren-Chaplin, Ransby, and Borden (1990) reported improved word recognition of trained words by poor readers in both groups. However, the two training groups did not differ significantly in the number of words they recognized and, surprisingly, the segmentation-training group showed no evidence of transfer of skills to untrained words. Thus, although the segmentation training did allow children to increase their reading vocabulary during training, they remained poor at transferring their existing knowledge to new reading experiences. In a subsequent attempt to reconcile the contradiction between the lack of transfer to new words following segmentation training and the well-known link between phonological ability and reading, Lovett et al. (1994) modified the phonological condition. They again used a phonemic segmentation and blending program, but this one placed more time and emphasis upon teaching individual letter-sound correspondences than did the segmentation program in the earlier study. This time, the segmentation program not only improved the children's ability to read the trained words, but it also improved their ability to transfer these skills to new words and their performance on standardized measures of reading ability. The authors suggested that the ability to transfer skills arose from the increased focus on letter-sound correspondence in the second study, as compared with their previous study, which had focused more on breaking words into phonemes without emphasizing the corresponding letter patterns. Again, this indicates that training children on phonological analysis alone will not teach them strong and general reading skills. Instead, if one aims to improve children's word reading in contexts outside that of training, it may be necessary to make clear not only the phonological segments of words, but also to emphasize the corresponding orthographic patterns.

Perhaps the strongest support for this argument comes from two meta-analyses of the effects of phonological awareness training on reading skills. In a meta-analysis of 52 studies investigating phonemic training on reading skills, training that paired phonemic awareness with letter training resulted in an effect size of 0.67 on immediate posttests and 0.59 on delayed posttests, while phonological training without letters resulted in effect sizes of 0.38 and 0.36, respectively, indicating a statistically significant benefit of phonemic training that is paired with letters (Ehri et al., 2001). An earlier meta-analysis revealed similar results (Bus & van IJzendoorn, 1999). Ehri et al. (2001) suggest that the benefits of letter training could be due to the fact that letters provide a lasting symbol which is more concrete and permanent than the ephemeral nature of sounds.

4. The role of orthography

Thus, it is clear that variance in reading skill is strongly influenced by phonological skills. Nonetheless, skills related to processing orthography also play an important role. Some interesting evidence for this argument comes from work by Wolf and Bowers and colleagues (Bowers & Wolf, 1993; Wolf & Bowers, 1999; Wolf, O'Rourke, Gidney, Lovett, Cirino, & Morris, 2002) with children whose reading skills fall at the bottom of the distribution. As many as 60% of these very poor readers have been shown to have deficits not only in phonological awareness, but also in naming speed (Wolf et al., 2002). Specifically, these readers have difficulty performing Rapid Automatized Naming (RAN) tasks, which involve rapid identification of letters and numbers, and thus require rapid, automatized access to visual symbol processing. Wolf et al. (2002) demonstrated that RAN performance and phonological awareness explain independent variance in reading ability in very poor readers, with naming speed being especially predictive of word identification, explaining 41% of the variance in this measure. Additionally, RAN has been shown to explain variance in reading skill beyond that explained by phonological awareness in children with a high familial risk of dyslexia (Cardoso-Martins & Pennington, 2004). Kindergarten students' ability to rapidly name a series of letters has also been shown to predict their word reading ability up to Grade 2 (Boscardin et al., 2008).

The relationship between RAN and reading ability is sensible when the processes underlying word identification are considered. Good readers fluently and automatically recognize orthographic patterns (Ehri et al., 2001; Wolf et al., 2002). Wolf and colleagues argue that similar processes would obviously be active during a task requiring rapid naming of letters and numbers (e.g. Bowers & Wolf, 1993). It is possible that training these double deficit readers on tasks designed to improve fluent, automatic identification of orthographic symbols would lead to improvements in reading skills (Bowers & Wolf, 1993). Indeed, it is possible that such tasks would help any beginning readers develop skills that will lead to fluent reading.

This possibility has led to the development of several programs that emphasize orthography through segmentation methods. As with phonological training programs, while there remains considerable debate among the supporters of segmentation regarding the optimal orthographic unit to segment in reading programs for children (Christenson & Bowey, 2005), we do know that experienced readers process written words at the level of these larger units (Landi, Perfetti, Bolger, Dunlap, & Foorman, 2006). Recall also that Ehri (2005) proposed that as children's reading improves, they form representations of these larger letter units by combining the individual letter-sound representations. What remains to be determined is whether training programs that emphasize larger letter units influence the likelihood that children will form representations of them so that they can use them to read fluently. It is possible that the earlier children become aware of these larger patterns, the stronger their reading skills will be; once they have developed an understanding of how the writing system works, children should be able to apply that understanding to new reading situations.

5. Training programs involving orthography

One way to emphasize the orthography of words is through colour highlighting of the shared letter units (Bruck & Treiman, 1992; Olson & Wise, 1992; Olson, Wise, Ring, & Johnson, 1997). Highlighting the shared units should make the consistent patterns very obvious to children, thus making it easier for them to learn about the consistency in English print and apply that knowledge to new reading situations. Highlighting can emphasize individual letters (i.e. which correspond to shared phonemes) or larger chunks, such as rimes. Arguments for the benefit of emphasizing rimes in such a manner have been based on the fact that they have a more consistent spelling-sound relationship than individual letters and phonemes (Treiman et al., 1995). Shared units can also be emphasized by presenting them in sets of multiple examples, or families (e.g. Christenson & Bowey, 2005); the emphasis seems especially strong when rimes are used. The use of one or both of these approaches in training programs should make the orthographic patterns in print obvious to children and thereby improve their word reading skills.

However, support for this hypothesis is variable. For example, Bruck and Treiman (1992) developed a paradigm to compare the effect of three types of segmentation methods on acquisition, retention, and generalization to new words. Beginning non-readers were trained to read a set of 'clue' words. Once they were well-learned, they were asked to read new words that could be identified through some analogy with the clue words. That is, the new words shared a certain orthographic pattern with the clue words: either a rime, an onset and vowel cluster, or a medial vowel. Furthermore, the orthographic similarity was emphasized by printing the shared letters in coloured ink, such that in the clue word 'rain' and the new word 'pain', the rime 'ain' was printed in red ink. This colour highlighting was done for all three conditions. Exposure during training implemented phonological (oral) segmentation as well as orthographic segmentation at the level of the shared unit. Performance during training seemed to indicate that rime analogies led to the best learning: children being trained to read using rime analogies learned all of the new words faster than children in either of the other two conditions. However, performance on tests of retention and generalization showed a reversal of this pattern. In fact, the best performance on tests of generalization came from children trained using vowel analogies. Thus, although performance during training was best when rime analogies were made visibly obvious, there was little evidence that the children had formed a strong representation of the sound units corresponding to the analogies they had been making.

Levy and her colleagues continued on a similar line of research in an effort to determine the mechanisms underlying the reversal between acquisition and retention or generalization of skills, and to clarify the debate regarding the optimal unit of segmentation. They completed a series of studies investigating several manipulations of concurrent phonological and orthographic segmentation (Levy, Bourassa, & Horn, 1999; Levy & Lysynchuk, 1997). Levy and Lysynchuk (1997) worked with beginning non-readers to compare the effects of four types of reading programs which differed in terms of the unit of segmentation. Using flashcards presented in sets of four, one group was trained to read a set of words as whole words, while other groups were trained to orally segment by phonemes, onset plus vowels, or rimes. For each segmentation condition, the relation between phonology and orthography was emphasized using a highlighting manipulation similar to that used by Bruck and Treiman (1992) and Olson and colleagues (Olson et al., 1997; Olson & Wise, 1992). For example, in the phoneme segmentation condition, each phoneme was printed in a different colour; in the onset condition, each onset was printed in a particular colour. This was intended to emphasize the orthographic similarities between words. Additionally, in the onset and rime conditions, the words were grouped into families, such that all four words in a set shared the same onset plus vowel but different final consonants (onset condition), or the same rime but different onsets (rime condition). Levy and Lysynchuk (1997) found that training on any of the segmentation methods led to faster and better acquisition than training on whole words. Although there was a trend for faster, better acquisition with segmentation at the level of larger units, this was not significantly better than acquisition with segmentation at the phoneme level. Furthermore, in contrast to the benefit of segmentation during training, retention analyses including only children who had learned all the training words revealed no benefit of segmentation over whole word training. Similar results were observed when children were asked to generalize their skills to new words and pseudowords containing the same onset and rime segments as the training words. Nonetheless, children who had received training did perform better on the tests of retention and generalization than did children in a control group who had received no training. In a second experiment testing Grade 2 readers who demonstrated reading delays, a similar pattern of results was observed. Thus, this additional work investigating the relative benefits of manipulations designed to make the orthography especially 'visible' to children (along with phonological emphases) did not reveal a clear benefit of any one type of segmentation.

It was possible that the effects of orthographic manipulations were simply less effective in children with typical or generally delayed reading skills, but would be more obvious in children who have suspected deficits in orthographic processing, such as the double deficit readers identified by Wolf and colleagues. Indeed, work by Olson and Wise (1992) supports such an argument. They compared children's word reading skills after programs that provided instruction at the onset-rime unit or at the syllable unit. Their results showed an interesting interaction between severity of reading impairments and the effect of the training program. Specifically, children who showed the worst reading skills benefited most from programs that emphasized syllable units, while children who were less severely delayed benefited more from programs that emphasized onsets and rimes. Levy et al. (1999) suggested that this finding could indicate that extremely delayed readers, such as those with weak phonological and orthographic processing, benefit from programs that emphasize larger units in print, such as whole word training, more so than from programs that emphasize smaller units, such as phonemes.

In a test of this hypothesis, they selected poor readers in Grade 2 and split them into two groups according to their performance on a RAN task. Children who performed well on the RAN task were identified as typical poor readers (i.e.

with delays in phonological processing), while those who performed badly on the RAN task were assumed to be similar to Bowers and Wolfs' (1993) double deficit readers (i.e. with delays in phonological and orthographic processing). Children were assigned to one of three flashcard training programs. The rime group saw words presented in groups of four, with all four words sharing a rime and each rime printed in a particular colour (colour highlighted). The phoneme and whole word groups also saw words presented in groups of four, but these were not grouped according to the shared orthographic unit. Words in the phoneme condition had each letter colour highlighted; words in the whole word condition included no colour highlighting of orthographic units. Children in the rime and phoneme conditions were asked to orally segment according to the unit of emphasis; children in the whole word condition were asked to read the words with no oral segmentation. A control group completed math activities during the training sessions.

During training, both RAN groups benefited from segmentation training, as compared to whole word training, but there was no differential benefit of the two segmentation approaches. However, the detriment of the whole word training was especially severe for the slow RAN group, indicating that processing at the level of larger orthographic units is especially difficult for these readers. Nonetheless, the relative benefits of the rime segmentation training, which also required processing at the level of larger orthographic units, conflicted with this conclusion. With consideration, a new possibility arose. The presentation in the rime segmentation training required not only that the common letter chunk be 'highlighted' in coloured ink, but also that it be presented, or 'blocked', with three other words containing that same chunk. The whole word condition contained no such manipulations. Perhaps, in fact, the relative benefits of rime training observed was due to the blocking and highlighting themselves, which resulted in a very 'visible' orthography to link to the corresponding phonology (Levy, 2001). In a direct test of this hypothesis, Levy et al. (2000) tested whether simply making the letter patterns in words more visible to children influences their reading performance, even if there is no corresponding oral segmentation in the training. Using flashcards presented in sets of four, poor readers were trained to read individual words using four types of orthographic segmentation at the onset-rime level. Training words were either blocked together according to their common rime, such that all four words shared the same rime, or were not blocked, such that none of the four words presented at once shared a rime. The rimes were also either highlighted in a different colour of print from their onset, with each rime being printed in a systematic colour throughout training, or each entire word was printed in the same colour. Orthogonal combination of these two manipulations gave rise to four training conditions. As described above, it was hypothesized that blocking would make the rimes 'visible', as the child would have four successive instances of a particular spelling pattern in the rime. Adding highlighting would make the common patterns even more visible. Crucially, this was a test of orthographic manipulations alone: the orthographic segmentation was not accompanied with oral segmentation. The children were to read the words as whole items without breaks between the onsets and the rimes.

During training, there was a clear benefit of blocking in terms of the number of words learned, and no additional benefit of highlighting. This appeared to support the possibility that training that allows children to easily see the common patterns in print improves their reading skills. However, tests of generalization and retention failed to confirm this; the blocked groups performed relatively poorly when the support of the training material was removed.

Clearly, training that emphasizes orthographic units often leads to a difference in performance for training versus retention. For example, Levy et al. (1999) found that although children trained on rime segmentation outperformed those using phoneme or whole word reading early in training and continued to outperform those trained on whole word reading throughout training, they identified significantly fewer words on tests of retention than did children in the other two training conditions. Recall that Bruck and Treiman (1992) similarly found that beginning readers improved more rapidly during training when asked to read words by making analogies to rimes than to onsets plus medial vowels or medial vowels, but these benefits did not last: retention was poorest following this rime family training. This pattern has been apparent in reading situations that require children to read words they have been trained on, as well as those that require them to read new words or pseudowords (Levy et al., 1999). Although there is some evidence that such reversals may not occur when children thoroughly learn all of the training words by the end of training (Levy & Lysynchuk, 1997), further work is needed to fully understand children's common failure to develop permanent and generalizable reading skills following training programs that emphasize the segments of the English writing system, and that support faster acquisition during the training phase.

6. The current experiments

The experiments presented here investigate the effects of Levy's 'visible orthography'. We know that orthographic skills explain variance in reading skills (e.g. (Bus & van IJzendoorn, 1999; Ehri et al., 2001; Levy et al., 2000). Children who are just beginning to read and have only rudimentary word decoding skills should be aided by print that has a visible orthography, in much the same way as delayed readers. Thus, exposure to a visible orthography should increase the efficiency with which they learn to read. As explained in detail in the chapters below, these experiments focus on whether a visible orthography, in the absence of concurrent phonological segmentation, leads to such an effect on reading skills. A further question that is investigated is the relation between training on a visible orthography and subsequent retention of reading skills.

Chapter 1

It has been well-established that the ability to manipulate and perceive the phonological components of words is correlated with reading ability. Correspondingly, making these components obvious to young readers has been shown to improve their online word identification (e.g. Lovett, Warren-Chaplin, Ransby, & Borden, 1990). However, there is evidence to indicate that pairing emphasis on phonological segments with emphasis on orthographic segments may further benefit online word identification (Levy, 2001; Walton & Walton, 2002). Indeed, work with poor readers indicates that emphasis on orthographic patterns alone, without concurrent emphasis on sound units, leads to dramatic improvements in online word identification (Levy, 2001). Clearly, when children make use of approaches that emphasize segmentation, whether at the oral or visual level, their online word identification skills improve.

However, the benefits to word identification are less reliable when the support of such training paradigms is removed. For example, Bruck and Treiman (1992) trained beginning readers to read words by making analogies to rimes, onsets plus medial vowels, or medial vowels. Improvements in word identification during training were most rapid in the rime analogy condition. However, these benefits did not last, with retention being poorest following this type of training. In contrast, Levy and Lysynchuk (1997) found that beginning non-readers' word identification following training emphasizing phonological and orthographic segmentation remained higher than that following whole word training one week after the end of training. Even several months later, there remained a benefit of training: although children trained using segmentation methods no longer outperformed those trained on whole words, they still outperformed children from a control group who had received no reading training. The same pattern was observed in young readers who were at risk for reading failure. However, Levy et al. (1999) failed to replicate such parallel performance during and after training. Specifically, although children trained on rime segmentation (at the oral and visual level) outperformed those using phoneme or whole word reading early in training and continued to outperform those trained on whole word reading throughout training, they identified significantly fewer words on tests of retention than did children in the other two training conditions. Levy (2001) reported the same costs to retention with very poor readers who had been trained on orthographic segmentation only: segmentation methods that led to the best performance during training led to the worst transfer to other reading situations which did not contain visual segmentation. This pattern has been apparent in reading situations that require children to read words they have been trained on, as well as those that require them to read new words or pseudowords (Levy et al., 1999).

Although there is some evidence that such reversals may not occur when children thoroughly learn all of the training words by the end of training (Levy & Lysynchuk, 1997), further work is needed to fully understand children's common failure to develop permanent and generalizable reading skills following training programs that emphasize the segments of the English writing system. The first experiment reported here was designed to meet this need. Following work indicating that poor readers' online word identification is dramatically improved by making orthographic and phonological patterns very obvious, it was important to investigate whether beginning non-readers would show similar benefits. These children, like children who are having difficulty learning to read, have under-developed decoding skills and few, if any, sight words. Since the English writing system comprises a complex orthography, learning to read is a very difficult task, and it is possible that making orthographic patterns obvious to these beginning readers will improve their ability to decode words containing those patterns. Furthermore, there is evidence that skilled readers are more likely than unskilled readers to process written words not on a letter-by-letter basis, but at the level of larger orthographic units (Brown & Deavers, 1999; Landi et al., 2006). Ehri (2005) also proposed that as children's reading improves, they form representations of these larger letter units by combining the individual letter-sound representations. Training approaches that emphasize those larger units may allow beginning readers to use word identification techniques that are more like those of skilled readers.

Alternatively, training may not influence beginning readers. Such an outcome would fit with evidence suggesting that poor readers rely more on orthography than children who are reading at an age-appropriate level (e.g. Foorman & Liberman, 1989; Snowling, 1980). Poor readers typically have delayed phonological skills for their age (Bruck & Treiman, 1992; Stanovich, Cunningham, & Cramer, 1984) and difficulty reading non-words (Baddeley, Ellis, Miles, & Lewis, 1982; Snowling, 1981), which requires phonological recoding. Nonetheless, their visual skills are typically unaffected: they demonstrate age-appropriate ability to recall letters from a novel orthography (Vellutino, Pruzek, Steger, & Meshoulam, 1973). Thus, manipulations that emphasize orthography, as opposed to phonology, may be especially helpful for poor readers, while normal beginning readers may find marginal benefit to emphasis on orthography, relative to emphasis on phonology. Previous manipulations indicating the benefits of orthographic emphasis (in the absence of phonological segmentation) to online word reading (e.g. Levy et al., 2000) could therefore be more effective with poor readers than beginning non-readers. Such an effect could be due to the fact that children whose reading skills are at an age-appropriate level naturally exhibit similar aptitude in phonological and orthographic tasks, while poor readers often demonstrate some discrepancy between these two reading-related skills. In that case, normal readers, but not poor readers, could make use of both types of information, while poor readers with phonological deficits may only have access to benefits of orthographic manipulations. Subsequently, the effects of orthographic manipulations may be more likely to emerge in a sample of poor readers than in a sample of normal readers. Fitting with this, there is evidence that, relative to children reading at an age-appropriate level, poor readers rely more on orthography than phonology (McNeil & Johnston, 2008). Their relative focus on orthographic patterns could lead them to have overdeveloped skills for orthographic tasks, relative to phonological ones, meaning that orthographic manipulations may be more effective with poor readers than normal readers.

Building on Levy's (2001) work indicating that emphasis of orthographic patterns explains variance in poor readers' online performance beyond that explained by emphasis of phonology, the training program used here included orthographic manipulations only. That is, the visibility of orthographic patterns was manipulated using colour highlighting and blocking of rime families, but all word identification was done at the whole-word level.

Thus, this first experiment addressed two main questions. Firstly, will a program that emphasizes the orthographic patterns in the English writing system, without concurrent phonological segmentation, lead to rapid improvements in beginning non-readers' online word identification? Secondly, if their online word identification is improved in the course of the training program that emphasizes orthography, will these beginning non-readers show lasting and generalizable reading skills following training?

Method

Participants

The final sample comprised 95 children (48 girls and 47 boys, mean age = 5 years, 6 months) from Senior Kindergarten classes in seven schools of a local school system. All children whose parents supplied written consent were screened for inclusion in the study. Children with diagnosed developmental delays or whose first language was not English were excluded from the study before screening. Participants were given a sticker or a baseball card each day they worked with the experimenter. Children knew they could withdraw from the study at any point they wished.

The initial screening for inclusion in the study tested the children's ability to read the target words to be used in the study. Words were printed individually on flash cards in black ink and presented randomly with the constraint that words with the same rime were never presented consecutively. Children who could read 8 or more of the 32 words were excluded from the study. The reading subtest of the Wide Range Achievement Test – Third Edition (WRAT-3) (Wilkinson, 1993) was used to assess letter knowledge and reading ability and the Test of Auditory Analysis Skills (TAAS) (Rosner, 1999) was used to assess phonological sensitivity. Children were required to demonstrate letter knowledge and phonological awareness that fell within one grade level of their current grade (Kindergarten) to be included in the study. Finally, the first two subtests from the Matrix Analogies Test – Expanded Form (MAT-EF) (Naglieri, 1985) were used to obtain a basic measure of non-verbal reasoning ability for all the children. All children tested scored within the normal range.

Design

There were four training conditions and one control condition. The training conditions were formed through orthogonal combination of orthographic blocking and highlighting manipulations. Children trained on blocked material read words organized into rime families, while those trained on unblocked material read words arranged so that the different rimes were distributed on a page, with no repetitions of a given rime on any page. Children trained on highlighted materials read words whose rimes were colour highlighted systematically, while those trained on material that was not highlighted read words which were printed entirely in coloured print, but with no systematic relation between colour and rime units. Orthogonal combination of blocking and highlighting thus formed four training groups: Blocked/Highlighted, Blocked/Not Highlighted, Unblocked/Highlighted, Unblocked/Not Highlighted. All groups received 20 training trials, spread over five training days. On the last four trials, the colour manipulation was removed and children read the words printed in black ink. The control group completed five sessions of math activities during the training sessions.

All five groups completed transfer posttests measuring reading of training words without the support of the training context immediately after the final training session and one week later. All groups also completed generalization posttests testing their ability to read words and pseudowords that shared rimes with the trained words.

Thus, each child completed a total of eight sessions with the experimenter. Pretests were given on Day 1. Any child who qualified for the experiment was then assigned to one of the five conditions. The five training sessions were completed on Day 2, Day 3, Day 4, Day 5, and Day 6. Following completion of the final training session on Day 6, all children completed the Immediate Transfer test. On Day 7, all children completed the two Generalization tests, and one week after Day 6, all children completed the Delayed Transfer test.

Materials

Training. The main training materials for the study were 10 books telling whimsical stories which were written by the author for the purpose of the study. Each storybook was 18 pages long and had four repetitions of each training word embedded in the story. Each book contained the same 32 training words (see Appendix A), based on the following eight rime families: *an, at, in, ig, it, ap, ill,* and *ash*, with each of the rimes repeated 16 times throughout each book. The children's attention was directed to the print during reading, but each page also had a colourful picture to go along with the story, which resulted in an age-appropriate style of book.

There were two different types of stories, one where training words with the same rimes were blocked together and the other where words with the same rime were distributed throughout the story (blocking manipulation). All training words in all books were printed in boldfaced type in AvantGarde Book font. Within each blocking manipulation (blocked or unblocked), one book presented the training words from each rime family systematically highlighted in a different colour from the onset of the word, and one did not contain highlighting of rime families (highlighting manipulation). To avoid a confound arising from the possibility that children may prefer coloured print to black print, the training words in the books without highlighting were written entirely in coloured ink. Thus, the 10 books resulted in the four reading conditions: Blocked/Highlighted, Blocked/Not Highlighted, Unblocked/Highlighted, and Unblocked/Not Highlighted.

The blocked books began by introducing a rime family; when this rime had been presented 16 times, a new rime family would be introduced[1]. Typically, when a rime family was introduced, all four training words from the family were repeated on that page. These words were then each repeated three more times over the next 2-4 pages. In contrast, although the unblocked books contained the same 32 training words as the blocked books, the members of each rime family were distributed throughout the book, such that a single page contained no two words having the same rime[2]. Examples from the four types of books appear in Appendix B.

For the final training day, each storybook had the highlighting manipulation removed; all of the training words were written in black boldface type. In all other aspects, these final training books were identical to those used on the previous training days.

Posttests. For all posttests, words were printed individually on flashcards in black ink. For the two transfer tests, the flashcards contained the 32 training words. For the generalization words test, 32 new words were presented on the cards, comprising four instances of each of the eight trained rime families paired with a new onset. For the generalization non-words test, another 32 new words comprising four instances of the eight trained rime families were presented, but these were paired with onsets that resulted in pseudowords.

---

[1] Note that to maintain a storyline, there were instances where a new rime family was introduced before the previous one had been repeated a full 32 times. In such cases, words from two different rime families were both presented on a page. Such occurrences were, however, rare.
[2] Again, to maintain storylines, there were rare instances when two words from the same rime family were presented on the same page in an unblocked book.

Procedure

Pretests. During this phase, the children were tested on several measures that are correlated with reading success. All children were tested on their ability to read the 32 target words; if they read eight or more, they were excluded from the study. For the purposes of matching ability across groups, the students were also tested on the TAAS, the reading subtest of the WRAT-3 (blue form only), and on the first two subtests of the MAT-EF. The means and ranges for each condition are displayed in Table 1. All children obtained scores within the normal range on all pretest measures. One-way ANOVAs indicated no reliable differences between the groups on any of the pretest measures ($p>.05$).

Table 1. Mean scores (M) and ranges (R) of groups on pretest measures.

		B/H	B/NH	UB/H	UB/NH	Control
N		18	21	18	20	18
	Boys	6	8	11	12	10
	Girls	12	13	7	8	8
Age (months)						
	M	65.9	66.5	65.9	66.6	66.3
	R	60-71	61-73	60-72	61-71	63-69
Words read pretest						
	M	1.1	1.1	1.7	2.2	1.4
	R	0-4	0-3	0-6	0-6	0-7
WRAT-3 standard score						
	M	110.8	109.8	110.8	110.2	109.4
	R	103-119	93-122	98-123	103-119	101-123
TAAS score						
	M	3.4	2.5	2.7	2.4	3.4
	R	0-9	1-6	1-8	0-5	1-9
MAT-EF 1 standard score						
	M	11.9	11.5	12.7	10.8	11.3
	R	8-16	6-17	8-17	6-16	6-16
MAT-EF 2 standard score						
	M	12.2	12.4	12.7	12.2	13.3
	R	7-16	5-19	8-19	7-19	7-19

Note. WRAT-3 = Wide Range Achievement Test, 3rd edition; TAAS = Test of Auditory Analysis Skills; MAT-EF = Matrix Analogies Test – Expanded Form, Subtests 1 and 2. Maximum possible standard score on WRAT-3 is 155. Maximum possible score on TAAS is 13. Maximum possible standard score on MAT-EF subtests is 19.

Training. Each child was trained individually with the experimenter. One attendance day after completing the pretests, children who qualified for the experiment completed their first training session with the experimenter. The experimenter modeled reading the target words for the child using flash cards presented in groups of four that corresponded to the child's experimental condition. The experimenter pointed to and read each word without oral segmentation and the child repeated it immediately.

For children in the experimental groups, a supported reading paradigm was implemented in each training session, wherein the experimenter read all words of the story except the target words, and the child read all instances of the 32 target words. If the child read a target word incorrectly or did not read it within 5 seconds, the experimenter immediately read the word for the child with no oral segmentation and continued reading the story words. The child's response (correct, incorrect, or omission) was recorded for each target word. Each session lasted between 12 and 30 minutes, depending upon the child. Children in the control group completed age-appropriate math activities with the experimenter during each training session.

Each child completed five days of training; on each day, a different story was read. On the first four days, the child read a book corresponding to his or her experimental condition with the target words printed in coloured ink. On the fifth training day, the colour highlighting manipulation was phased out; target words in all the books were printed in black ink in boldface type. The blocking condition was maintained as on the previous four days of training. The order of the five training books was counterbalanced within each group across children.

There were two Senior Kindergarten classes at each school; the children in each class attended school every other day of the week and on alternate Fridays. Thus, training was not completed over five consecutive days, but over five consecutive attendance days.

Posttests. All children, including those in the control group, completed all post-tests.

On the final training day, immediately following the reading of the fifth book, the children were tested on their ability to transfer their reading of the 32 training words to a no-context situation. They were asked to read each of the training words, presented one at a time in random order on individual cards and printed in black ink.

One attendance day after the final training day, the children were tested on their ability to generalize their knowledge of the trained material to new situations. First, they were asked to read 32 new words comprising four instances of the same eight rimes as the training words. The children were then shown 32 pseudowords comprising four instances of the same eight rimes. For both generalization tests, the words were presented one at a time in random order on flash cards and printed in black ink.

One week after the final training day, the children were tested on their ability to read the training words again, in a no-context situation[3]. This test was carried out in the same manner and using the same materials as the transfer test immediately following the final training session on Day 5.

In all posttests, no corrective feedback was given and all words read correctly were recorded. The words were presented in random order and the participants were allowed 10 seconds to read each word. As with the training sessions, the scheduling of generalization and retention tests was influenced by the attendance schedule of the different classes.

Results

Due to the nature of the design (the main analysis is a 4 (groups) x 15 (trials) mixed design ANOVA), the assumption of sphericity (one assumption of repeated measures ANOVA) was frequently violated. In that case, the Greenhouse-Geisser correction[4] was used to avoid inflation of false positive (Type I error) rates in analyses measuring change over training trials. Although this correction can sometimes be problematic if group sizes vary, the group sizes in this set of experiments were similar enough to allow its use. For simplicity of reading, if the p value was identical for the standard and Greenhouse-Geisser test, the standard degrees of freedom are reported (Baguley, 2004).

---

[3] All children also completed two 'in-context' retention tests, in the form of a Blocked book and an Unblocked book, each of which had only one repetition of each training word. Performance on this test was similar to that seen on the no-context retention test, so analyses have not been included here.

[4] This correction provides a more conservative test of the null hypothesis by modifying the degrees of freedom on which the hypothesis is evaluated. The degrees of freedom are calculated by multiplying the original df by epsilon. Epsilon is a statistic that indicates the extremity of the violation of sphericity; the further it falls from 1, the more severe is the violation (Baguley, 2004). When sphericity is violated, epsilon is smaller than 1, which leads to evaluation of the ANOVA on degrees of freedom with that involve decimals.

Online performance

Performance over training was analysed by comparing the mean number of words read correctly by each training group. There were five books, each containing four instances of the 32 training words. Thus, across training, a child had 20 attempts (or trials) to read each word; in the last four trials, the highlighting manipulation was removed and the words were printed in black ink. Figure 1 shows the number of words correctly identified on each trial.

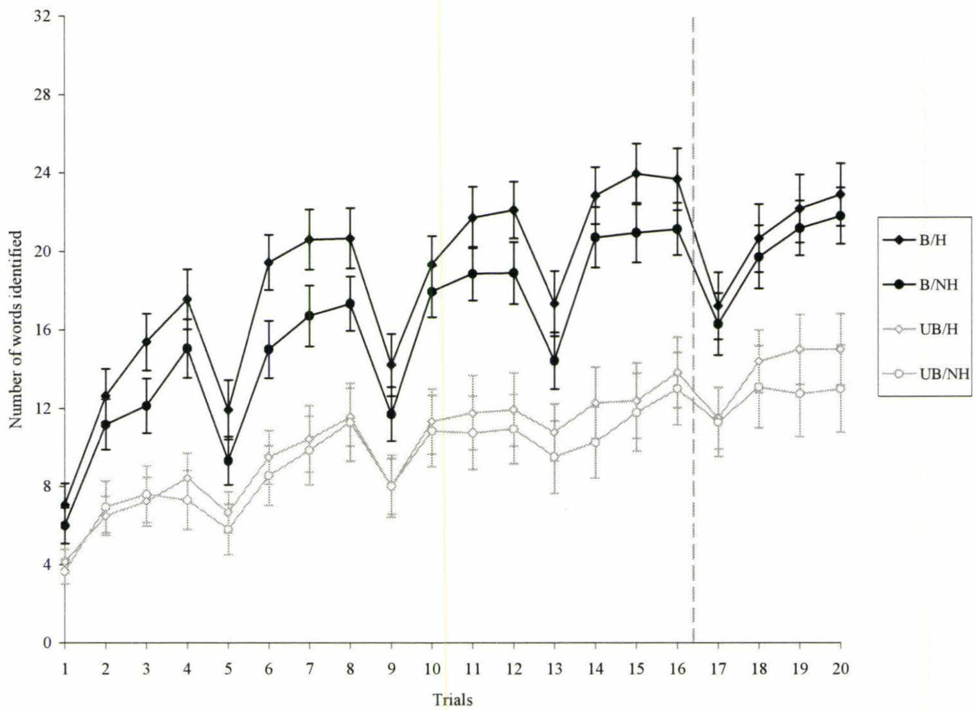


Figure 1. Mean number of words correctly identified on each trial of training (SE bars). Children completed four trials/day. Trials to the left of the dashed line used coloured text for the target words; those to the right used black text. (B/H=Blocked/Highlighted; B/NH=Blocked/Not Highlighted; UB/H=Unblocked/ Highlighted; UB/NH=Unblocked/Not Highlighted).

Data from the first four books (coloured trials) were analyzed using a 2 x 2 x 16 mixed design ANOVA, where the variables were Blocking, Highlighting, and Trials. The assumption of sphericity was not met, so repeated-measures data were evaluated using the Greenhouse-Geisser correction; this correction was used in all analyses that violated this assumption. There were main effects of Trials, $F(6.99, 509.89) = 110.86$, $MSE = 18.74$, $p < .001$, and of Blocking, $F(1, 73) = 28.36$, $MSE = 583.09$, $p < .001$, which were qualified by a significant Trials x Blocking interaction, $F(6.99, 509.89) = 11.08$, $MSE = 18.74$, $p < .001$. No other effects approached significance.

Due to the relative drop in performance on the first trial of each training day (Trials 5, 9, and 13), the Trials x Blocking interaction was further investigated by separating the first trial of each day from the subsequent trials of that day, with each blocking manipulation collapsed across the highlighting manipulation. A 2 x 4 mixed design ANOVA comparing the effects of the blocking manipulation on the first trial of each day indicated significant main effects of Trials, $F(2.26, 169.47) = 92.40$, $MSE = 11.89$, $p < .001$, and of Blocking, $F(1, 75) = 13.17$, $MSE = 111.35$, $p = .001$, which were again qualified by a significant Trials x Blocking interaction, $F(2.26, 169.47) = 3.65$, $MSE = 11.89$, $p < .025$. Thus, the number of words read on the first trial of each day improved in Blocked and Unblocked groups across training, with children in the Blocked groups reading more words on those trials than those in the Unblocked groups. To investigate the source of the interaction, difference scores between Trial 1 and Trial 4 were calculated for each group, providing a measure of their improvement over the course of training on the first trials of each day. An independent samples t-test comparing the two groups on these difference scores indicated that the interaction was driven by a different rate of improvement between the groups over the course of training. Specifically, children trained on blocked material showed more rapid increases in word identification on the first trial per day than children trained on unblocked material, $t(75) = -2.59$, $p = .01$.

From Figure 1, it was also apparent that performance on the first and subsequent trials of each day could be qualitatively different. To determine if it was only the first trial of each day that was driving the interaction seen in the omnibus analysis, the subsequent three trials of each day were collapsed together and the effects of the blocking manipulation across these trials were compared using a 2 x 4 mixed design ANOVA. Again, there were significant main effects of Trials, $F(2.70, 202.35) = 103.52$, $MSE = 6.39$, $p < .001$, and of Blocking, $F(1, 75)$

= 32.05, $MSE = 162.41$, $p < .001$, qualified by a significant Trials x Blocking interaction, $F(2.70, 202.35) = 6.26$, $MSE = 6.39$, $p = .001$. Thus, as on the first trial of each day, the number of words identified on the subsequent trials of each day increased in both groups across coloured training, with children in the Blocked groups reading more words on those trials than those in the Unblocked groups, and with groups improving at different rates. As on the first trial measure, a difference score indexing the difference between their performance on Day 1 and Day 4 for these subsequent trials was calculated for each group. An independent samples t-test comparing the two groups in terms of these difference scores indicated that word identification increased more rapidly in children trained on blocked material than in those trained on unblocked material, $t(75) = -3.63$, $p = .001$.

Thus, throughout coloured training, word identification during training on blocked material was superior to and improved more rapidly than during training on unblocked material, though both types of training led to significant increases in this measure. Colour highlighting did not influence performance.

The data for the final day of training, when all colouring was removed from the books, were analysed in a similar fashion. A 2 x 2 x 4 mixed design ANOVA revealed significant main effects of Trials, $F(2.24, 163.19) = 45.67$, $MSE = 7.72$, $p < .001$, and of Blocking, $F(1, 73) = 16.95$, $MSE = 220.85$, $p < .001$, as well as a significant Trials x Blocking interaction, $F(2.24, 163.19) = 6.09$, $MSE = 7.72$, $p < .005$. No other effects were significant. Thus, as during coloured trials, both blocked and unblocked training led to improvements in word identification and the effects of highlighting earlier in training were not significant. Furthermore, relative to training on unblocked material, training on blocked material led to better and more rapid improvements in word identification.

However, the interaction indicated that word identification by blocked and unblocked groups may have improved at different rates over these last four trials. To investigate the interaction, difference scores were calculated for the Blocked and Unblocked groups, collapsed across highlighting. Pairwise comparisons indicated that both groups showed similar large increases in the number of words read between Trials 17 and 18 (Blocked: mean increase = 3.44; Unblocked: mean increase = 2.32). However, between Trials 18 and 19, the number of words identified by the Blocked group continued to increase (mean increase = 1.49), while the Unblocked group's rate of increase leveled off (mean increase = .11); at this stage of training, there was a significant difference in the rate of increase between the groups, $t(75) = -2.63$, $p = .01$. Finally, between Trials 19 and 20, the two groups showed small and significant increases in the number of words identified (Blocked mean increase = .67; Unblocked mean increase = .13). Thus, while both groups showed substantial and similar increases in word identification between the first and second trials on the final day of training, children reading words blocked into rime families showed a more prolonged increase than did children reading unblocked words. Furthermore, as in earlier trials, in these final training trials, children reading blocked material identified more words than did children reading unblocked material.

Thus, by the end of training, children reading blocked material were identifying significantly more words than children reading unblocked material. Highlighting did not influence this pattern. Analysis of posttest performance addressed the second key question in this experiment, namely whether the benefits of blocking persisted when the support of the training material was removed.

Posttest performance: Immediate and Delayed Transfer (out of context)

The mean number of words read on posttests is presented in Table 2. A 2 x 2 x 2 mixed design ANOVA was used to compare how blocking and highlighting influenced the number of words identified when no story context was provided; this was tested immediately after the final training session and one week later. In contrast to the patterns observed at the end of training, the type of training received had no significant influence on posttest performance (all $ps > .13$). The four training groups performed similarly on the posttests, and performance did not significantly change over the course of the week between posttests.

Table 2. Mean number of words read and standard error (*SE*) by groups on post-test measures. All tests had a maximum possible score of 32.

		B/H	B/NH	UB/H	UB/NH	Control
Immediate transfer						
	Mean	9.17	8.48	10.67	11.00	2.28
	SE	*1.72*	*0.99*	*2.08*	*2.31*	*0.64*
Delayed transfer						
	Mean	10.06	6.95	11.78	11.79	3.67
	SE	*1.97*	*0.76*	*2.11*	*2.58*	*1.16*
Generalization - words						
	Mean	4.00	1.43	3.00	3.85	1.67
	SE	*1.52*	*0.68*	*0.89*	*1.30*	*0.63*
Generalization – pseudowords						
	Mean	3.89	1.29	2.61	3.70	2.06
	SE	*1.48*	*0.65*	*0.96*	*1.71*	*0.72*

A 2 x 5 mixed design ANOVA was used to compare posttest performance of the training and control groups. This revealed only a significant effect of group, $F(4, 89) = 4.17$, $MSE = 451.63$, $p < .005$. Posthoc analyses conducted using Dunnett's t-test revealed that all training groups, with the exception of the Blocked and Not Highlighted group, identified significantly more words than the control group on these tests (all *ps* $< .05$).

Posttest performance: Generalization

The majority of the children were unable to identify any words or pseudowords on the tests of generalization (see Table 2). Thus, the data were severely positively skewed and did not meet the standards for parametric analyses and were instead analysed using Kruskal-Wallis non-parametric analyses. These indicated that the groups did not differ significantly in terms of the number of generalization words or pseudowords they read (*p*s $> .30$). Thus, the training paradigm did not influence children's ability to identify new words or non-words, and indeed, they were no better at this task than their classmates who had received no training.

Discussion

This experiment was designed to address two questions: whether a training program that employed orthographic segmentation without concurrent phonological segmentation would lead to rapid improvements in online word identification, and whether improvements in online performance would be associated with permanent and general improvements in reading skills. Each question will be addressed in turn.

During training, blocked material not only led to more words being read, but also to more rapid improvement in word identification performance. Such improvements in online reading are consistent with earlier empirical work showing similar benefits during training using paradigms that emphasize consistent orthographic patterns in the English writing system concurrently with phonological segmentation (e.g. Berninger et al., 2000; Bradley & Bryant, 1983; 1985; Bruck & Treiman, 1992). These data provide clear evidence that beginning non-readers, like delayed or poor readers, show benefits to online performance when trained on material that emphasizes the segments of written English. They further demonstrate that orthographic segmentation, in the absence of phonological segmentation, is sufficient to bring about such improvements. This effect of orthographic segmentation on its own was the first of its kind with beginning non-readers.

These benefits are also consistent with Goswami's work (e.g. Goswami, 1986; Goswami & East, 2000) which has indicated that beginning readers are able to make use of rime analogies to aid in word identification. Indeed, the clarity of available analogies was the main difference between the blocked and unblocked conditions in this current experiment. The substantial differences in online word identification between children trained in these conditions indicate that the ready availability of rime analogies substantially aided performance on the training task. Thus, the first question this experiment was designed to address can be answered in the affirmative: training that emphasizes orthographic patterns through the use of rime families, without concurrent phonological segmentation, is sufficient to lead to rapid and substantial improvements in the online word identification skills of beginning non-readers.

It is interesting to note that the effects of orthographic emphasis are carried entirely by the blocking manipulation; colour highlighting the rimes did not influence online performance. The mechanisms underlying this lack of influence may differ for the two blocking conditions. That is, it is possible that rime family blocking makes rimes so salient to children that little effect is left for the highlighting. In contrast, when words are unblocked, children may find it difficult to recall the colour of a particular rime the last time they saw it, since a given rime would not be repeated in close proximity to earlier instances. Such a conclusion would be consistent with other work that has implemented colour highlighting with blocking for poor readers (Levy et al., 2000) and demonstrated no benefit of highlighting.

The second interest of this study was whether the online improvements seen with orthographic blocking are associated with permanent and general improvements in reading skill. In fact, the benefits of blocked training disappeared when transfer of word reading skills was examined. When children attempted to read the trained words without the story context immediately following their last training session, and one week later, there were no group differences in performance (although training did lead to better performance than was demonstrated by untrained peers in the control group). Furthermore, all children performed poorly on tests of generalization, wherein they were asked to read words and non-words comprised of the trained rimes paired with untrained onsets. Thus, although children benefited from blocking in the context of that supportive training paradigm, such benefits did not transfer to other reading situations, even for the trained material. In relation to training performance, this means that children trained on blocked material read fewer words on posttests than during training, while those trained on unblocked material read the same number of words on posttests as during training.

The most likely explanation for this effect is that children were able to use the blocked context during training to predict the identity of rimes in upcoming words, and thus did not fully consolidate the skills required to read those words in the absence of such a predictable context. This hypothesis is supported by work with pre-readers (Walton & Walton, 2002). One of their training paradigms employed extensive instruction using rime analogies. At the end of the training program, the children were asked to read several sets of words. These words were first presented in a manner similar to that used in transfer tests in this current study: each word was shown individually and the child was asked to read it. The authors then used an additional posttest wherein children were taught to read two 'clue' words that shared a rime and then were presented with a third word from the same rime family and asked to read that (with the two clue words still visible). Word reading was significantly better in the presence of clue words than when the word to be read was presented in isolation, indicating that children were making use of the clue provided by the rime family in their online reading performance. As in the current study, this strategy was apparent even after extensive training on rime analogies.

Thus, blocked training led to a dependence on the strong support provided by the training environment. When that support was removed, children demonstrated substantial drops in performance, and showed little evidence of having abstractly represented the letter-sound relations. These patterns of retention are again consistent with work by Bruck and Treiman (1992), wherein rapid acquisition of trained material was associated with poor retention. Similar patterns have also been elicited in older poor readers (Lemoine, Levy, & Hutchinson, 1993). Children trained on material organized into rime families with four exemplars made fewer errors in word identification during training than did children reading the same words in a scrambled order (comparable to the unblocked condition in this current experiment). However, as was observed here, when the children were asked to read the training words in a scrambled format as a test of retention, the two training groups performed very similarly.

Indeed, blocked, or massed, training programs often lead to rapid skill acquisition but fail to show corresponding enhancements to retention and generalization. This robust pattern has been demonstrated not only in reading development (e.g. Bruck & Treiman, 1992; Levy et al., 2000), but also on a broad range of cognitive and motor tasks (e.g. Carlson & Yaure, 1990; Dempster, 1987; Durgunoglu, Mir, & Arino-Marti, 1993; Glover & Corkill, 1987; Greene, 1989; Li & Wright, 2000; Toppino & Bloom, 2002). Several theories have been proposed to explain this widespread effect. These theories can be divided into two classes: contextual-variability and deficient processing theories. While the former is able to account for performance on free recall tasks following massed training, the latter class accounts for massed training effects when learning is tested using recognition or cued recall (Greene, 1989), as was done in this current experiment, and is thus particularly relevant to this training paradigm. Deficient processing theories are based on the argument that the relatively poor retention of material taught using massed training paradigms is due to deficient processing of that material during training. The general idea is that when participants are given a set of items to learn, they must allocate a certain amount of effort to learning each item. When similar or identical items are massed together, as they were in the case of blocked training in this experiment, participants will be more likely to think they know that item, and this will attenuate processing of that item. The benefits of repetitions are thus decreased substantially. Spaced, or unblocked items, on the other hand, will seem less familiar, since they will not have been recently processed. Due to this lack of familiarity, participants will be more likely to allocate additional processing effort to identify or encode the item (Greene, 1989).

Stated a slightly different way, when an item is repeated during training, participants will attempt to retrieve the information they encoded for that item on a previous trial or trials. If the earlier trial was recent, participants will still have the information very accessible in memory, and thus will retrieve what they previously encoded, allowing the correct response to be given without complete encoding of the incoming stimulus. However, if time or trials intervene between the previous and current experience with that item, participants will have forgotten the response they encoded, and must fully encode the incoming item (Cuddy & Jacoby, 1982). Using Jacoby's (1978) terms, these situations represent the difference between remembering the solution, as opposed to solving the problem. Deficient processing theories argue that this disparity in the extent of processing leads to disparities in memory for items learned in blocked or unblocked contexts. This idea is typically applied in list-learning contexts, but it can easily be applied to the learning task from the current experiment. In blocked training, children have recently paired a particular sound with a particular rime. When they encounter that rime shortly thereafter, they can recall the sound without processing the visual input with which it is associated. When they subsequently encounter a new rime, they will be forced to fully process the association between those letters and the corresponding sound, such that in blocked training, full processing will be likely to occur only on the first instance of each rime. In unblocked training, each instance of a particular rime is spread between several other rimes. The more items that intervene between a previous encounter with that rime and the current trial, the more likely the child will be to have forgotten the sound that corresponds to the letters. Thus, processing on each unblocked trial will be similar to the processing required on the first instance of a rime in blocked training. This more complete processing during training will lead to stronger memory traces for each rime taught during training, which will allow better performance on posttests.

Following from this, it seems possible that in the current study, the children reading blocked material were recalling the sound unit associated with a particular rime throughout the series of trials for that rime. Instead of decoding the letters corresponding to that sound each time they encountered them, they may have simply focused on decoding the onset, and added the predictable rime sound. Indeed, Bruck and Treiman (1992) proposed this possibility as an explanation for their results. In contrast, children reading unblocked material would not have been able to use the previous several words to predict what each word's rime sound was, and thus would have had to use their knowledge of letter sounds to decode the letters in both the onset and the rime in order to read a word.

It seems likely that whenever material is blocked, children will be able to use an approach that allows them to remember the solution, instead of solving the problem. The very nature of blocked material, namely, similar trials repeated in close succession, will encourage such a strategy. This in turn may very well lead children to identify rimes without processing the corresponding letter-sound association on each repetition, thus leading to the commonly documented disparity between online performance and consolidation. Thus, it may be possible to improve consolidation following blocked training by drawing children's attention to the letter patterns more directly. This hypothesis was investigated in the next experiment.

Chapter 2

Experiment 1 demonstrated a disparity between online word identification and transfer to new reading situations when children were trained on blocked material: blocked training benefited online word reading, but this benefit disappeared when the supportive training context was removed. Training on unblocked material did not lead to a similar disparity between online and transfer performance. The relatively poor transfer following blocked training could stem from the fact that children reading blocked material can use a strategy based on the predictability of the rime sounds so that rimes for later words in a block can be identified without decoding their constituent letters. Roberts and McDougall (2003) described such a strategy succinctly by suggesting that children may be using a *rhyme* analogy instead of a *rime* analogy. That is, children using a rhyme analogy identify words based on sound analogies, while children using a rime analogy identify words based on letter analogies. A rhyme analogy strategy would not be available to children reading unblocked material, since the sounds of the rimes in unblocked material are not predictable from the organizational context. The use of a rhyme analogy strategy has been posited by Bruck and Treiman (1992) to explain the disparity between online reading and memory for blocked training and by Landi et al. (2006), who found similar contrasts between word identification in online and transfer situations when readers were trained using a semantically-predictable context. To date, no approaches to avoid the detriment of such contextually-driven strategies have been investigated.

If, during blocked training, children are using a rhyme analogy (i.e. an analogy based on sounds, not letters) and thus not thoroughly processing the letters associated with each rime's sound, it is possible that transfer could be improved by pairing the shared reading task with tasks that require more thorough processing of the letters within each word, such as a spelling task. It has been suggested that spelling training can have an influence on reading development (Bishop, Adams, Lehtonen, & Rosen, 2003), and there is evidence to support such an argument (Ouellette & Senechal, 2008). In fact, Conrad (2005) found that training Grade 2 average readers in spelling led to better reading of trained words than did training in reading, and suggested that pairing training of spelling with reading of words in rime families may lead to maximal benefits in reading. Most spelling programs, including Conrad's, have been completed with children who are older than the ones used in this current line of research. It is unreasonable to expect the beginning non-readers used in this set of studies to be able to spell words, since they have not yet formed strong representations of letter-sound relations. However, modified spelling tasks that require recognition, as opposed to production of the correct spelling could still encourage children to focus on letter-sound relations in an analytical manner similar to that seen with traditional spelling tasks and could thereby lead to improved transfer of word identification skills.

The following experiments involved two such modified spelling tasks. The first task (Experiment 2A) was designed to direct children's attention to the visual patterns in the words. The second task (Experiment 2B) was designed to direct their attention to the letter-sound relations within the rimes, and thus more closely approximated the skills used in a traditional spelling task. The two experiments used a common comparison (control) group that replicated the Blocked/ Highlighted training group in Experiment 1.

Experiment 2A

Method

Participants. As in Experiment 1, any students who returned consent forms were screened for inclusion in the study. Selection criteria again excluded any children who read eight or more training words or who demonstrated letter knowledge below a kindergarten level (tested using the WRAT-3 reading subtest) at the outset of the study. In this experiment, knowledge of the training words was tested by having the words appear one at a time on a laptop computer screen, instead of on flashcards. The final sample comprised 39 children (25 girls and 14 boys, mean age = 5 years, 7 months); this sample was divided by random assignment into two training conditions. As in Experiment 1, the children attended school every other day and alternate Fridays, and worked individually with the experimenter.

Materials and procedure. Pretests. These were identical to the pretest materials used in Experiment 1 (see Table 2). As in Experiment 1, both groups received modeling appropriate to their training condition before training commenced. In this experiment, both groups received modeling with blocked and highlighted flashcards.

Training. The training materials for this study were the five Blocked storybooks used in Experiment 1. However, the books were now presented via laptop computer, such that one page at a time appeared on the computer screen. The experimenter or child pushed the space bar to go to the next page. There were two types of books: Reading Only and Colour Choice. In the Reading Only books, the rimes of target words were highlighted in the same way as Experiment 1, with the highlighting maintained throughout the entirety of training. Children assigned to this condition served as the control group (they provided a replication of the Blocked condition in Experiment 1). They completed a standard shared reading task for five days, as in Experiment 1.

The stories in the Colour Choice books were identical to those in the Reading Only books. The same target words were embedded in these stories. However, the presentation of the target words differed. Initially, each target word appeared on the screen written entirely in black boldface print with two dots of different colours beside it. At the top of the page were two model rimes, one of which matched the rime in the current target word. These rimes were written in two different colours that corresponded to the colours of the two dots beside the target word (see Appendix C). The children in this group were asked to complete the shared reading task (identical to that done by the Reading group). Additionally, after reading each target word children in this group were asked to use the computer mouse to click on the adjacent dot that matched the colour of the target word's rime modeled at the top of the page. When the children used the mouse to select a coloured dot, the rime in the target word automatically changed to the colour that matched the model rime at the top of the screen. This colour change was independent of whether or not the child selected the correct dot colour, though the children received corrective feedback on their choice if necessary. Accurate completion of the colour choice task thus required that children take note of the letters in the target word's rime and find those same letters at the top of the page. The colour choice task was designed to draw the children's attention to the common letter patterns in the words, but did not emphasize the common sound patterns any more so than in the Reading Only condition.

To encourage children in the Colour Choice group to look at the letters themselves and not simply form an association between a particular rime or sound and a given colour, the colour and position of the coloured dots was randomized throughout each book. Thus, each rime appeared in several different colours throughout the course of training. The position of the model rimes also changed across pages, so that, for example, a given rime would not always appear at the left side of the page.

Before training began, children in the Colour Choice group completed a task to ensure they understood the experimental task and knew how to use a mouse. This task was similar to the colour choice task that would be used during training, except that instead of matching colours based on letter sequences, the child was required to use a mouse to match colours based on number sequences. For example, a black digit was displayed at the bottom of the computer screen with two differently coloured dots beside it. At the top of the screen were that same digit and a different digit, written in the two colours of the dots. The child was asked to select the dot that would make the digit at the bottom of the screen match the colour of that same digit at the top of the screen. When a dot was selected, the target digit would change to the colour of the model digit. To emulate the multi-letter units that comprise rimes, the task began with single-digit numbers and progressed to multi-digit numbers, where the child was required to make the colour of the target digit match that of a model digit embedded in a multi-digit number. Feedback was provided on this task and all children performed well by the end of the training trials.

All children (Reading Only and Colour Choice groups) read all five Blocked storybooks over the course of the training period. The order of the storybooks was counterbalanced, and each child read a single story per consecutive school day. Feedback was provided on both reading performance and colour choice performance on a trial-by-trial basis.

Posttests. For the Transfer tests, the training words appeared one at a time in random order on the computer screen, displayed in black print. As in Experiment 1, the Immediate Transfer test was completed immediately after the final training session and the Delayed Transfer test was completed one week later. Generalization tests (words and pseudowords) were administered one school day after the final training session. The generalization flashcards from Experiment 1 were used as they were before, but because the children found these two tests long in Experiment 1, only two instances of each rime (instead of four) were used for each generalization test. Again, children received no corrective feedback on any posttest performance.

Results

As in Experiment 1, due to the nature of the design (the main analysis is a 2 (groups) x 20 (trials) mixed design ANOVA), the assumption of sphericity was frequently violated. The Greenhouse-Geisser correction was used to avoid for inflation of Type I error rates in within-subjects analyses. For simplicity of reading, if the p value was identical for the standard and Greenhouse-Geisser test, the standard degrees of freedom are reported (Baguley, 2004).

Matching of groups

Independent t-tests indicated that the two training groups did not differ on the pretest measures (all *p*s > .19; see Table 3).

Table 3. Mean scores and ranges of groups on pretest measures.

	Reading Only	Colour Choice
N	19	20
Boys	8	6
Girls	11	14
Age (months)		
Mean	67.6	67.4
Range	64-72	60-72
Words read pretest		
Mean	1.2	0.8
Range	0-4	0-4
WRAT-3 standard score		
Mean	107.4	109.5
Range	101-112	98-119
TAAS score		
Mean	3.2	2.8
Range	0-8	0-9
MAT-EF 1 standard score		
Mean	11.5	11.6
Range	6-17	8-16
MAT-EF 2 standard score		
Mean	12.4	11.9
Range	10-16	7-16

Note. WRAT-3 = Wide Range Achievement Test, 3rd edition; TAAS = Test of Auditory Analysis Skills; MAT-EF = Matrix Analogies Test – Expanded Form, Subtests 1 and 2. Maximum possible standard score on WRAT-3 is 155. Maximum possible score on TAAS is 13. Maximum possible standard score on MAT-EF subtests is 19.

Online word identification

As in Experiment 1, online performance was measured as the number of words correctly identified on each trial, where each book contained four trials, or instances, of each target word. Each child read five books, comprising 20 trials per target word over the course of training. There was a maximum possible score of 32 (8 rimes, 4 instances of each rime) on each trial. The data were analysed using a 2 (groups) x 20 (trials) mixed design ANOVA, which revealed only a main effect of Trials, $F(19, 703) = 74.32$, $MSE = 11.78$, $p < .001$. Performance improved across training. There were no significant differences between the two groups, and no significant interaction (all ps > .58; see Figure 2).

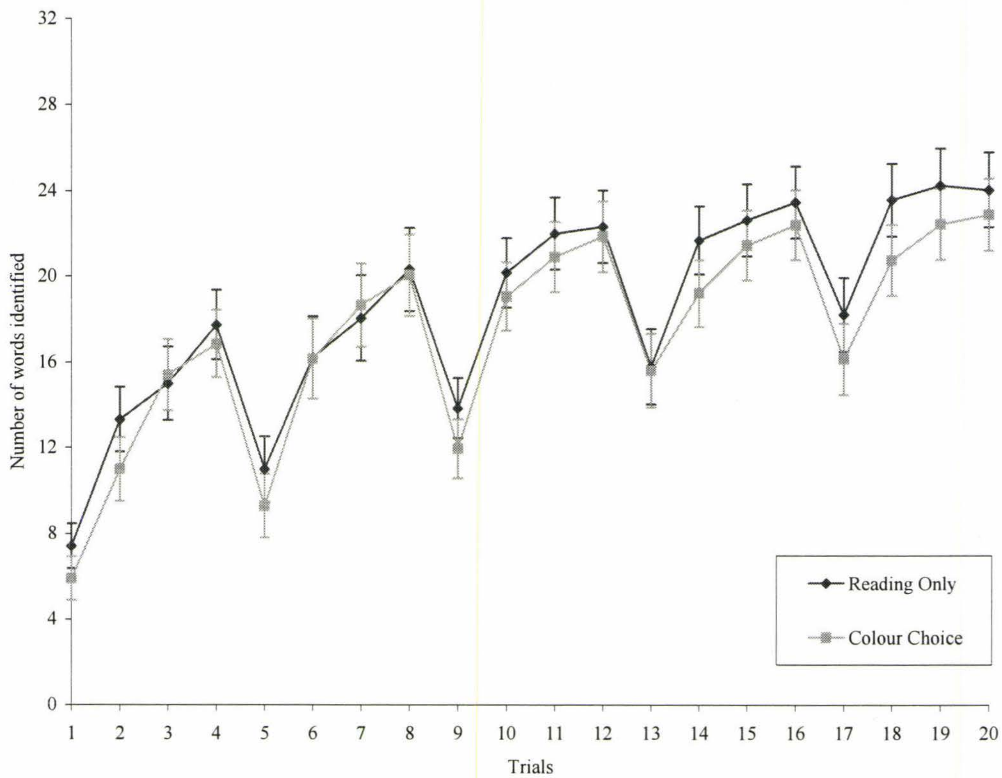


Figure 2. Mean number of words correctly identified on each trial of training (SE bars). Children completed four trials/day.

As in Experiment 1, performance on the first trial of each day was substantially lower than that at the end of the previous training day. A 2 (group) x 5 (trial) mixed design ANOVA indicated that while there were improvements in word identification on this first trial of each day, $F(4, 148) = 71.36$, $MSE = 9.82$, $p < .001$, the groups did not differ on this measure (all $ps > .44$). To ensure that the improvements in word identification did not occur only on the first trial of each day, data from the last three trials of each day were collapsed and analysed with a 2 (group) x 5 (trial) mixed design ANOVA. Again, this revealed that while performance on these trials improved significantly over the course of training, $F(4, 148) = 42.98$, $MSE = 9.62$, $p < .001$, there were no differences between the groups on this measure (all $ps > .62$). Thus, experience with the shared reading paradigm led to improvements in online word identification in a manner similar to that seen in Experiment 1, but the additional emphasis of the common letter patterns offered by the colour choice task did not exert a unique effect on online word identification.

Colour Choice Performance

Performance was measured using the number of correct colour choices per storybook (i.e. four trials); the maximum possible score per session was 128. Performance on the colour choice task was high from the onset of training; over the course of the five training days, average performance ranged between approximately 104 and 114 correct choices. A one-way repeated measures ANOVA revealed no significant change in performance over training, $F(2.55, 48.52) = 2.63$, $MSE = 119.50$, $p > .05$. Overall reading performance was not correlated with overall colour choice performance, $r(17) = .09$, $p > .05$.

Posttests

As in Experiment 1, performance on posttests following blocked training was substantially lower than performance at the end of training (see Figure 2 and Table 4). Critically, the additional emphasis on the letter patterns that was offered through the colour choice task did not prevent this pattern: a 2 (training condition) x 2 (transfer posttest) mixed design ANOVA indicated that the groups did not differ in their performance on the tests of immediate and delayed transfer ($p > .80$). As in Experiment 1, there were no significant changes in performance over the week that elapsed between the immediate and delayed transfer tests and no interaction ($ps > .40$). Thus, when children completed a shared reading task with materials that were blocked by rime, the addition of a task that further emphasized the letter patterns did not allow better transfer of skills beyond the training context.

Table 4. Mean number of words read and standard error (*SE*) on post-test measures. Transfer tests had a maximum possible score of 32; Generalization tests had a maximum possible score of 16.

	Reading Only	Colour Choice
Immediate transfer		
Mean	9.47	9.85
SE	*1.79*	*1.75*
Delayed transfer		
Mean	8.74	9.60
SE	*1.86*	*1.81*
Generalization - words		
Mean	1.89	1.90
SE	*0.62*	*0.64*
Generalization – pseudowords		
Mean	1.47	2.75
SE	*0.65*	*0.99*

Performance on the two generalization tests did not meet the assumptions of parametric tests, as the numbers of words read were not normally distributed; most children did not read any words at all, and those who did generally read very few. Thus, data from these tests were converted into categorical variables coding simply whether a child read or did not read words on the test. Any child who read words on these tests received a score of 1, regardless of the number of words read; children reading no words received a score of 0. A Pearson chi-squared test of independence revealed no significant differences between the two groups in terms of their ability to read new words, $\chi^2(1) = .03$, $p > .05$, or pseudowords, $\chi^2(1) = .03$, $p > .05$. Clearly, neither training program led to strong decoding skills for new letter combinations.

Discussion

As mentioned above, the Colour Choice manipulation was designed to direct the children's attention to common visual patterns in the print, but did not directly encourage further analysis of the sounds associated with those letter patterns. Although this manipulation led to improvements in online word identification over the course of training, the improvements were not superior to those observed in children who simply completed the shared reading program.

This outcome contrasts with hypotheses suggesting that relatively poor patterns of performance following blocked training could be due to children using a rhyming strategy that allows them to identify the answer without solving the problem during training (e.g. Bruck & Treiman, 1992; Jacoby, 1978; Roberts & McDougall, 2003). Such hypotheses imply that training that requires deeper analysis of the letters forming a word, as the colour choice task did, would increase consolidation of reading skills above that seen after the traditional blocked manipulation. For example, the generation effect (Jacoby, 1978) refers to the finding that retention for an item is better when it has been generated by the individual than when it has simply been read or presented in its entirety. Jacoby proposed that generation of a response requires more cognitive effort than reading the item, and that the increase in cognitive effort during training leads the item to be more accessible during tests of retention. Following from this reasoning, the Colour Choice manipulation in this experiment was designed to increase cognitive effort above that necessary for the Reading Only manipulation.

These data also contrast with empirical work in reading development. For example, Conrad (2005) demonstrated that manipulations requiring generation of a letter corresponding to a sound, as is the case in spelling tasks, lead to better reading performance than manipulations requiring generation of a sound corresponding to a letter, as is the case in reading tasks.

Nonetheless, given the similar performance during training for the two groups, it is not surprising that they yielded similar performance in transfer. It seems likely that this manipulation did not influence post-training performance because it was not sufficient to influence online performance. Although it was designed as a modified spelling task, it was substantially different from a traditional spelling task, and may simply have been too far removed from a letter-generation task to exert a notable influence on performance.

By this explanation, a task more akin to a traditional spelling task, where children generate letters that correspond to a particular sound, may lead to better skill consolidation. Related to this argument is evidence indicating that pairing oral and visual segmentation leads to somewhat better performance during training than using visual segmentation alone (Levy, 2001). Several studies have also shown that adding orthographic emphasis to phonological awareness training leads to better improvements in reading ability than phonological training alone (e.g. Berninger et al., 2000; Ehri et al., 2001; Hatcher et al., 1994; Lovett et al., 1994).

Given such data and the widespread evidence that phonological skills are strongly correlated with reading ability, it seemed likely that a manipulation encouraging thorough processing of the common letter units as well as their corresponding sounds would improve consolidation of the rimes' letter-sound relations and thus would lead to better transfer following blocked training. Such a manipulation was implemented in Experiment 2B.

Experiment 2B

Method

Participants. Students in Experiments 2A and 2B were recruited from the same school board and were screened in an identical manner for both experiments. As mentioned above, both experiments used the same comparison group, which completed only the shared reading manipulation. The final sample for this experiment comprised 37 children (22 girls, 15 boys; mean age = 5 years, 8 months); of these, 19 belonged to the Reading Only group described in Experiment 2A.

Materials and procedure. Pretests. Materials for pretests were identical to those used in Experiment 2A. As in previous experiments, both groups received modeling of blocked and highlighted words.

Training. The Reading Only group in this experiment was the one that was used in Experiment 2A; children in this group completed the blocked and highlighted shared reading task, as described above. Children in the Rime Choice group completed the shared reading paradigm as well, but also completed a modified spelling task. After each page in the story, children saw each onset from the previous page appear on the screen, one at a time, followed by a blank (e.g. c___, r___). Children were told by the experimenter what each word was supposed to be, and were asked to use the computer mouse to choose the rime that would correctly complete that word from two rimes that were printed in two different colours at the top of the page. When either rime was chosen, the correct rime appeared in the target word, added to the onset (e.g. c*at*, r*at*) and printed in colour so as to be highlighted (see Appendix D). If necessary, corrective feedback was given on this task. The onset from the next target word then appeared on the screen, and the child worked through rime selection for all of the words on the previous page before continuing with the next page of the story.

The identity of the alternative rime options differed randomly from trial to trial, as did their positions at the top of the page (left or right). The two rimes were written in different colours, and with each new onset, the colours changed. Thus, for a single page, the same rime may always have been present as an option, due to the blocked nature of the material, but the position of that rime and the alternative rime varied, as did their colours. This variation was implemented because pilot work indicated that without it, children used position and print colour as cues to the correct rime, which could decrease focus on the letters of each rime.

As in Experiment 2A, children in the Rime Choice group completed familiarization exercises with the computer mouse before training commenced (these are described it the Method section of Experiment 2A). This confirmed that they knew how to use the mouse and understood the experimental task.

Posttests. These were identical to the posttests administered in Experiment 2A, with the addition of a modified spelling test which was comparable to the Rime Choice task described above. Children were given a sheet of paper with 32 onsets, each of which was followed by a blank and two rimes. They were told by the experimenter the identity of each word and were asked to circle the rime that would complete the word. This test was completed after the immediate transfer test. No corrective feedback was given on any posttests.

Results

Pretests. Independent t-tests indicated that the groups did not differ on any of the pretest measures (see Table 5).

Table 5. Mean scores and ranges of groups on pretest measures.

	Reading Only	Rime Choice
N	19	18
Boys	8	7
Girls	11	11
Age (months)		
Mean	67.6	69.3
Range	64-72	64-73
Words read pretest		
Mean	1.2	1.1
Range	0-4	0-5
WRAT-3 standard score		
Mean	107.4	104.7
Range	101-112	100-117
TAAS score		
Mean	3.2	3.0
Range	0-8	0-8
MAT-EF 1 standard score		
Mean	11.5	10.9
Range	6-17	6-14
MAT-EF 2 standard score		
Mean	12.4	11.6
Range	10-16	9-14

Note. WRAT-3 = Wide Range Achievement Test, 3rd edition; TAAS = Test of Auditory Analysis Skills; MAT-EF = Matrix Analogies Test – Expanded Form, Subtests 1 and 2. Maximum possible standard score on WRAT-3 is 155. Maximum possible score on TAAS is 13. Maximum possible standard score on MAT-EF subtests is 19.

Online word identification. Patterns of online performance were very similar to those seen throughout blocked training in Experiments 1 and 2A (see Figure 3). Peformance was analysed using a 2x20 mixed design ANOVA, with Training Condition as a between-subjects factor and Trials as a within-subjects factor. There was a significant main effect of trials, $F(19, 665) = 72.13$, $MSE = 11.90$, $p < .001$, such that word identification performance improved over the training trials. There were no significant differences between the groups and no interaction (all *p*s > .32).

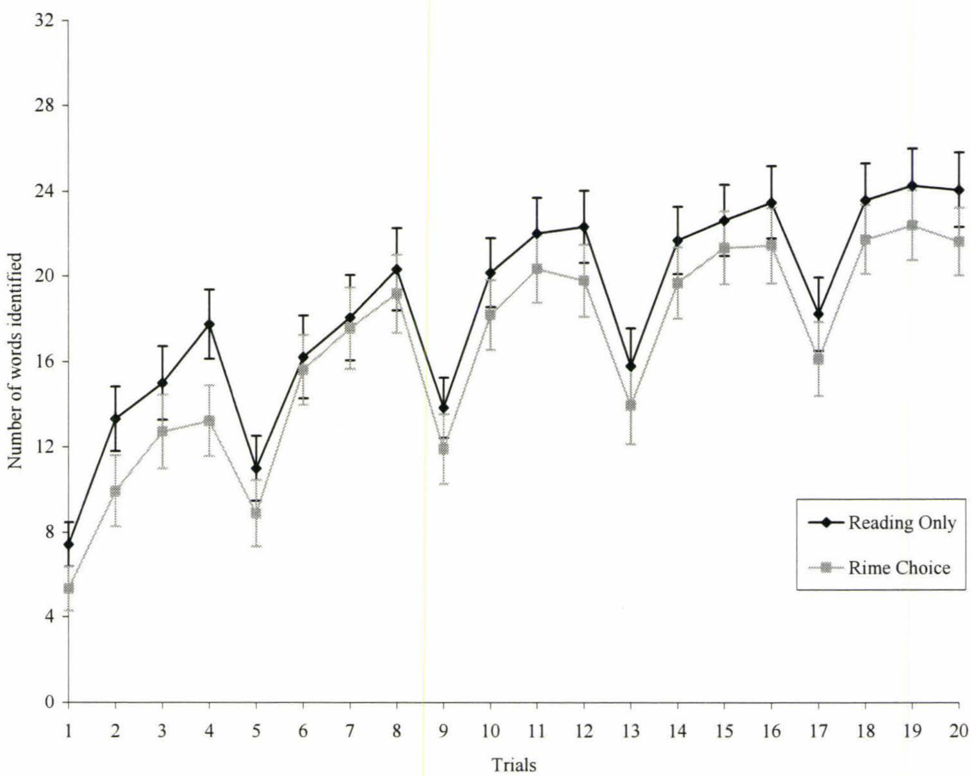


Figure 3. Mean number of words correctly identified on each trial of training (SE bars). Children completed four trials/day.

Again, further analyses were completed to determine whether improvement over the course of training occurred only on the first trial of each day, or whether improvement also occurred in later trials each day. A 2x5 mixed design ANOVA analysing performance on the first trial of each day indicated only a significant main effect of Trials, $F(4, 140) = 61.07$, $MSE = 10.77$, $p < .001$, such that word identification performance was improving over the course of the five training days on the first trial of each day. Similarly, a 2x5 mixed design ANOVA comparing the two groups on the number of words identified on the subsequent three trials of each day (collapsed together) indicated only a significant main effect of Trials, $F(4, 140) = 51.24$, $MSE = 9.81$, $p < .001$, again showing an improvement over the course of the five training days. Thus, the groups did not differ on any measures of online word identification.

Rime choice. The children demonstrated very good performance on the rime choice task throughout training. Performance was quantified on a daily basis, with each day scored out of 128. A repeated-measures ANOVA indicated that, contrary to the patterns shown by children in the Colour Choice group in Experiment 2A, performance on the modified spelling task did significantly improve throughout training, $F(4, 68) = 12.59$, $MSE = 104.56$, $p < .001$. However, even on the first day of training (i.e. collapsed over the first four trials), accuracy on this measure was high (73%), indicating that children were performing near ceiling with very little experience on this task. Despite this, overall performance on the rime choice task was correlated with overall reading performance, $r(16) = .47$, $p = .05$. During training, children who identified more rimes correctly read more words correctly.

Posttests. Again, children identified substantially fewer words on the posttests than at the end of training (see Figure 3 and Table 6). A 2x2 mixed design ANOVA revealed no differences between the Reading Only and Rime Choice groups on the two post-training tests of transfer, nor any change in this over the week following training (all *ps* > .31). As in Experiment 2A, performance on the generalization tests was extremely low and did not meet the assumptions of parametric analyses. Performance was thus categorized so as to code whether the child read any words. Pearson's chi-square revealed no significant effect of training condition on children's ability to read new words, $x^2(1) = .03$, $p > .05$, or pseudowords, $x^2(1) = 03$, $p > .05$. Spelling performance was relatively high, but an independent samples t-test revealed no significant difference between the Rime Choice and Reading groups.

Overall performance on the rime choice task was significantly correlated with performance on the Delayed Transfer test, $r(17) = .49, p < .05$, and the spelling posttest, $r(15) = .62, p < .005$, but not with performance on the immediate transfer test, $r(17) = .46, p > .05$. Thus, children who performed better on the rime choice task read and spelled more words correctly following training. This provides evidence for a connection between the rime choice performance and skill transfer. However, since there were no significant group differences on the two transfer posttests, it is likely that the correlation between rime choice performance and posttest performance was driven by individual differences that existed prior to the experiment, and not by skills developed during training.

Table 6. Mean number of words read and standard error (*SE*) on post-test measures. Transfer tests and the spelling test had a maximum possible score of 32; Generalization tests had a maximum possible score of 16.

	Reading Only	Rime Choice
Immediate transfer		
Mean	9.47	6.78
SE	*1.79*	*1.83*
Delayed transfer		
Mean	8.74	6.44
SE	*1.86*	*1.80*
Generalization - words		
Mean	1.89	0.72
SE	*0.62*	*0.28*
Generalization – pseudowords		
Mean	1.47	0.33
SE	*0.65*	*0.18*
Spelling		
Mean	23.70	20.56
SE	*2.20*	*2.06*

Thus, although children's word identification improved during training, children in both groups improved similarly. The inclusion of a modified spelling task which encouraged processing of letter-sound associations did not make children's online word identification nor their ability to transfer word identification skills beyond the training context superior to that of children trained on the shared reading task only.

Discussion

The modified spelling manipulations in these two experiments were designed to test whether training manipulations that focused attention on the letter patterns could offset the poor transfer that has been observed following blocked training. It was suggested that, consistent with Jacoby's (1978) work, this poor transfer may be driven by readers' tendency to use the contextual support of the paradigm to identify the solution without working through (decoding) the problem (words). While such a strategy would permit strong performance in the presence of the contextual support offered by the training paradigm, it could lead to a dependence on that support. This dependence may be due to poorly-formed representations of the letter-sound relations in the trained words, which would result in relatively poor transfer beyond that context. Following such logic, the two experiments described above aimed to increase the probability that children would work through the problem during training with the expectation that this would lead to better skill transfer in the absence of the contextual support which was available during training. That is, requiring children to perform modified spelling tasks was expected to improve their ability to identify words at posttest. The spelling tasks should require processing of the letter-sound correspondences beyond that required in the shared reading task, and should in turn lead to better representations of those correspondences than simply reading.

Contrary to the hypotheses, neither of the modified spelling manipulations improved transfer following training. This contrasts sharply with work by Conrad (2005), which demonstrated that training in spelling led to better reading of trained words than training in reading. It is possible that this contrast is due to differences in the training programs; Conrad's training was not completed in the context of a story. There is evidence to indicate that maximal transfer of word identification skills occurs when training and testing are completed in the same context (Martin-Chang, Levy, & O'Neil, 2007). While Conrad's work involved training and testing of words in isolation, the experiments presented here trained words in the context of a story but tested transfer of word reading using words in isolation. This was also a difference between studies that have demonstrated good transfer following training and testing on flashcards (Levy et al., 1999; Levy & Lysynchuk, 1997) and studies which have demonstrated poor transfer following training on storybooks and testing on flashcards (Levy, 2001). While it could thus be argued that the difference in context between training and test could account for some of the differences seen between these experiments and Conrad's work, pilot work conducted for this dissertation counters such an argument. In that work, children were asked to read the trained words in the context of a blocked story and an unblocked story. Each story was six pages long and contained only one repetition of each target word used during training, printed in boldface type. These books were read using a supported reading paradigm similar to that used during training, but no corrective feedback was given by the experimenter. Thus, if the child read a word incorrectly or failed to read it, the experimenter simply continued with the reading of the book. Despite the fact that children were trained and tested on similar materials, their word reading for unblocked books mirrored that for flashcards, indicating that the patterns persist even when the children are tested in the same type of context in which they have received training.

It is also important to consider that Conrad (2005) used older children and required that they actually write out each word, thus generating the spelling from scratch. There were no constraints on the letters they could choose, so these generated spellings required consideration of many options. The children in the present experiments had alphabetic skills that were too weak to allow them to generate spellings, so the modified spelling tasks which required that they choose from two alternatives were used. These did not require much generation of the spelling. Jacoby's (1978) work has demonstrated retention benefits provided by the generation of a solution during training. It may be that the generation of letters based on sounds is necessary to see the transfer observed by Conrad (2005).

Related to this, it is possible that presenting words in a blocked fashion could have driven the poor transfer following the tasks. Although Conrad (2005) also presented spelling tasks in a blocked fashion and demonstrated good transfer, the fact that she used a traditional (i.e. more challenging) spelling task could have allowed children to overcome the poorly formed representations that typically derive from blocked training. Indeed, Berends and Reitsma (2007) trained Dutch poor readers on an orthographic task similar to the colour choice task used in Experiment 2A that did not block words by orthographic unit and demonstrated good generalization to reading new words following training. It may be that unblocked presentation of the modified spelling tasks in these experiments would have led to better transfer than reading alone.

It is also interesting to note that the children in both experiments performed very well on the modified spelling tasks from the beginning of training. Compared to the reading task, the spelling task was much easier for them to do correctly. It is possible that making the spelling task more difficult (for example, by making it a 4-alternative choice instead of a 2-alternative choice) would require children to process the letter-sound correspondences more thoroughly and would thus lead to better transfer.

Chapter 3

The experiments reported in this dissertation have shown a consistent benefit of rime family training in measures of online word reading, but a loss of that benefit upon transfer to a new context that is not organized by rimes. The manipulations discussed in Chapter 2 were unsuccessful in determining the mechanisms underlying the transient nature of this benefit. It is important to note that the experiments reported here have not been completed with a special population (i.e. poor readers), thus countering any argument that these patterns are due to characteristics specific to poor readers, such as deficits in phonological awareness (Bradley & Bryant, 1983) or verbal working memory (Cormier & Dea, 1997). Instead, this pattern of effects arising from rime family training is pervasive across several samples of children with typical beginning reading skills. Furthermore, this pattern is not only persistent in this dissertation work with the current version of rime family training, but has been seen in a variety of other studies using rime family training (e.g. Bruck & Treiman, 1992; Levy, 2001; Levy et al., 1999).

The present study took a different approach from the earlier experiments described here. The previous experiments strove to determine the mechanisms driving the effects of rime family training by identifying ways to improve transfer of word reading beyond the training context. In contrast, this study focused on determining those mechanisms by investigating whether rime family training was allowing children to develop skills beneficial to word identification that may have been overlooked in earlier studies. This possibility is based again on the idea that children may be using a rhyme analogy (based on sounds) instead of a rime analogy (based on letters) during blocked training (Bruck & Treiman, 1992; Roberts & McDougall, 2003). Underlying this view is the argument that blocking words by rimes may lead children to focus on pronouncing the analogy, instead of decoding the word, since the supportive context provided by the rime families could allow children to predict upcoming rimes without learning the letters associated with those sounds. Such a strategy would lead to precisely the patterns observed with rime family training: good online performance when the contextual support is there, but poor performance when the support of the training materials is removed during transfer tasks. Recall that Landi et al. (2006) proposed a similar idea. They observed poorer retention for words learned in context than for those learned in isolation, and suggested that this pattern could be driven by readers' use of top-down semantic clues to identify words without decoding them. Doing so would lead to a failure to form representations of the letter-sound relations; in the absence of strong representations of the trained material, transfer would suffer substantially.

Applying these ideas to the training materials used here leads to a rather counterintuitive hypothesis. Training programs that encourage the use of orthographic analogies may actually discourage analysis of the analogized unit. Upon removal of the supportive training context, performance on the analogized unit may thus suffer. Indeed, it is possible that instead of seeing good performance on the segment that seems to be emphasized in the analogy, better performance may occur in the segment that is not repeated throughout the analogy. For example, if a child reads the word 'cat', and is then able to read the word 'bat' by making an analogy about the shared rime, one may expect that the child has formed a representation of the /at/ rime, and could apply that to new reading situations. However, according to the reasoning presented above, the child may not have fully processed the rime, and will thus fail to form a strong representation of the letter-sound relations in that rime. Nonetheless, to identify the word the child must still have processed the onset, despite the availability of the rime analogy. The more complete processing of the onset would allow formation of a strong representation of the letter-sound relations for that orthographic unit, while the less complete processing of the rime would lead to a poor representation of the letter-sound relations comprising it. Thus, transfer may be superior for the onset compared to the rime.

Such a hypothesis fits with de Jong and Share's (2007) argument, wherein they explain that Share's self-teaching hypothesis of reading (Share, 1995) requires that children actively decode words in order to form the strong orthographic representations that are recruited in skilled word identification. When active decoding of the printed words is not required, children will be less likely to develop strong representations of letter-sound correspondences for the letter combinations with which they have had experience. The hypothesis is also consistent with Ehri's (1999) model of how children learn to read words by sight. In her model, Ehri proposes that children use phonological skills and knowledge of grapheme-phoneme correspondences to form connections between the spellings and sounds of words. Experience reading words leads to stronger formation of these connections. If children are using a rhyming strategy when reading the rimes of words that are organized into rime families, they may be much less likely to form connections between the letters and sounds within a word, and thus would be expected to show poor abilities to read words containing those letters. In contrast, they would likely form strong letter-sound connections for the letters comprising the onsets (when being trained on rime families) because they would be working through that connection each time they encountered the rime. Indeed, there is work to indicate that training children to read words in a very analytic fashion leads to improvements in reading other words. Ehri and Wilce (1985) demonstrated that children who had received reading training using words that were very similar and so required analysis of each grapheme-phoneme correspondence (e.g. curt, cart, carve) read new words more accurately after training than children who had been trained on grapheme-phoneme correspondences in isolation. Attention during training to each component of the word helped with transfer to new words. Rime family training may very well allow children to identify words during training without attending to the letter-sound relations of the rime component, and thus decrease the probability that they will form connections between the graphemes and phonemes comprising the rime.

This final experiment was designed to investigate the consequences of this possibility. In the strongest form of the argument, children who are trained on words blocked into rime families should form strong representations of trained onsets, but not rimes, since solving the analogy would require only that they decode the onset during training. Conversely, children trained on unblocked material should show no substantial advantage in rime versus onset identification, since neither is especially emphasized by unblocked training. Framing these predictions in terms of this body of work, rime family training would allow benefits to online word identification through benefits to rime identification (via analogy) and would lead to detriments in word identification on tests of transfer through relatively poor representation of rimes (because children are solving the analogy instead of decoding the rime). This argument would lead to the expectation that on tests of transfer following rime family training, 1) onset identification will be equal to or exceed that following unblocked training and 2) rime identification will be lower than that following unblocked training.

These types of effects should not be limited to rime family training; comparable effects should be seen with training programs that block by other orthographic units. For example, in comparison to rime family training, training with onset families should lead to precisely the reverse effect. That is, according to the logic above, children who are trained on material that is organized into onset families should form strong representations of trained rimes but not onsets, since only the onsets can be identified by analogy. An onset family training manipulation was included in the following experiment to allow further investigation of the mechanisms underlying the effects seen in this dissertation.

This experiment adds a novel level of analysis to this area of study. The previous studies in this dissertation, as well as previous work in the field (e.g. Bruck & Treiman, 1992; Goswami, 1986; Lemoine et al., 1993) have used identification of entire words as the measure of children's reading skills. According to the hypothesis laid out above, doing so could lead to an underestimate of the effectiveness of training programs that incorporate orthographic families. Children may have acquired sublexical knowledge that is valuable in learning to read but that is not well-reflected in a whole word measure. Additionally, while there is some work that has compared the effectiveness of training via onset and rime families (e.g. Goswami, 1986, Levy & Lysynchuk, 1997), that work has again focused on children's ability to identify entire words. The inclusion of onset family and rime family training, as well as the analysis at the level of onset and rime identification, allows investigation of different types of training in a more detailed manner than has previously been described.

Method

Participants

The sample comprised 60 children (24 boys, 36 girls, mean age = 5 years, 9 months) in Senior Kindergarten classes at four local schools which were located in upper-middle class neighbourhoods. Children were randomly assigned to the Rime Families training condition, the Onset Families training condition, or the Unblocked condition. All children who had at least a Kindergarten level of letter knowledge (assessed using the WRAT-3; Wilkinson, 1993) and could read fewer than six of the training words were included in the study. The experiment was completed in the Winter and Spring terms.

Materials

Pretests. The Test of Auditory Analysis Skills (TAAS) (Rosner, 1999), a syllable and phoneme deletion task, was used as a measure of phonological awareness. The Blue form of the Wide Range Achievement Test – Third Edition (WRAT-3) (Wilkinson, 1993) was used as a measure of letter knowledge and word reading ability. This test requires children to name 15 letters and to read a set of words. Knowledge of the training words was tested using 32 flashcards, each of which contained a single training word printed in black ink. These flashcards were presented in random order, and were not organized into onset or rime families. Three subtests from the Pre-Reading Inventory of Phonological Awareness (PIPA) (Dodd, Crosbie, McIntosh, Teitzel, & Ozanne, 2003) were used to assess children's letter-sound knowledge, rime awareness, and onset awareness. On the latter two subtests, children must identify the one item from a set of four that has a different onset or rime, respectively, from the other items in the set.

Training. The word set contained eight different four-word rime families and eight different four-word onset families. Orthogonal combination of these onsets and rimes led to 32 training words that could be organized into 4-item onset or rime families, or presented in sets such that no words shared an onset or a rime. The training materials were 32 flashcards with these words printed using Levy's visible orthography. Appendix E contains the words. That is, for rime family training, each rime was printed in coloured ink and each onset was printed in black ink. Rimes were systematically colour-highlighted, such that the different rimes were printed in different colours, but each particular rime was always printed in the same colour. Unblocked training was also completed using these cards. For onset family training, the onsets were highlighted systematically, while the rimes were printed in black ink.

Posttests. Children were tested on their ability to transfer skills beyond the training context immediately after the final training session and one week later. Both tests of transfer to a new context were completed using flashcards with the training words printed in black ink and presented one at a time in random order, not in onset or rime families. Generalization to new material was tested using 16 flashcards containing new words with the trained onsets or rimes and 16 flashcards containing pseudowords with the trained onsets or rimes. Again, all items were printed in black ink and presented one at a time in random order.

Procedure

Sessions were completed one-on-one with the experimenter in a room at the children's school, during regular school hours. Because the children attend school on alternate days, training sessions were completed on consecutive school days, but not consecutive weekdays (as in the previous studies).

On the first day, each child was asked to read the 32 training words from flashcards, printed in black ink. Any child who could read six or more words was not included in the study. Each child also completed the TAAS, the three PIPA subtests, and the WRAT-3; children with scores lower than a kindergarten level on the WRAT-3 were not included in the study. The experimenter modeled reading the training words for the child, with the words organized according to the child's training condition (described below).

Over the course of five days, children completed either Onset Families, Rime Families, or Unblocked training. For the Onset Families condition, flashcards were presented four at a time in onset families, such that each of the four words in a family shared a common onset, but had a different rime. For the Rime Families condition, flashcards were presented four at a time in rime families, such that each of the four words in a family shared a common rime, but had a different onset. For the Unblocked condition, flashcards were presented four at a time such that none of the words in a set shared an onset or a rime. The order of the 4-word sets was randomized on each trial. Children in all conditions read all 32 training words twice each day, so that by the end of training, children had 10 training trials with each word. This was only half as many trials as were completed in previous studies, but this was sufficient to see improvement over the course of training on all measures (see analyses below). As in the previous experiments in this thesis, children read all words at the whole-word level, and any errors that they made were corrected at the whole-word level.

For posttests, words were printed in black ink on flashcards and were presented one at a time in random order. Immediately following the final training session, children were given the test of immediate transfer to a new context, wherein they were asked to read the 32 training words to determine whether they could read the words without any highlighting or organization by orthographic unit. One week following the final training session, children completed a test of delayed transfer that was identical to the immediate transfer test. One day after the final training session, children completed a test of generalization to new words and pseudowords that contained the trained onsets or rimes. Performance for this test was extremely low and was not included in the main analysis.

Crucially, in contrast to the previous studies in this dissertation, the exact response of each child was written down by the experimenter during all pretest, training, and posttest trials. This allowed evaluation not only of the children's ability to identify entire words, but also subunits in those words (i.e. onsets and rimes).

Results

For pretest, online (training), and post-test measures, performance was assessed first at the level of word identification and then at the level of rime identification and onset identification. Measures of rime and onset identification were not independent of measures of word identification: correct identification of a word would count as a correct word identity, rime identity, and onset identity. However, use of rime and onset identification scores allowed for analysis of performance beyond trials wherein children correctly identified entire words; trials wherein children correctly identified only *part* of a word (e.g. correct identification of the onset but not the rime, or vice versa) were now also included in the analyses. For example, if a child correctly identified the word 'cat', this would lead to a score of correct word, correct rime, and correct onset. However, if the child identified the word 'cat' as 'car', he would obtain a score of incorrect word, incorrect rime, and correct onset. In this way, scoring of identification of the two subunits was correlated with scoring of word identification. The onset identification measure gives the number of onsets correctly identified regardless of whether the word's rime was correctly identified, and the rime identification measures give the number of rimes correctly identified regardless of whether the word's onset was correctly identified. As described above, analysis at this level has important theoretical implications for the effects of training programs that emphasize orthographic families.

As in the previous chapters, the assumption of sphericity that is required for repeated measures was frequently violated. Again, the Greenhouse-Geisser correction was used to avoid for inflation of Type I error rates. Recall that this correction provides a more conservative test of the null hypothesis by modifying the degrees of freedom on which the hypothesis is evaluated. Again, for simplicity of reading, if the p value was identical for the standard and Greenhouse-Geisser test, the standard degrees of freedom are reported (Baguley, 2004).

Pretest performance

One-way ANOVAs indicated no significant differences between the groups on any of the pretest measures ($ps > .30$; see Table 7).

Table 7. Mean scores and ranges of groups on pretest measures.

	Rime families	Onset families	Unblocked
N	20	20	20
Age (months)			
Mean	69.0	68.7	68.0
Range	61-74	63-77	60-74
Words identified pretest			
Mean	1.1	1.0	1.1
Range	0-5	0-3	0-5
Onsets identified pretest			
Mean	8.1	4.5	7.6
Range	0-28	0-17	0-23
Rimes identified pretest			
Mean	1.4	1.2	1.5
Range	0-5	0-3	0-7
WRAT-3 standard score			
Mean	107.7	106.1	107.1
Range	99-122	95-120	93-119
TAAS score			
Mean	3.2	3.1	3.0
Range	0-8	0-9	0-8
PIPA-RA percentile score			
Mean	43.5	35.0	33.5
Range	7-97	1-77	7-77
PIPA-OA percentile score			
Mean	27.3	31.8	38.5
Range	2-77	2-77	2-87
PIPA-LSK percentile score			
Mean	39.5	36.3	41.5
Range	12-87	7-97	2-97

Note. Number of onsets and rimes identified on pretest was not independent of number of words identified. The onset and rime measures indicate the number of components correctly identified regardless of whether the entire word was correctly identified. WRAT-3 = Wide Range Achievement Test, 3rd edition; TAAS = Test of Auditory Analysis Skills; PIPA = Pre-Reading Inventory of Phonological Awareness – Rime Awareness, Onset Awareness, Letter-Sound Knowledge. Maximum possible standard score on WRAT-3 is 155. Maximum possible score on TAAS is 13. Maximum possible percentile score on PIPA subtests is 99.

Online performance

Online performance was measured as the number of units (words, rimes, or onsets) correctly identified on each trial, where each child completed two trials per day (two repetitions of all 32 flashcards) with each target word. Each child completed five training sessions, comprising 10 trials per target word over the course of training. For each measure (word, rime, and onset identification), there was a maximum possible score of 32 (8 rimes, 8 onsets, 4 instances of each rime and each onset) on each trial. The data for each unit were analysed using a 3 (Training Condition) x 10 (Trials) mixed design ANOVA.

For word identification, this analysis revealed a main effect of Trials, $F(9, 513) = 65.99$, $MSE = 9.70$, $p < .001$, indicating that performance improved across training (see Figure 4). There was also a main effect of training condition, $F(2, 57) = 7.00$, $MSE = 370.27$, $p < .005$, but no significant interaction between training condition and trials ($p > .90$). Post hoc analyses were completed using Dunnett's t, and indicated that Rime Family training led to significantly better word identification than Unblocked training (mean difference = 6.24, $p < .005$) and Onset Family training (mean difference = 6.23, $p < .001$). However, there was no significant difference between Onset Family and Unblocked training ($p > .66$).

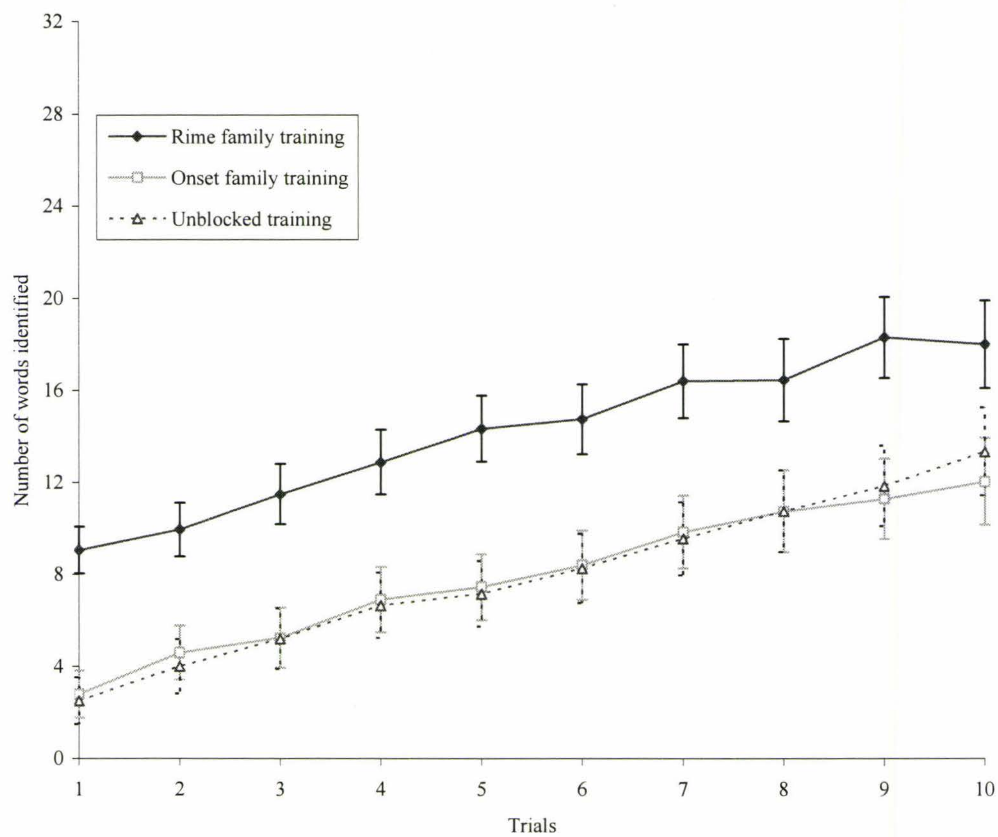


Figure 4. Mean number of words correctly identified over the course of training (SE bars). Children completed two trials/day.

For onset identification, analyses revealed a main effect of Trials, $F(9, 513) = 38.96$, $MSE = 10.95$, $p < .001$, indicating that performance on this measure also improved across training. There was no significant main effect of Training Condition and no interaction (*p*s>.19; see Figure 5).

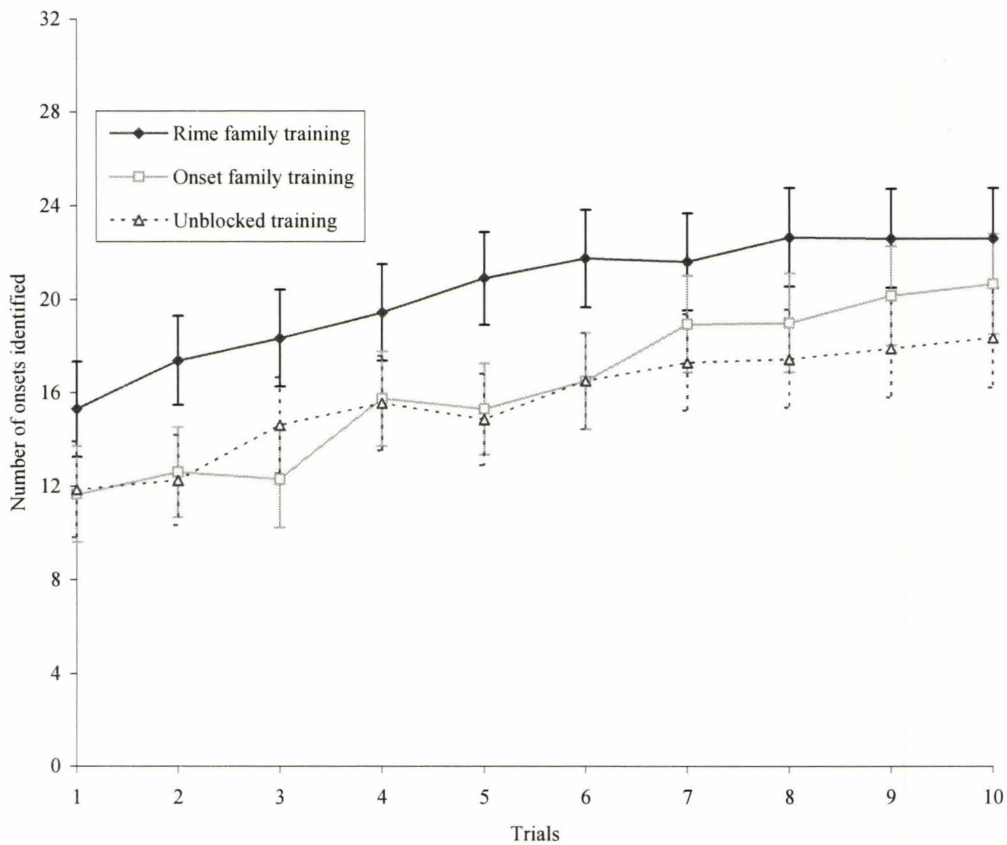


Figure 5. Mean number of onsets correctly identified over the course of training (SE bars). Children completed two trials/day.

For rime identification, analyses revealed a main effect of Trials, $F(9, 513) = 72.10$, $MSE = 10.04$, $p < .001$, indicating again that performance improved across training. There was also a main effect of training condition, $F(2, 57) = 26.40$, $MSE = 240.06$, $p < .001$, but no significant interaction between training condition and trials ($p > .80$; see Figure 6). Post hoc analyses were completed using Dunnett's t, and indicated that Rime Family training led to significantly better rime identification than Unblocked training (mean difference = 8.57, $p < .001$) and Onset Family training (mean difference = 10.61, $p < .001$), but that there was no significant difference between Onset Family and Unblocked training ($p > .96$). This was the same pattern observed on the word identification measure.

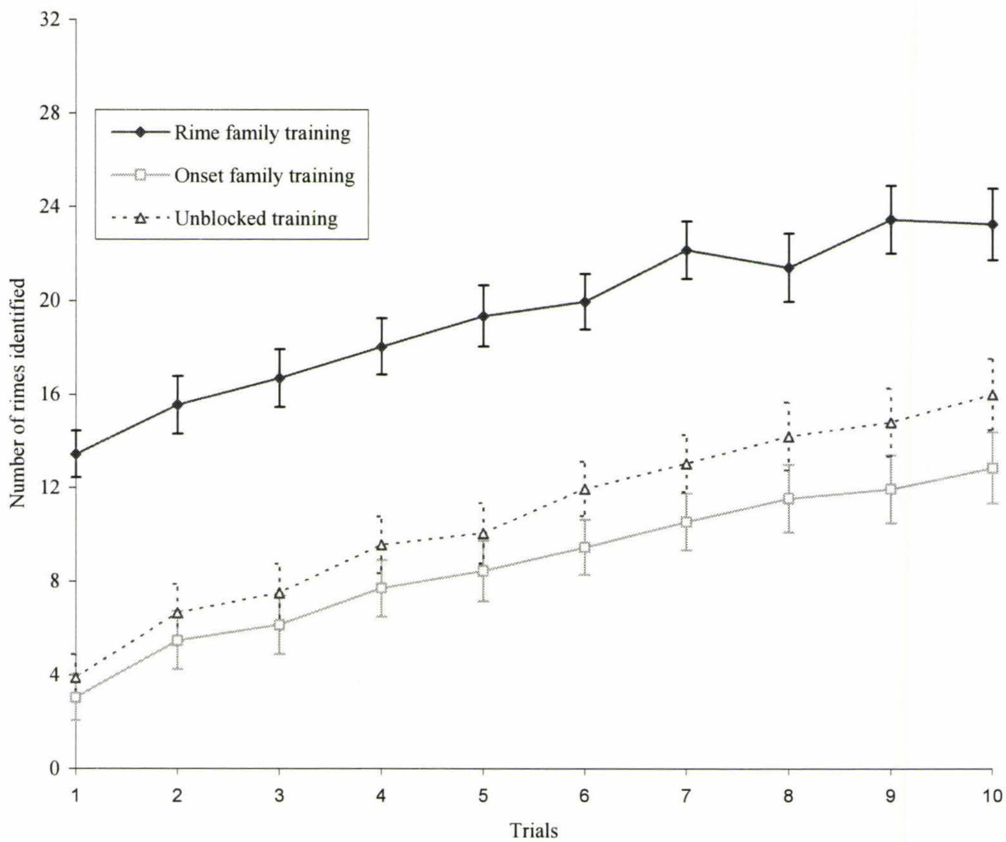


Figure 6. Mean number of rimes correctly identified over the course of training (SE bars). Children completed two trials/day.

Posttest performance

Again, the data for words, onsets, and rimes were analyzed in three separate analyses. Mean group performance on each posttest measure is presented for each orthographic unit (words, onsets, rimes) in Table 8. The data were analyzed using three 2 (posttest) x 3 (training condition) mixed design ANOVAs.

Table 8. Mean scores *(and standard errors)* of groups on posttest measures.

	Rime families	Onset families	Unblocked
Words			
Immediate Transfer	9.30 *(1.55)*	9.25 *(1.57)*	13.50 *(2.02)*
Delayed Transfer	7.70 *(1.56)*	9.80 *(1.81)*	12.00 *(1.90)*
Onsets			
Immediate Transfer	21.90 *(2.33)*	20.40 *(1.86)*	20.25 *(2.47)*
Delayed Transfer	22.05 *(2.16)*	20.05 *(2.00)*	20.95 *(2.36)*
Rimes			
Immediate Transfer	10.75 *(1.48)*	10.55 *(1.57)*	16.20 *(1.75)*
Delayed Transfer	8.75 *(1.56)*	11.20 *(1.78)*	14.15 *(1.84)*

On the measure of word identification, this analysis revealed no significant main effect of training condition and no interaction between training condition and posttest (*p*s>.16). Thus, consistent with the work presented in earlier chapters, the benefits of the Rime Family training over Unblocked training did not persist on posttests, nor did the benefits over Onset Family training persist. Furthermore, there was no significant effect of posttests (*p*>.10); children performed similarly on the immediate and delayed tests of transfer to a new context.

As can be seen in Table 8, children identified substantially more onsets than words, but aside from this difference, performance on the onset identification measure was similar to that on the word identification measure. The ANOVA revealed no significant main effects and no significant interaction (*p*s > .76). This is consistent with the patterns seen during training.

Finally, on the rime identification measure there was a significant effect of posttest ($F(1, 57) = 4.72$, $MSE = 8.17$, $p < .05$), indicating that children performed better on the test of immediate transfer than on the test of delayed transfer. There was also a significant effect of training condition ($F(2, 57) = 3.18$, $MSE = 103.05$, $p < .05$). Crucially, posthoc tests revealed that this was driven by a significant difference between the Rime Family and Unblocked conditions, however contrary to online performance, this was such that the Unblocked group outperformed the Rime Family group (mean difference = 5.43, $p <.02$). Thus, the benefit to rime identification during training was reversed on posttests, indicating that while Rime Family training benefited rime identification during training, it led to a detriment in rime identification beyond the training context.

Within-group comparisons

Finally, analyses were done to determine whether the number of onsets and rimes identified differed within each training group in a similar way for online and transfer measures. This was a crucial comparison, as it would allow direct testing of the effects of each type of training within individual groups. The measure of interest here was a comparison of the major trends between training and posttest performance in terms of onset versus rime identification. Thus, difference scores were calculated to measure the difference between rime and onset identification on the immediate and delayed transfer tests. A difference score for rime and onset identification was also calculated from the average number of each orthographic unit identified over the entire course of training (i.e. this was the difference between the average number of rimes identified over 10 training trials and the average number of onsets identified over 10 training trials). Collapsing online performance over trials in this way allowed for a more concise analysis of differences between online and transfer performance.

All groups showed a benefit to onset identification over rime identification on pretests. Thus, the difference between rime identification and onset identification on pretest was used as a covariate in a 3 x 3 mixed design ANCOVA comparing difference scores (rime identification-onset identification) for the three training groups on measures of online performance, immediate transfer, and delayed transfer. This revealed a significant main effect of test, $F(2,112) = 40.52$, $MSE = 14.00$, $p < .001$, a main effect of training condition, $F(2,56) = 9.71$, $MSE = 94.62$, $p < .05$, and an interaction, $F(2,112) = 9.98$, $MSE = 14.00$, $p < .001$ (see Table 9).

Table 9. Adjusted mean difference scores (rime identification - onset identification) for the three training groups at three times.

	Rime families	Onset families	Unblocked
Online	-0.21	-8.73	-4.45
Immediate transfer	-10.84	-10.34	-3.86
Delayed transfer	-13.12	-9.16	-6.68

Note. Difference scores were calculated as Rime identification-Onset identification. Negative numbers denote a benefit of onset identification over rime identification. Numbers near 0 denote similar rates of identification for rimes and onsets. Adjusted means were calculated using Pretest difference score = -5.35.

All three training groups showed better onset identification than rime identification on both tests of transfer to a new context. Children who completed Onset Family or Unblocked training also showed better onset identification than rime identification during training, however, as is indicated by their small difference score, children trained on Rime Families did not.

Recalling that all three groups identified similar numbers of onsets on tests of transfer (see Table 6), these patterns are again consistent with the argument that training on rime families inflates children's ability to identify words online by increasing their ability to identify rimes. When the training support is removed, rime identification ability drops.

Discussion

This experiment replicated findings from the earlier studies, and clarified some of the mechanisms underlying the patterns of performance seen during and after orthographic family training. To begin, the benefits to online word identification observed for rime family training relative to unblocked training replicated the findings of experiments in Chapters 2 and 3, as well as Levy's work on this issue (e.g. Levy, 2001). The finding that those benefits were lost upon removal of the supportive training materials also replicated the earlier studies in this document, as well as work by others (e.g. Bruck & Treiman, 1992; Levy, 2001; Levy et al., 1999). Together, these findings indicate that the transient benefits to word identification of rime family training are consistently experienced by beginning non-readers. Although measures of online word identification seem to indicate substantial benefits of Rime Family training, when asked to read the trained words in a new context, children trained on rime families perform no better than children trained on unblocked material.

The analyses at the level of onset and rime identification clarified some of the mechanisms underlying this robust finding. These analyses indicated that the benefits observed during Rime Family training are driven by benefits to online rime identification, but not online onset identification. This pattern is consistent with the hypothesis laid out earlier in this chapter. When children are asked to identify words that are organized by rime families, their ability to identify rimes is substantially better than if those words are presented in a way that makes the rimes less predictable (in onset families or unblocked). Presumably, it is the availability of an analogy that allows children this benefit. However, solving these analogies during training leads to no lasting benefit to word identification, and in fact led to a lasting detriment to rime identification, relative to Unblocked training. This is consistent with Roberts and McDougall's (2003) proposal that rime family training leads children to make rhyme (sound-based) analogies, not rime (letter-based) analogies. In turn, this supports the arguments (e.g. de Jong & Share, 2007; Landi et al., 2006) that active decoding of words at the level of grapheme-phoneme correspondences is the driving force for the formation of strong representations of letter-sound relations, and Ehri's (1999) model which suggests that connections between graphemes and phonemes must be formed before children can read words by sight.

Comparisons of onset and rime identification by children trained on rime families and unblocked material thus present data that are entirely consistent with the argument that solving an analogy for a particular unit leads to a relatively poor representation of that unit. In contrast, the patterns exhibited by children trained on onset families are only partially consistent with the hypothesis. That is, the fact that Onset Family training does not benefit online rime identification is consistent with the hypothesis, as rimes cannot be identified by analogy during Onset Family training, and so should be more difficult to identify than during Rime Family training. However, the fact that the three training conditions did not differ in terms of online onset identification indicates that there is indeed some qualitative difference in the way in which children go about identifying onsets and rimes. All groups improved significantly in onset and rime identification over the course of training, but despite the availability of analogies for this unit during Onset Family training, that group showed no significant benefit to online onset identification, and no detriment to onset identification when the support of the analogies was removed.

Several factors could play a role in this effect. One possibility is that onset and rime identification involve qualitatively different strategies. For example, schoolchildren are often taught to read words by identifying the first sound and then filling in possible endings. Research shows that beginning readers reliably use this strategy (Bowman & Treiman, 2002). If children in this experiment are using this strategy, it could lead to an overall benefit to onset identification relative to rime identification, regardless of the training program. That is, children in all three training conditions would actively decode onsets during training, but only children receiving training on onset families and unblocked words would actively decode rimes. Such a discrepancy could explain the differential transfer in onset and rime identification following training on analogies.

Furthermore, all rimes used in this experiment comprised two letters, while all onsets comprised only one letter. This relative simplicity of onsets, paired with onset-based word identification strategies they have learned in the classroom, could have encouraged or allowed the children to process the onsets, even if they were available through analogy. Such a strategy would explain the high onset identification performance on posttests for all groups, and would indicate that children from all three training conditions formed strong representations of the onsets used in this experiment. Further work could include complex onsets (two letters) to investigate whether this leads to differences between the groups that are more consistent with the original hypotheses.

Finally, there has been some work arguing that onsets are less salient to children than rimes (Goswami, 1986). She demonstrated that children who were trained to use rime analogies developed better word reading skills than children who were trained to use onset analogies. It could be that blocking by rimes makes the shared unit much more phonologically salient than does blocking by onsets, and that this increased salience leads to more use of sound-based analogies in rime family training than in onset family training.

One finding that appears puzzling at first glance is the fact that children trained on onset families and rime families identified the same number of rimes on posttests. In its strongest form, our hypothesis would predict that children trained on onset families would identify more rimes than children trained on rime families, as Onset Family training requires active decoding of rimes, while Rime Family training allows rime identification through analogy. Analysis of within-group effects clarifies this finding. Independent samples t-tests reveal a non-significant trend for the benefits to onset identification (over rime identification) to be greater in the Rime Family group than in the Onset Family group. Thus, children trained on rime families did show relatively more benefits to onset identification than children trained on onset families. This provides some support for the argument that blocking by any orthographic unit will lead to benefits for the unit that is not available through analogy. A larger sample size which would permit increases to statistical power may allow more convincing data in support of this conclusion.

General Discussion

This set of experiments explored how manipulating the visibility of spelling patterns in English print, without concurrent oral segmentation, influences the word identification skills of beginning non-readers. Experiment 1 investigated the effects of variations in emphasis on the orthographic patterns in print on these young readers' word reading skills. Levy's (2001) visible orthography was used, which implemented visible segmentation of onsets and rimes in printed words. Previous work indicated that training programs employing rime families benefit children's online word identification performance, but that those benefits can become unreliable beyond the support of the training context. This effect had been demonstrated with beginning readers who were trained using visual and oral segmentation together (Bruck & Treiman, 1992) and also with older poor readers who were trained using visual segmentation, without concurrent oral segmentation (Levy, 2001). Experiment 1 replicated the disparity between online performance and transfer to new contexts. First, it demonstrated that when compared with unblocked training (where rimes were distributed across word sets), blocked training (i.e. rime family training) led to more rapid improvements and ultimate superiority in online performance. Secondly, it demonstrated that transfer to a new context was poor following blocked training, revealing the same cost to transfer that had been demonstrated by Levy's (2001) poor readers. Importantly, these effects were observed in the absence of concurrent oral segmentation. Bruck and Treiman's (1992) training program that led to similar transfer failure had involved segmentation at the level of phonology and orthography.

In addition to replicating previous findings, Experiment 1 revealed the transient nature of blocked training benefits when training was completed in the context of a story, thus broadening the range of reading settings in which these effects have been observed. Previous work involving manipulations that make the orthography 'visible' employed training of words in isolation, with a notable exception: Levy (2001) embedded the visible orthography in age-appropriate storybooks. However, her training required that children complete a shared reading activity with the same storybook five times consecutively. This procedure raised questions about whether improvements in children's online word identification were due to improvements in word identification per se or due to memorization of the text. In Experiment 1 described here, the visible orthography was presented in the context of age-appropriate storybooks, but each child engaged in shared reading with a different storybook in each training session. The use of different stories each day decreased the likelihood that children's improvements in word identification were due to rote memorization of the stories themselves. Furthermore, this allowed a training regime that was both palatable to the children and representative of the typical manner in which they may be exposed to new words, thus lending ecological validity to the work.

Thus, Experiment 1 established that the benefits of rime family training to online word identification are robust, demonstrating them in a new population and in the absence of oral segmentation. This ascertained that improvements in beginning non-readers' online word identification can indeed be influenced by orthographic manipulations alone, providing a clear example of the importance of orthographic skills in early reading acquisition. The experiment also demonstrated that the disparity between online and transfer performance seen with rime family training is robust, as the effects of the orthographic manipulations did not persist beyond the training context. Crucially, these effects were observed using materials that are typically used in classrooms and by parents. The mechanisms underlying these transfer failures merited investigation, specifically to determine whether the online benefits of rime family training could be prolonged so that young readers could generalize their skills to new reading situations.

One explanation that has been proposed to account for the poor transfer following rime family training is that children fail to fully process the letter-sound relations in the rime and instead generate the rime from memory of prior word naming within the training block (e.g. Bruck & Treiman, 1992; Roberts & McDougall, 2003). Such deficient processing of the orthography would lead to poorly formed representations of the letter-sound relations (Greene, 1989), despite allowing correct identification of the rimes in the context of the orally predictable training material. Indeed, similar disparities in word identification between online and transfer measures were reported by Landi et al. (2006) using a semantically predictable context, and they argued that a similar mechanism (poor representations of letter-sound relations deriving from predictable training material) could explain their effects. In a related vein, Levy and Lysynchuk (1997) found that when children reached a criterion of perfect performance by the end of blocked training, they maintained good word reading beyond the training context. They argued that until children have reached this criterion, they are less likely to have abstracted a strong representation of the letter-sound relations in the words, so their learning is quite unstable. When the support of the training material is removed before they reach consistently perfect performance, that lack of stability becomes apparent and children demonstrate poor transfer. To date there have been no empirical attempts to encourage more complete processing during rime family training. If the hypothesis of incomplete processing is true, manipulations that lead to more thorough processing of the components of the words during training should lead to improved transfer following training that blocks together words sharing rimes. Following this hypothesis, Experiments 2A and 2B investigated ways to maintain the online benefits of rime family training, allowing more permanent gains in word identification.

Experiment 2A investigated whether including a modified spelling task that directed attention to the visual patterns in the print improved transfer following training on a blocked shared reading paradigm. It is important to note that this spelling task required that children actively attend to the letters in the rimes, but it did not explicitly draw their attention to the corresponding sounds; it was essentially a visual matching-to-sample task. Experiment 2B implemented a modified spelling task that more closely approximated a traditional spelling task in that it directed children's attention to the letter-sound relations in the words. As in a traditional spelling task, in order to complete this task, children needed to know which letters would make the sounds they heard in a word. This better processing should lead to the formation of stronger representations of the trained rimes and allow better transfer beyond the training context. However, in contrast to previous work (Conrad, 2005) that demonstrated more accurate reading of trained words following spelling training compared with reading training, the spelling manipulations did not lead to improved transfer in either of these experiments.

These failures in transfer can be interpreted in terms of a generation effect which leads to more extensive processing and better memory for the item (Jacoby, 1978). It is possible that previously-observed benefits of spelling to reading derive from the generation of a response choice from many options, as was necessary in Conrad's (2005) work using a spelling task where children were required to generate spelling from sounds. Children in Experiments 2A and 2B reported here were given two alternatives from which to choose; the task did not require much generation of the spelling. The fact that their performance on the spelling tasks improved rapidly and was near ceiling for most of the training trials is consistent with this argument: if the task was too simple for the children, it would not lead to additional processing during training, and thus would fail to influence transfer.

The mechanism underlying the generation effect (Jacoby, 1978) is that, compared to reading a response, "solving the problem" to generate the correct response requires additional cognitive processing. Blocked and unblocked training can be compared in terms of that same mechanism. Blocked training allows children to solve the problem and provide the correct spelling or pronunciation without generating or decoding the response from the information on the oral or visual input alone. In contrast, unblocked training forces children to solve the problem by decoding the word from the letters on the page. Thus, the fact that the words in the modified spelling tasks were presented in a blocked fashion could explain the persistence of poor transfer seen in these experiments. Consistent with this, training on an orthographic task similar to that used in Experiment 2A but that presented words in an unblocked fashion has been shown to lead to improved generalization to new words with poor readers (Berends & Reitsma, 2007). Furthermore, although Conrad (2005) presented spelling tasks in a blocked fashion and obtained good transfer, her use of a more challenging spelling task could have allowed children to overcome the poorly formed representations that typically derive from blocked training.

One remaining possibility for the poor transfer seen in this dissertation was that training and testing were completed in different contexts: while children were trained to read the words in the context of storybooks, they were tested using flashcards. Other work in the field had found effects that support this possibility: good transfer was observed following training and testing on flashcards (Levy et al., 1999; Levy & Lysynchuk, 1997) while poor transfer was observed following training on storybooks and testing on flashcards (Levy, 2001). Furthermore, Martin-Chang et al. (2007) demonstrated that transfer (measured as reading fluency) is maximized when children are trained and tested in the same context. To investigate the possibility that contextual differences between training and test were driving poor transfer, Experiment 3 implemented training and testing on flashcards. Despite this, the disparity between online and transfer word identification persisted.

Experiment 3 also involved a different approach to the previous experiments. The previous experiments had attempted to determine the mechanisms underlying poor transfer following blocked training by improving transfer. This final experiment focused instead on determining those mechanisms by ascertaining precisely what children learn from rime family training. A notable pattern in the previous experiments was that children would often correctly read a portion of some words on the transfer tests. Because the variable of interest in those experiments was whole-word reading, such responses were scored as being incorrect. For this final experiment, children were still scored on their ability to identify whole words; this ascertained that word identification followed similar patterns to those observed in the previous studies despite the change from storybook to flashcard training. However, children were also scored on their ability to correctly identify onsets and rimes, regardless of whether they correctly identified the entire word. This allowed a more precise measure of what children were learning from the different training programs.

Consistent with the previous experiments and work by others (Bruck & Treiman, 1992; Levy, 2001), word identification was better during Rime Family training than during Unblocked training, but this benefit disappeared on tests of transfer to a new context. This indicated again that the transient benefits to word identification of Rime Family training are consistently experienced by beginning non-readers. Analyses at the level of onset and rime identification clarified the mechanisms driving this effect. When children were trained with Rime Families, the benefits and detriments to rime identification mirrored those for word identification: online rime identification benefited, but transfer suffered. This contrasted with performance following Unblocked training, indicating that the online benefits were likely driven by the availability of the rhyme analogy during training. Consistent with arguments that active decoding of words is the driving force for the formation of strong representations of letter-sound relations (e.g. de Jong & Share, 2007; Landi et al., 2006), use of this analogy led to no lasting benefit to rime identification, and in fact led to a detriment relative to Unblocked training. Although others had suggested that transfer failures following Rime Family training could be due to children's tendency to use a rhyme analogy during training instead of a rime analogy (e.g. Bruck & Treiman, 1992; Roberts & McDougall, 2003), this was the first empirical work supporting this hypothesis.

In contrast to the evidence that children make use of the analogies available in Rime Family training, Experiment 3 revealed no evidence that they make use of analogies in Onset Family training. On tests of transfer, children trained with onset families demonstrated similar word identification to children trained with unblocked words and thus offered no opportunity for analogies. Furthermore, there were no differences among the three groups in terms of their online onset identification. These patterns indicate that there is some qualitative difference in the way that children approach onset and rime identification. As mentioned in Chapter 4, it is possible that classroom-taught strategies that focus on identifying the beginning sound of a word lead children use this approach consistently, regardless of whether an analogy for that beginning sound is available. This would explain the similar online onset identification between the three groups, and would also explain the relatively strong transfer on the onset identification measure, since children in all three groups were actively attending to the letter-sound associations for the onsets during training.

It is important to interpret these data in terms of the generalizability of the effects of blocked training. According to the argument that emphasizing an orthographic unit through blocking leads to less complete processing of that unit than unblocked training, unblocked training should lead to better identification of a unit than when it was blocked. That is, unblocked training should lead to better rime identification than rime family training and better onset identification than onset family training. This hypothesis was partially supported: relative to unblocked training, Rime Family training led to poor transfer of rimes. However, comparisons of transfer following Onset Family and Unblocked training did not show a benefit to onset identification following Unblocked training. This discrepancy merits future investigation. The onsets used in these experiments were all single-letter onsets, while the rimes were all comprised of two letters. The fact that there is less orthographic overlap between words sharing an onset (e.g. bat, bug) and words sharing a rime (e.g. bat, sat) may make children less likely to note the availability of an analogy in the former situation. Future work could investigate whether, relative to unblocked training, training that blocks by complex (i.e. two-letter) onsets leads to detriments to onset identification. Such detriments would indicate that the effects of blocking on the processing of grapheme-phoneme correspondences during training generalize across various orthographic units. In contrast, failure to observe such detriments would indicate that children approach identification of onsets and rimes in qualitatively different ways, and work to decode even complex onsets when an analogy is available.

Conclusion

This set of experiments has provided several contributions to the literature. First, the effects of training with rime families have been extended to a new population and the effects of a visible orthography in the absence of phonological segmentation have been replicated. Secondly, the need for substantial increases in processing to prevent transfer failures following rime family training have been implied by the null results reported in Chapter 3. Finally, the mechanisms underlying transfer failures typically observed following rime family training have been elucidated as failures to transfer the emphasized rime unit. Further work can determine the generalizability of this effect through training programs that include complex onsets.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Adams, Treiman, & Pressley (1998). Reading, writing, and literacy. In W. Damon, I. E. Sigel, & K. A. Renninger (Eds.), *Handbook of child psychology, 5th ed. Vol 4: Child psychology in practice* (pp. 275-355). Hoboken, NJ: John Wiley & Sons Inc.

Baddeley, A. D., Ellis, N. C., Miles, T. R., & Lewis, V. J. (1982). Developmental and acquired dyslexia: A comparison. *Cognition, 11,* 185-199.

Baguley, T. (2004). An introduction to sphericity. Retrieved June 22, 2009 from: http://homepages.gold.ac.uk/aphome/spheric.html.

Berends, I. E., & Reitsma, P. (2007). Orthographic analysis of words during fluency training promotes reading of new similar words. *Journal of Research in Reading, 30,* 129-139.

Berninger, V. W., Abbott, R. D., Brooksher, R., Lemos, Z., Ogier, S., Zook, D., et al. (2000). A connectionist approach to making the predictability of English orthography explicit to at-risk beginning readers: Evidence for alternative, effective strategies. *Developmental Neuropsychology, 17,* 241-271.

Biemiller (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly, 6,* 75-96.

Bishop, D., Adams, C., Lehtonen, A., & Rosen, S. (2003). Effectiveness of computerised spelling training in children with language impairments: A comparison of modified and unmodified speech input. *Journal of Research in Reading, 28,* 144-157.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95,* 153-193.

Boscardin, C. K., Muthen, B., Francis, D. J., & Baker, E. L. (2008). Early identification of reading difficulties using heterogenous developmental trajectories. *Journal of Educational Psychology, 100,* 192-208.

Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing, 5,* 69-85.

Bowman, M., & Treiman, R. (2002). Relating print and speech: The effects of letter names and word position on reading and spelling performance. *Journal of Experimental Child Psychology, 82,* 305-340.

Bradley, L., & Bryant, P.E. (1985). *Rhyme and reason in reading and spelling.* Ann Arbor, MI: University of Michigan Press.

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

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

Brown, G. D. A., & Deavers, R. P. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Child Psychology, 73,* 208-242.

Brown, I., & Felton, R. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading and Writing: An Interdisciplinary Journal, 2*, 223-241.

Bruck, M., & Treiman, R. (1992). Learning to pronounce words: The limitations of analogies. *Reading Research Quarterly, 27,* 374-388.

Bus, A. G., & van IJzendoorn, M H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91,* 403-414.

Cardoso-Martins, C., & Pennington, B. F. (2004). The relationship between phoneme awareness and rapid serial naming skills and literacy acquisition: The role of developmental period and reading ability. *Scientific Studies of Reading, 8,* 27-52.

Carlson, R. A., & Yaure, R. G. (1990). Practice schedules and the use of component skills in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 484-496.

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

Chall, J. S. (1983). *Learning to read: The great debate (Updated edition).* New York, NY: McGraw-Hill.

Chall, J. S. (1989). From language to reading and reading to language. *American Speech Language Hearing Association Reports Series, 17,* 28-33.

Christenson, C. A. (1997). Onset, rhymes, and phonemes in learning to read. *Scientific Studies of Reading, 1,* 341-358.

Christenson, C. A., & Bowey, J. A. (2005). The efficacy of orthographic rime, grapheme-phoneme correspondence, and implicit phonics approaches to teaching decoding skills. *Scientific Studies of Reading, 9,* 327-349.

Conrad, N. J. (2005, June). Examining the relation between reading and spelling:

A training study. Poster session presented at the annual meeting of the Society for the Scientific Study of Reading, Toronto, ON.

Cormier, P., & Dea, S. (1997). Distinctive patterns of relationship of phonological awareness and working memory with reading development. *Reading and Writing: An Interdisciplinary Journal, 9,* 193-206.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior, 21,* 451-467.

de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading, 11,* 55-71.

Dempster, F. N. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79,* 162-170.

Dodd, B., Crosbie, S., McIntosh, B., Teitzel, T., & Ozanne, A. (2003). Pre-Reading Inventory of Phonological Awareness (PIPA). San Antonio, TX: Pearson Education, Inc.

Durgunoglu, A., Y., Mir, M., & Arino-Marti, S. (1993). Effects of repeated readings on bilingual and monolingual memory for text. *Contemporary Educational Psychology, 18,* 294-317.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9,* 167-188.

Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), Reading development and the teaching of reading: A psychological perspective. (pp. 79–108). Oxford: Blackwell Science.

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,* 250-287.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71,* 393-447.

Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27,* 13-26.

Ehri, L., C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20,* 163-179.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing, 8,* 453-485.

Foorman, B. R., & Liberman, D. (1989). Visual and phonological processing of words: A comparison of good and poor readers. *Journal of Learning Disabilities, 22,* 349-355.

Glover, J. A., & Corkill, A. J. (1987). Influence of paraphrased repetitions on the spacing effect. *Journal of Educational Psychology, 79,* 198-199.

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6,* 126-135.

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

Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart (2002). *Journal of Experimental Child Psychology, 82,* 47-57.

Goswami, U., & Bryant, P. (1990). *Essays in Developmental Psychology: Phonological skills and learning to read.* Lawrence Erlbaum Associates Ltd.: East Sussex, UK.

Goswami, U., & East, M. (2000). Rhyme and analogy in beginning reading: Conceptual and methodological issues. *Applied Psycholinguistics, 21,* 63-93.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 371-377.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65,* 41-57.

Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology, 96,* 150-166.

Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005). Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9,* 351-365.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Behavior, 17,* 649-667.

Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly, 35,* 458-492.

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology, 94,* 114-133.

Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology, 58,* 88-111.

Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology, 55,* 297-328.

Levy, B. A. (2001). Moving the bottom: Improving reading fluency. In M. Wolf (Ed.), *Dyslexia, Fluency, and the Brain,* Timonium, MD.: York Press, pp. 367-379.

Levy, B. A., Bourassa, D. C., & Horn, C. (1999). Fast and slow namers: Benefits of segmentation and whole word training. *Journal of Experimental Child Psychology, 73,* 115-138.

Levy, B. A., Gong, Z., Hessels, S., Evans, M. A., & Jared, D. (2005). Understanding print: Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology, 93,* 63-93.

Levy, B. A., & Lysynchuk, L. (1997). Beginning word recognition: Benefits of training by segmentation and whole word methods. *Scientific Studies of Reading, 1,* 359-387.

Levy, B. A., Martin-Chang, S., Graydon, S., Horn, C., Minnett, S., & Breukelman, C. (2000). Developing reading skill: Making orthography visible. Unpublished manuscript, McMaster University, Hamilton, Ontario, Canada.

Li, Y., & Wright, D. L. (2000). An assessment of the attention demands during random- and blocked-practice schedules. *The Quarterly Journal of Experimental Psychology, 53A,* 591-606.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B., (1974). Reading and the awareness of linguistic segments. *Journal of Experimental Child Psychology, 18,* 201-212.

Lovett, M. V., Borden, S. L., DeLuca, T., Lacerenza, L., Benson, N. J., & Brackstone, D. (1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. *Developmental Psychology, 30,* 805-822.

Lovett, M. W., Warren-Chaplin, P. M., Ransby, M. J., & Borden, S. L. (1990).

Training the word recognition skills of reading disabled children: Treatment and transfer effects. *Journal of Educational Psychology, 82,* 769-780.

Martin-Chang, S. L., Levy, B. A., & O'Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology, 96,* 37-56.

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

McNeil, A. M., & Johnston, R. S. (2008). Poor readers' use of orthographic information in learning to read new words: A visual bias or a phonological deficit? *Memory and Cognition, 36,* 629-640.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24,* 45-64.

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

Naglieri, J. A. (1985). *Matrix Analogies Test-Expanded Form.* San Antonio: The Psychological Corporation.

Oakhill, J., & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology, 75,* 152-164.

Olson, R. K., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the posttraining development of word recognition. *Scientific Studies of Reading. Special Issue: Components of Effective Reading Instruction, 1,* 235-253.

Olson, R. K., & Wise, B. W. (1992). Reading on the computer with orthographic and speech feedback. *Reading and Writing: An Interdisciplinary Journal, 4,* 107-144.

Ouellette, G.,& Senechal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development, 79,* 899-913.

Parrila, R., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading, 8,* 3-26.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly. Special Issue: Children's Reading* *and the Development of Phonological Awareness, 33,* 283-319.

Roberts, L., & McDougall, S. (2003). What do children do in the rime-analogy task? An examination of the skills and strategies used by early readers. *Journal of Experimental Child Psychology, 84,* 310-337.

Rosner, J. (1999). *Test of Auditory Analysis Skills.* Academic Therapy Publications: Novato, CA.

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

Snowling, M. J. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. *Journal of Experimental Child Psychology, 29,* 294-305.

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

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-406.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38,* 175-190.

Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 437-444.

Torppa, M., Poikkeus, A. M., Laakso, M. L., Tolvanen, A., Leskinen, E., Leppanen, P. H. T., et al. (2007). Modeling the early paths of phonological awareness and factors supporting its development in children with and without familial risk of dyslexia. *Scientific Studies of Reading, 11,* 73-103.

Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124,* 107-136.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology, 83,* 437-443.

Vellutino, F. R., Pruzek, R. M., Steger, J. A., & Meshoulam, U. (1973). Immediate visual recall in poor and normal readers as a function of orthographic-linguistic familiarity. *Cortex, 9,* 370-386.

Walton, P. D., & Walton, L. M. (2002). Beginning reading by teaching in rime analogy: Effects on phonological skills, letter-sound knowledge, working memory, and word-reading strategies. *Scientific Studies of Reading, 6,* 79-115.

Wilkinson, G. S. (1993). *WRAT3 Administration Manual.* Wilmington, DE: Wide Range, Inc.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415-438.

Wolf, M., O'Rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 15,* 43-72.

Wood, C., & Farrington-Flint, L. (2002). Orthographic analogy use and phonological effects in non-word reading. *Cognitive Development, 16,* 951-963.

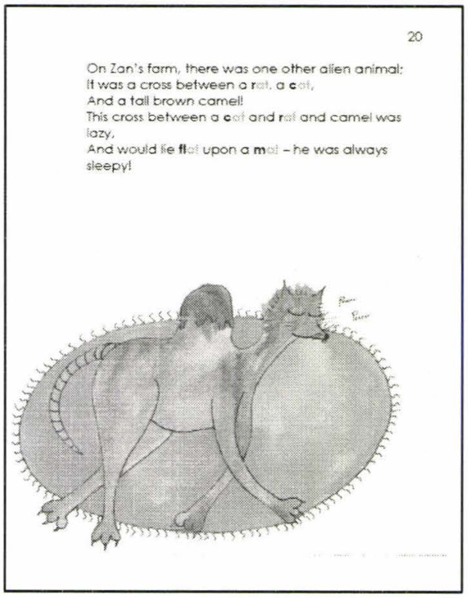
Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3-29.

Appendix A. Word lists for Experiments 1, 2A, and 2B. Note that Experiments 2A and 2B used only 16 Generalization Words and 16 Generalization Pseudowords (*), such that each list comprised two instances of each rime family instead of four instances.

Rime family	**Training words**	**Generalization words**	**Generalization pseudowords**
ig	big	dig*	mig*
	pig	jig	kig
	fig	twig	sig
	wig	rig*	tig*
an	can	pan*	han*
	ran	ban	zan
	flan	fan*	gan*
	man	van	san
in	pin	din	rin*
	bin	grin	lin*
	win	kin*	hin
	fin	tin*	vin
at	mat	bat*	jat*
	flat	hat	dat*
	cat	sat	lat
	rat	pat*	gat
ap	cap	rap*	hap
	flat	lap	bap*
	nap	gap	fap*
	map	tap*	vap
it	fit	wit*	dit*
	bit	kit*	jit
	sit	hit	vit
	pit	lit	rit*
ill	will	pill	zill
	bill	dill*	nill
	fill	gill	rill*
	hill	mill*	vill*
ash	dash	clash	fash*
	rash	bash*	jash
	flash	sash*	pash
	cash	lash	tash*

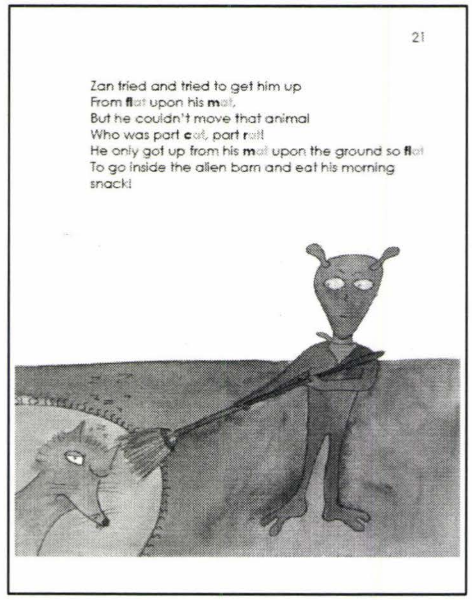
Appendix B. Excerpts from books used in Experiment 1.

Blocked/Highlighted



20

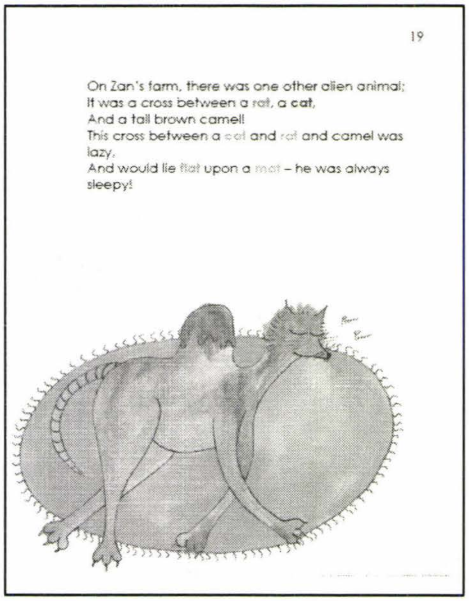
On Zan's farm, there was one other alien animal:
It was a cross between a rat, a cat,
And a tall brown camel!
This cross between a cat and rat and camel was lazy,
And would lie flat upon a mat – he was always sleepy!

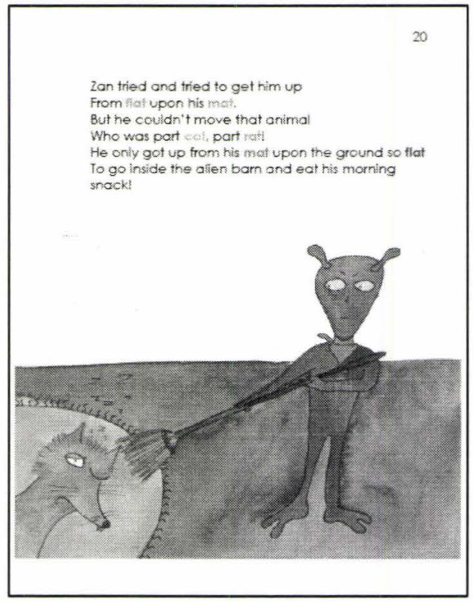


21

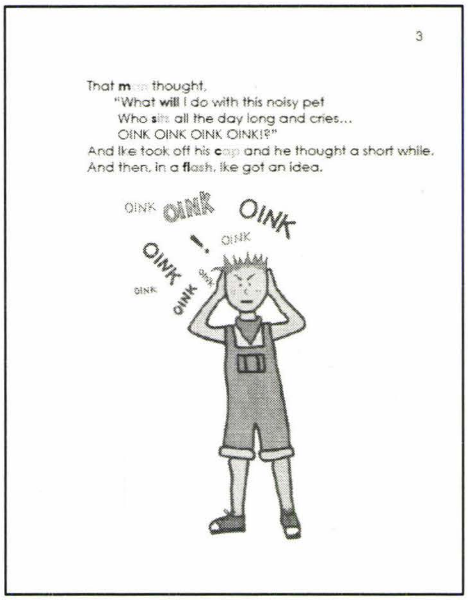
Zan tried and tried to get him up
From flat upon his mat,
But he couldn't move that animal
Who was part cat, part rat!
He only got up from his mat upon the ground so flat
To go inside the alien barn and eat his morning snack!

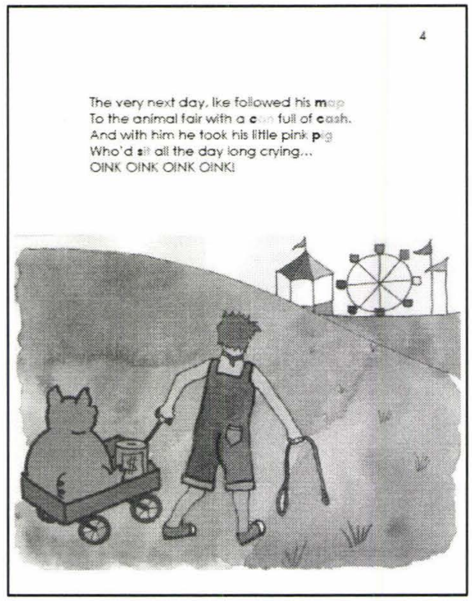
Blocked/Not Highlighted



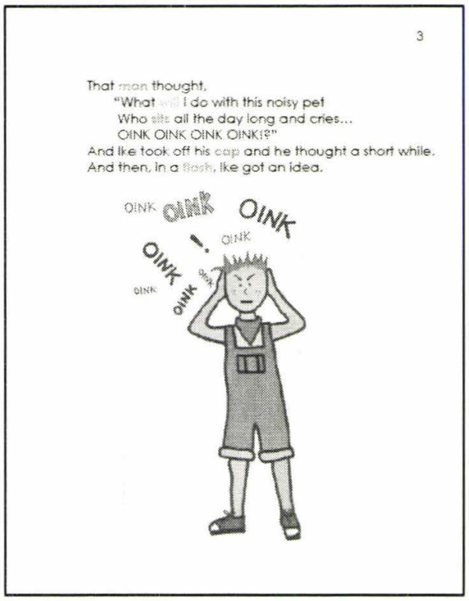


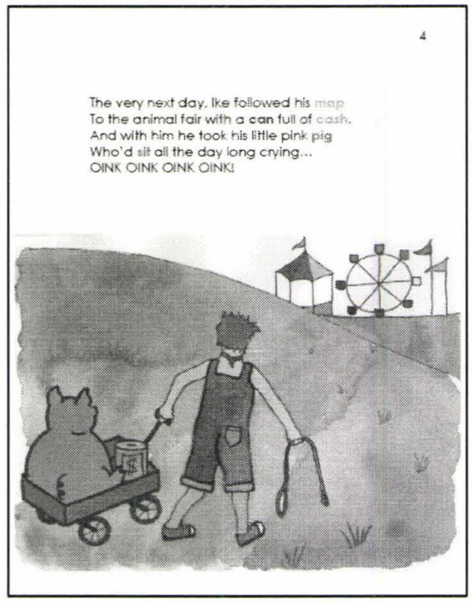
Unblocked/Highlighted



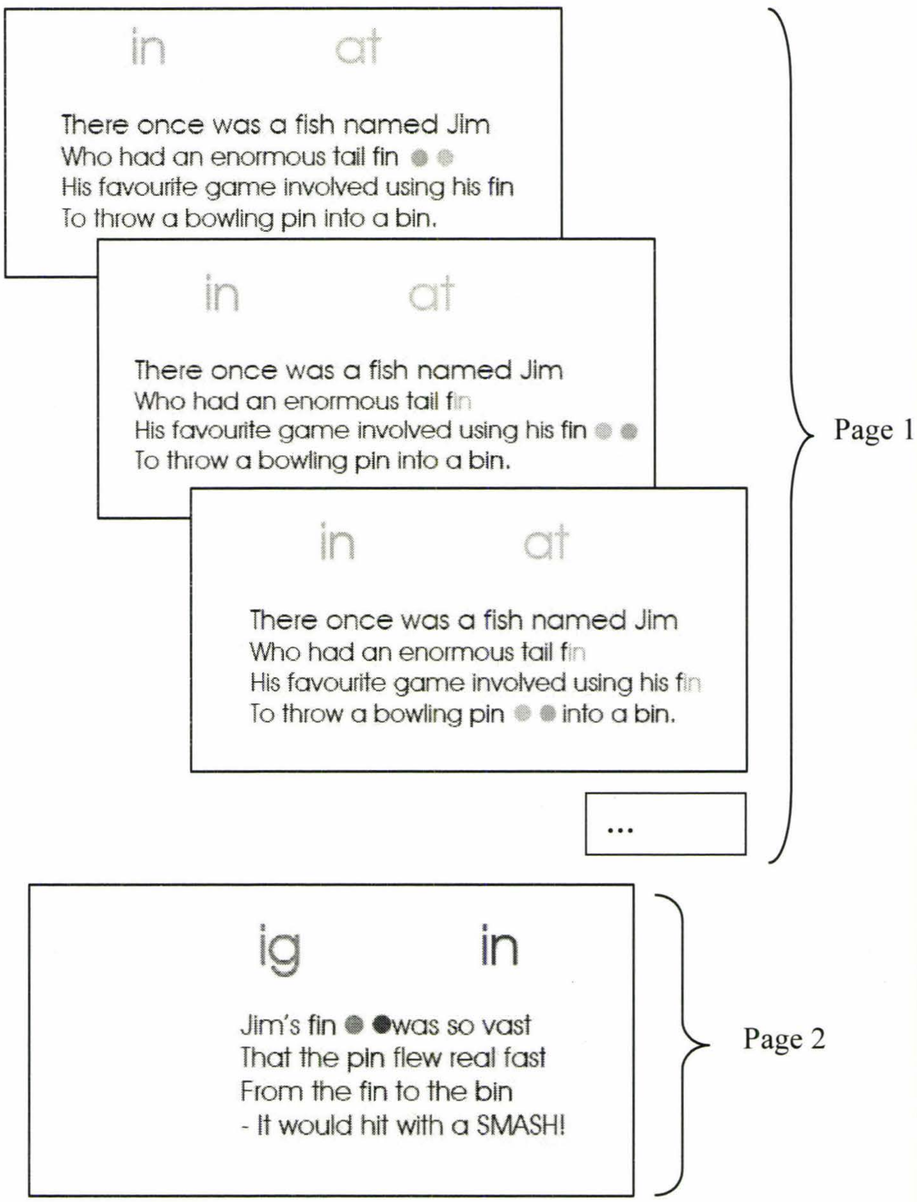


Unblocked/Not Highlighted

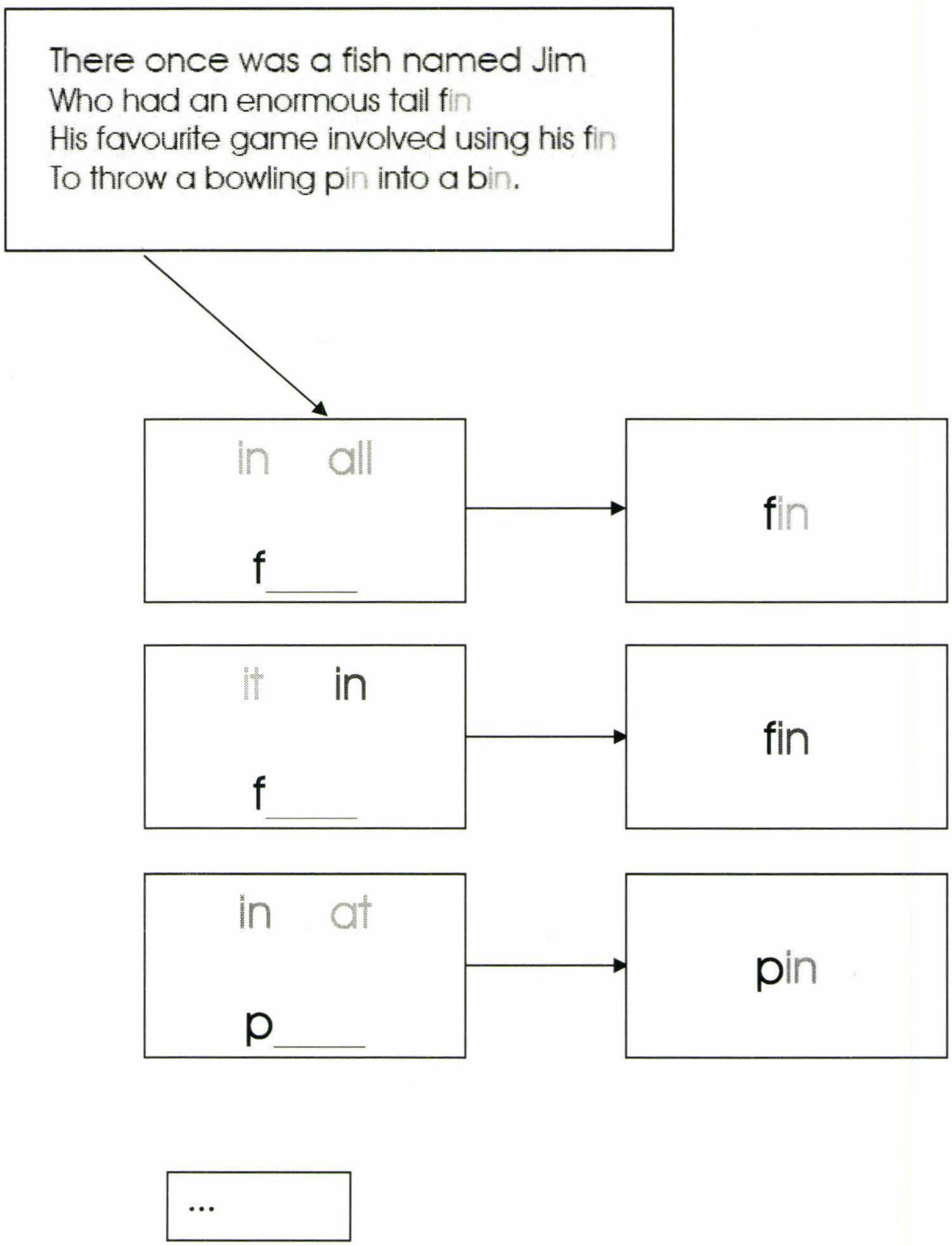




Appendix C. Example of Colour Choice task used in Experiment 2A. Note that all pages included pictures above the text (as in Experiment 1), but that pictures have been excluded here due to space constraints.



Appendix D. Example of Rime Choice task used in Experiment 2B. Note that all pages included pictures above the text (as in Experiment 1), but that pictures have been excluded here due to space constraints.



Appendix E. Word lists for Experiment 3.

Training words

	Onset families							
Rime families	d	h	j	l	b	f	p	r
ot	dot	hot	jot	lot				
ay	day	hay	jay	lay				
ug	dug	hug	jug	lug				
og	dog	hog	jog	log				
an					ban	fan	pan	ran
ad					bad	fad	pad	rad
ig					big	fig	pig	rig
un					bun	fun	pun	run

Generalization words

Rime families						**Onset families**				
ot	cot	not				d	den	dim		
ay	say	may				b	bus	bed		
ug	mug	tug				f	fit	fin		
og	bog	fog				h	ham	hip		
an			can	tan		p			pal	pod
ad			sad	mad		j			jet	job
ig			gig	wig		r			rag	rib
un			sun	nun		l			lab	lid

Generalization pseudowords

Rime families						**Onset families**				
ot	mot	sot				d	dag	dit		
ay	vay	tay				b	bep	bom		
ug	nug	sug				f	fud	fam		
og	mog	tog				h	hin	heb		
an			gan	san		p			pob	pum
ad			gad	kad		j			jek	jaf
ig			kig	nig		r			rij	rog
un			tun	mun		l			lar	lub