Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
Does Vocabulary Influence Word Recognition and Reading Comprehension Through Set for Variability?

A thesis presented in partial fulfilment of the requirements for the degree of Master of Education at Massey University, Manawatū, New Zealand

Cheryl-Anne Mackay 2017
Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
Vocabulary is recognised as a key contributor to literacy development and comprehension. Children cannot make meaning from text if they lack the vocabulary to support what is being read. This study investigated the independent contribution of vocabulary to word recognition, and whether the contribution was direct or indirect through set for variability. A second aim of the study was to determine if a direct relation exists between vocabulary knowledge and reading comprehension.

Unpublished data from the longitudinal study (Chapman, Arrow, Tunmer, & Braid, 2016) were analysed to find predictive links between vocabulary and later reading outcomes, for a cohort of 374 5-year-old children in New Zealand primary schools. The results identified that word recognition and vocabulary both directly contributed to reading comprehension for these children in the middle of their second year at school. Word recognition explained a greater amount of the shared variance of reading comprehension in the middle of Year 2 as the children were at the stage of still trying to read a range of unfamiliar words. Set for variability was found not to directly contribute to reading comprehension when word recognition was added to the model. This finding suggests that set for variability mediates the relationship between vocabulary and word recognition but not for overall reading comprehension. In terms of practical teaching, it is suggested that language comprehension abilities should be acquired alongside the development of word recognition skills.

**Keywords:** alphabetic principle, decoding, language comprehension, lexicon, morphology, oral vocabulary, orthography, orthographic representation, phonological awareness, phonological recoding, phonological representation, phonology, reading
comprehension, semantics, set for variability, sight words, Simple View of Reading, syntax, word recognition.
Acknowledgements

I would like to take this opportunity to acknowledge the support and contribution made by my supervisors Dr Alison Arrow, Professor Margaret Walshaw and Dr Tara McLaughlin. Their guidance has encouraged me to delve deeply into new areas of learning and I have acquired so much through the process. All three supervisors have helped to broaden my knowledge and thinking in ways I could have never imagined possible.

I am grateful to Massey University for the Masterate Scholarship and the Institute of Education for the Pathway Scholarship. This funding has been beneficial in assisting in the implementation and completion of this study.

I wish to acknowledge the leaders of the Early Literacy Project, Professor James Chapman, Dr Alison Arrow, Distinguished Professor William Tunmer and Christine Braid, for sharing their unpublished data with me.

I would also like to thank my amazing family and friends for their ongoing support. Their presence and encouragement has been immense, especially during the difficult times. Finally, I would like to thank Danielle and Brayden for the extra effort and help they gave around home which enabled me to concentrate on my study. Without all the incredible support, this would never have been completed.
Contents

Abstract ................................................................................................................................................................. iii
Acknowledgements ................................................................................................................................................... v
Contents ................................................................................................................................................................ vi
List of Tables .......................................................................................................................................................... vii
Chapter One ........................................................................................................................................................... 1
   Introduction .......................................................................................................................................................... 1
   Rationale ......................................................................................................................................................... 3
   The Present Study ............................................................................................................................................. 4
   Key Terms and Definitions .............................................................................................................................. 4
   Thesis Overview ........................................................................................................................................... 8
Chapter Two ............................................................................................................................................................ 10
   Literature Review ........................................................................................................................................... 10
      Introduction ................................................................................................................................................ 10
      Theories of Early Reading Development and Acquisition ...................................................................... 11
      Decoding .................................................................................................................................................. 13
      Language Comprehension ........................................................................................................................ 18
      Vocabulary ................................................................................................................................................. 19
      Reading Comprehension .......................................................................................................................... 20
      Set for Variability ...................................................................................................................................... 26
      Aotearoa, New Zealand Research: Tunmer and Chapman (2012) .............................................................. 26
      The Present Study ..................................................................................................................................... 28
Chapter Three .......................................................................................................................................................... 30
   Methodology .................................................................................................................................................... 30
      Research Design ........................................................................................................................................ 30
      Sampling ................................................................................................................................................... 31
      Setting and Participants ............................................................................................................................ 33
      Data Collection Procedures ....................................................................................................................... 34
      Measures ................................................................................................................................................... 35
      Ethics ........................................................................................................................................................... 38
      Summary ................................................................................................................................................... 38
Chapter Four ............................................................................................................................................................ 40
   Results ............................................................................................................................................................... 40
      Descriptive Results .................................................................................................................................. 40
      Contributions to Word Recognition ........................................................................................................... 43
      Contributions to Reading Comprehension ................................................................................................. 45
      Summary ................................................................................................................................................... 48
Chapter Five ............................................................................................................................................................ 49
   Discussion ......................................................................................................................................................... 49
      The Direct and Indirect Influence of Vocabulary on Word Recognition ...................................................... 50
      The Direct Influence of Vocabulary on Reading Comprehension .............................................................. 52
      Limitations ................................................................................................................................................ 55
      Future Research ....................................................................................................................................... 56
      Implications for Practical Applications ..................................................................................................... 59
      Conclusion ................................................................................................................................................ 62
References ............................................................................................................................................................... 65
List of Tables

Table 1 Gender and Ethnicity ........................................................................................................... 33
Table 2 Correlations, Means and Standard Deviations for all Measures ................................. 41
Table 3 Model Summary and Coefficients .................................................................................... 44
Table 4 Model Summary and Coefficients .................................................................................... 46
Chapter One

Introduction

The ability to read text in multiple forms allows access to information, participation and active engagement in the world. Tunmer and Prochnow (2009) claim becoming literate is arguably the most important goal of schooling. Substantially limited life options are imminent for those who do not have the ability to read and write. The ability to read is essential to success in most aspects of the school curriculum at increased levels of complexity and independence. In addition, reading is a prerequisite skill needed for nearly all jobs and is the primary key to lifelong learning. Educational and life chances are somewhat determined through a person’s literacy level.

The Simple View of Reading (SVR) offers an important model of reading. The model is increasingly being drawn upon to explain the nature of reading development. According to the SVR, reading comprehension is the product of decoding and language comprehension (Gough & Tunmer, 1986). The reader must be able to decode words and hold a level of language comprehension for reading success (Herrera, Nielsen, Bridges, & Catts, 2015). As part of the Simple View, oral vocabulary is an important component of language comprehension (Gough & Tunmer, 1986) and is increasingly becoming recognised as a contributor to literacy development and reading comprehension. The term vocabulary is used for the purposes of this study to refer to oral vocabulary. Evidence suggests that learning to read is closely connected to children’s oral language skills and plays an important
role in the development of word recognition and reading comprehension (Ricketts, Nation, & Bishop, 2007). Vocabulary is crucial for a reader to gain meaning from text as children cannot make meaning from text if they do not have the vocabulary to support the text being read (Sénéchal, Ouellette, & Rodney, 2006).

This study investigates the influence of vocabulary, a component of language comprehension, on word recognition (as a measure of decoding) and reading comprehension. Set for variability is a skill understood to be helpful in measuring the cognitive reading process. Set for variability is important to measure as the skill contributes to the development of both decoding and word recognition skills which enable children to identify unknown words based on partial decodings (Venezky, 1999).

While the role of word recognition (as a measure of decoding) and language comprehension has been firmly established in learning to read in previous investigations, more recent research has focused on the role of vocabulary in reading comprehension and, moreover, the role of word recognition. Much of the evidence gathered from these studies has involved older children with only a limited number of studies providing evidence from children upon their commencement at school (Kendeou, Savage, & van den Broek, 2009; Lorch & van den Broek, 1997; Oakhill, Cain, McCarthy, & Nightingale, 2012; Ricketts et al., 2007; Sénéchal et al., 2006; Stuart, Staintorp, & Snowling, 2008; Vellutino, Tunmer, Jaccard, & Chen, 2007). Further evidence around the influence of vocabulary on word recognition and reading comprehension is needed from this younger age group to contribute to the knowledge gained from the small number of studies which have previously been carried out.
The current research could have implications for instructional practices in New Zealand primary schools. Identifying children with low vocabulary and who experience difficulty with word recognition and reading comprehension at the emergent literacy learning stage can allow for specific teaching strategies to be recognised and implemented. Such an approach may go some way to preventing the current achievement gap from widening.

Rationale

Acquiring reading comprehension is the fundamental product of the Simple View of Reading. Each component strives to maintain equal importance in the quest to achieve this (Gough & Tunmer, 1986; Hoover & Gough, 1990). While previous studies have established the role of word recognition (as a measure of decoding) and language comprehension in learning to read, further research with younger children 5 to 6 years of age, focusing on the role of vocabulary in reading comprehension and word recognition is needed. Knowing about how reading develops and the key components needed for reading development is important. When weaknesses are identified within the components of vocabulary or word reading skills at the emergent literacy stage, support can be implemented by way of specific teaching strategies. Such strategies may lead to a closing of the achievement gap and prevent these children falling further behind. The most recent National Standards data from the Ministry of Education (2016) shows there is still 30-40% of children not achieving at the standard for reading and writing. Early identification is important because effective strategies will assist these large numbers of children who do not seem to be improving in literacy to be supported.
The Present Study

The present research aims to demonstrate the importance of vocabulary to both reading comprehension and word recognition using the Simple View of Reading as a model to explain these relations across development. The research also aims to determine, if vocabulary knowledge directly influences word recognition, or if it is indirectly related through set for variability. The evidence from this study is intended to support previous national and international studies which consistently demonstrate best evidenced based practice in order to prevent the failings presently occurring.

Key Terms and Definitions

*Alphabetic principle* - the alphabetic principle is the understanding there are systematic and predictable relationships between letters and sounds. The knowledge that letters in written words correspond to phonemes in spoken words and the ability to translate the letters into phonological forms which, in turn, enable children to explicitly, and implicitly decode print (Arrow & Tunmer, 2012; Byrne & Fielding-Barnsley, 1989).

*Decoding* - decoding is the ability to sound out words based on phonics rules, including fast and accurate reading of familiar and unfamiliar words in text allowing access to the appropriate entry in the lexicon (Hoover & Gough, 1990). The term word recognition is used for the purposes of this study. Word recognition can be conceptualised in three different ways. Words can be “sounded out,” or decoded using letter-by-letter processing, phonologically recoded where unfamiliar printed words are translated into phonological forms or read as sight words which can be read automatically and do not need to be decoded.
Language comprehension (sometimes known as linguistic comprehension and listening comprehension) - language comprehension is the ability to gain meaning from spoken words, parts of sentences or other discourse (Herrera et al., 2015).

Lexicon - the lexicon is the part of the linguistic memory which contains knowledge of known words, much like a mental dictionary (Moats, 2010).

Morphology - morphology is the study and description of the meaning components of words. A morpheme is the smallest linguistic unit (part of a word) that has meaning. So, a word can contain a single morpheme: happy (the root morpheme) and its meaning can be changed by adding morphemes: un+happy, happi+ness, etc., to create words. A morphemic analysis can help with the derivation of complex, multi-morphemic words, e.g. antidisestablishmentarianism. Plurals and suffixes are also morphemes (dog+s); though in this case the base meaning of the word does not change (Oakhill, Cain, & Elbro, 2014).

Oral vocabulary - oral vocabulary is defined for the purposes of this study, as words used in speaking and listening. This extends to an expandable, stored set of word meanings in the lexicon (Stahl & Nagy, 2005). The term vocabulary is used for the purposes of this study to refer to oral vocabulary.

Orthography/Orthographic Representation - orthography is a writing system which relates to letters and spellings and the representation of speech sounds in writing. English orthography is fundamentally alphabetic which contain symbols that mostly
directly relate to speech sounds or phonemes. The alphabetic base contains 26 letters plus punctuation marks (Oakhill et al., 2014).

**Phonological awareness** - phonological awareness is the ability to hear and consciously break words into various parts. These parts include syllables, rhyme, onset and rime, and individual sounds or phonemes (Moats, 2010).

**Phonological Recoding** - phonological recoding is the implicit ability to translate letters and letter patterns into phonological forms. Unfamiliar printed words are translated into spoken equivalents which are thought to facilitate the reading acquisition process by acting as a self-teaching mechanism (de Jong & Share, 2007). For example, gaining an understanding that ‘at’ makes a specific sound and being able to use this knowledge to recode other words, like ‘cat’.

**Phonological Representation** - a phonological representation is the spoken form of a word. An accurate identification of individual phonemes is necessary because misclassification of a single phoneme can produce a word which is both semantically and syntactically different (e.g. ran vs van) (Adams, 1994).

**Phonology** - phonology is the basic sound units of language called phonemes. Phonology is a smaller component of language comprehension (Moats, 2010; Snow, Burns, & Griffin, 1998).

**Reading Comprehension** - reading comprehension is the product of decoding and language comprehension according to the Simple View of Reading. The multidimensional process involves the reader constructing a representation by
quickly and accurately locating individual words and sentences in lexical memory (Kendeou, Savage, et al., 2009).

*Semantics* - semantics involves understanding the meaning of individual words and sentences being spoken and the meaningful relations between them at three different levels, the discourse and sentence level, the vocabulary level and the morphology level (Moats, 2010; Snow et al., 1998).

*Set for Variability* - set for variability is the ability to determine the correct pronunciation of mispronounced spoken English words derived from regular pronunciations of irregularly spelled words (e.g., stomach pronounced as “stow-match”), the incorrect pronunciation of words contained polyphonic spelling patterns (e.g., glove pronounced like “clove”), and approximations to correct pronunciations based on the use of context-free spelling rules (e.g., kind pronounced like “pinned”) (Venezky, 1999).

*Sight words* - sight words are words automatically recognised from memory as whole words. Sight words are stored in the lexicon and do not require explicit sounding out (Ehri, 2005a).

*Simple View of Reading* - the Simple View of Reading (SVR) offers a theory of what reading is and how reading development occurs. According to the SVR, reading comprehension is the product of decoding and language comprehension, that is RC = D x LC. Each component is equally important to attaining reading comprehension (Gough & Tunmer, 1986).
Syntax - syntax encompasses the rules of language that denote how to combine different classes of words such as nouns, verbs and adjectives, to form sentences. Syntax defines the structural relationship between the sounds of a language or the phonological combinations, and the meaning of those combinations. Syntax can also assist people in working out unfamiliar words by developing ideas about what the unfamiliar word could be (Moats, 2010; Snow et al., 1998).

Word Recognition - word recognition, for the purposes of this study, is defined as the term which accurately reflects and includes the reading of words automatically through sight because of repeated exposures as well as phonological decoding skills (Ehri, 2014). Word Recognition can be conceptualised in three different ways. Words can be “sounded out,” or decoded using letter-by-letter processing, phonologically recoded where unfamiliar printed words are translated into phonological forms or read as sight words which can be read automatically and do not need to be decoded.

Thesis Overview

This thesis is comprised of five chapters. The second chapter reviews the literature and includes a discussion about two prominent theories of word reading and reading comprehension. The chapter discusses the role in which vocabulary is increasingly being recognised as a contributor to both word reading and reading comprehension. Also discussed in this chapter is the skill of set for variability and the ways in which this skill can be instrumental in the process of reading. Chapter 3 explains the methodology and research design used in the study. The results are presented in Chapter 4. Finally, Chapter 5 discusses the findings of the study in relation to the literature and considers the practical implications of the study. The study seeks to
provide further evidence on the role of vocabulary and set for variability in relation to the literature and the nature of reading and instruction within the New Zealand context. Suggested implications for further research are also noted.
Chapter Two

Literature Review

Introduction

Learning to read is a main objective for children living in a literate society. Reading skills build a foundation from which children can strive to achieve future academic success. Children who read well, will read more and gain additional knowledge and enhanced cognitive skills (Whitehurst & Lonigan, 2002). Effective word identification strategies, focused on rapid recognition of unknown words, are vital for successful reading acquisition. Vocabulary is progressively being recognised as a key contributor to literacy development and comprehension. Children cannot make meaning from text if they lack the vocabulary to support the text being read (Sénéchal et al., 2006). This study looked at the role vocabulary plays in the development of word reading which, in turn, will predict reading comprehension.

The study was conducted through a secondary analysis of literacy data collected in years 2015 and 2016 in Aotearoa, New Zealand. The study tests a model of relationships among key variables related to the Simple View of Reading (SVR) first proposed by Gough and Tunmer (1986). This chapter will briefly outline the context for literacy instruction in Aotearoa, New Zealand - highlighting historic, current, and emerging trends. Then the SVR is presented, focusing on key aspects including research related to decoding, language comprehension, reading comprehension and vocabulary. To further set the context for the present study, a detailed description of past New Zealand research conducted by Tunmer and Chapman (2012) related to the SVR is provided. As part of this, set for variability is introduced as a mediator...
between key variables in the SVR. The chapter ends with the context for the current study.

**Theories of Early Reading Development and Acquisition**

New Zealand has predominantly followed the whole language approach over the past 25 years which supports teaching children to read by recognising words as whole pieces of language (Tracey & Morrow, 2006). The Ministry of Education has adopted and strongly promotes the approach which has been embedded in literacy education through publications distributed to schools since the 1970s (Tunmer, Chapman, Greaney, Prochnow, & Arrow, 2013). The multiple-cues and psycholinguistic guessing theory of reading underpins the whole language approach (Goodman, 1967). The Simple View of Reading (SVR) offers a contrasting theoretical base to early reading acquisition and is increasingly being drawn upon as an alternative to explaining the nature of reading development.

The SVR offers a theory of what reading is and how reading development occurs. While the SVR acknowledges the value of the whole language approach by positioning reading as a linguistic activity, Gough and Tunmer (1986) maintain phonics and the cognitive processes used in decoding are equally prominent in the relationship between decoding and language comprehension.

According to the SVR, reading comprehension is presented as the product of decoding and language comprehension, that is $RC = D \times LC$ (Gough & Tunmer, 1986). While not denying reading is a complex process, the Simple View maintains decoding and language comprehension separately hold equal importance (de Jong &
van der Leij, 2002; Megherbi, Seigneuric, & Ehrlich, 2006; Muter, Hulme, Snowling, & Stevenson, 2004; Oakhill, Cain, & Bryant, 2003). Decoding denotes the ability to recognise and process written information. But decoding in isolation is not sufficient to gain meaning from print (Ehri, 2005b). The reader must be able to decode words and possess a level of language comprehension for reading success. Language comprehension is the ability to use linguistic knowledge to gain meaning from sentences and discourses (Herrera et al., 2015). Vocabulary is an important component of language comprehension (Gough & Tunmer, 1986). The two main components of the SVR, decoding and language comprehension, are underpinned by several cognitive developmental processes (orthographic, phonological and semantic) which develop congruently and reinforce each other. Further analysis of this model is worthy of investigation and therefore the present study explores the influence of vocabulary, a component of language comprehension, on word recognition (as a measure of decoding) and reading comprehension. The influence under investigation can be explained through the dynamic, evolving relationships between the two components. Research and theory for each component are outlined in the following sections to allow for a better understanding of the relationship(s) under investigation.

A large and growing body of evidence supports the Simple View, both in English (Aaron, Joshi, & Williams, 1999; Catts, Herrera, Nielsen, & Bridges, 2015; Kirby & Savage, 2008; Storch & Whitehurst, 2002; Tunmer & Chapman, 2012) and other languages (de Jong & van der Leij, 2002; Kendeou, Papadopoulos, & Kotzapoulou, 2013; Megherbi et al., 2006; Protopapas, Simos, Sideridis, & Mouzaki, 2012). This research work has documented the contributions of word recognition (as an outcome of decoding) and language comprehension to reading comprehension, and how
these contributions change across grades and/or skill levels (Catts, Hogan, & Adlof, 2005; Language and Reading Research Consortium, 2015). The Language and Reading Research Consortium (2015) used multiple measures to assess the basic SVR model with first, second, and third grade children. Grade 1 and 2 is the equivalent to the New Zealand new entrant to Year 2 and beginning Grade 2 is the equivalent to the New Zealand Year 3 levels. Findings from the Language and Reading Research Consortium confirmed the influence of decoding skill decreased with increasing grade. The influence of language comprehension increased and vocabulary skills indirectly affected reading comprehension through both decoding skill and language comprehension.

**Decoding**

Decoding is defined as the ability to sound out words based on phonics rules, including fast and accurate reading of familiar and unfamiliar words in text allowing access to the appropriate entry in the lexicon (Hoover & Gough, 1990). While some equate decoding with sounding out, Gough and Tunmer (1986) used the label decoding to emphasise the importance of letter-sound correspondence rather than word recognition. The term word recognition is used for the purposes of this study as the term more accurately reflects and includes the reading of words automatically through sight because of repeated exposures as well as phonological decoding skills (Ehri, 2014). Word recognition can be conceptualised in three different ways. Words can be “sounded out,” or decoded using letter-by-letter processing, phonologically recoded where unfamiliar printed words are translated into phonological forms or read as sight words which can be read automatically and do not need to be decoded.
Explaining how beginners acquire the ability to recognise words rapidly is a major task faced by researchers. The developmental theory of word reading from Ehri (1992) depicts a succession of distinct phases where a series of skills emerge, develop and change while involving connections. The four phases are characterised by the degree of cognitive skills involved. These skills include alphabetic knowledge and phonological awareness (the ability to identify and manipulate units of oral language including recognition of a word with the same initial sound, syllables and onsets and rimes). Each facilitates the alphabetic principle (the understanding that letters in written words correspond to phonemes in spoken words and being able to translate the letters into phonological forms) enabling children to explicitly, and implicitly decode print (Arrow & Tunmer, 2012). Children entering school with poor cognitive skills such as weak phonological awareness, alphabet knowledge and vocabulary, miss opportunities to develop reading comprehension strategies (Kendeou, van den Broek, White, & Lynch, 2009; Oakhill et al., 2012; Vellutino et al., 2007; Whitehurst & Lonigan, 2002). A connection forming process is at the heart of word learning. Connections form and link the written word to the pronunciation and word meaning. This information is stored in the reader's word memory or lexicon (Ehri, 1992; Rack, Hulme, Snowling, & Wightman, 1994). The following explores the general pattern which enables these connections to form.

Some words encountered in print need to be “sounded out,” or decoded using letter-by-letter processing. Explicit decoding, the first way to decode a word, involves reading and processing every single letter while applying grapheme-phoneme relationships to figure out how to read and spell. The written form (orthography) of an unknown word is translated into spoken form (phonology) through interactions between the orthographic processor, the phonological processor (which processes
phonemes), and the semantic processor when attempting to sound out a pronounceable word (Adams, 1994). Attempts are made to make sense of these letter-sound relationships by making connections between the symbols on the page and the sounds of speech which requires knowledge of the alphabetic principle (Ehri, Nunes, Stahl, & Willows, 2001). The alphabetic principle is applied to words read in text and words spelt when they are written. Children do this through sounding out and are aware they are trying to match letters to sounds in order to say a word (Arrow & Tunmer, 2012). A correct phonological representation necessitates an accurate identification of individual phonemes because misclassification of a single phoneme can produce a word which is both semantically and syntactically different (e.g., ran versus van).

The second type of decoding is developed through a process called the “self-teaching” hypothesis (Share, 2004) or phonological recoding. Phonological recoding requires an implicit ability to translate letters and letter patterns into phonological forms. Unfamiliar printed words are translated into spoken equivalents which are thought to facilitate the reading acquisition process by acting as a self-teaching mechanism (de Jong & Share, 2007). For example, having an understanding that 'at' makes a specific sound allows for this knowledge to be used to recode other words, like 'cat'.

Skilled readers use phonological recoding to read unseen words (including pseudowords), and the spellings of exception words involving patterns or rules. The aim of applying letter-sound knowledge to unknown exception words is to make a partial decoding which is sufficiently close to the correct phonological form for the word to be correctly identified. This occurrence is more likely if the corresponding
spoken word is in the vocabulary and if a skill called set for variability is applied (Tunmer & Chapman, 2012). Set for variability is further explained in a later section.

As a result of phonological recoding, cognitive resources are freed up and allocated to higher order cognitive functions such as determining the meanings of unknown words and gaining meaning from text. Children who lack phonological recoding skills will experience difficulty determining the phonological representation and meanings of any unknown words encountered (Tunmer & Hoover, 1993).

Studies have demonstrated that kindergarten children are capable of learning words on a phonetic basis rather than a visual one if they have some knowledge of spelling-sound relationships. A child capable of generating words beginning with a given sound and who has a basic knowledge of letter-sound correspondences should be able to generate a reasonable guess for an unknown word. Whereas a child who perceives words as wholes and is incapable of segmenting spoken words will have no way of generating a pronunciation for an unfamiliar word (Ehri & Sweet, 1991; Ehri & Wilce, 1985, 1987; Rack et al., 1994; Scott & Ehri, 1990). The gradual process of learning words on a phonetic basis continues over time until a solid foundation of words can be read without being sounded out. These words are referred to as sight words (Adams, 1994).

Sight word recognition (also known as orthographic representations) is the third way to decode a word. The skilled reader can read words quickly and accurately with no semantic context cues (e.g., on a flash card with only the to-be-read word) (Share, 1995). Frequent sightings of particular letter arrangements strengthen the associations between the combinations and the order of the corresponding letters in
the memory which gives the perception of a whole unit or whole word (Adams, 1994). These connections, resulting from phonological recoding, are retrieved from memory when the word is further encountered in print. Words such as the and of are not easily sounded out and require the visual memory to remember and recall them. Recognising words by sight, from memory, automatically activates the meaning and allows the reader to concentrate on constructing meaning (comprehension) rather than stopping to decode words as this slows the reading down and interrupts the reader's train of thought (Ehri, 2005a).

Throughout the connection forming process, the semantic processor is an integral part of decoding words which searches for words similar to that being pronounced (Whitehurst & Lonigan, 2002). An accurate word identification connects and stores the orthographic representation (the spelling and sound pattern) and the semantic representation (the meaning) of the target word in the lexical memory (the reader's long term memory) and contributes to the database from which further letter-sound patterns can be produced (Adams, 1994; Ehri & Saltmarsh, 1995; Jorm & Share, 1983; Share, 1995; Tunmer & Chapman, 2012). The semantic representations stored in the lexicon assist in the process of assigning a meaning to the word being read using the context of the sentence or paragraph in which the word is located to support the reader to make a correct identification (Adams, 1994).

A large vocabulary requires the lexicon to restructure and enable individual words to be distinguished. The ability to segment words into smaller parts (subcomponents) such as syllables and rime using phoneme awareness (the ability to manipulate individual phonemes within spoken words) is necessary (Shankweiler & Fowler, 2004). Skills include isolating or deleting individual phonemes (first, middle, last or
individual) from words. For example, children with phoneme awareness know that
the word ‘bait’ is made up of three phonemes, and that the words ‘pad’ and ‘map’
both contain the phoneme /p/ (Ehri et al., 2001). Phoneme awareness and phonics
are identified as being predictors of future reading acquisition (Ehri et al., 2001; Share, Jorm, Maclean, & Matthews, 1984).

This section examined the cognitive processes used to decode a word three different
ways from its printed form. Skilled readers accurately and efficiently decode familiar
and unfamiliar words they encounter using either explicit or implicit decoding skills
while drawing on the corresponding entry in the lexicon to gain meaning. Letters and
words are the primary cues used to decode text. The more skills the reader
possesses, the faster and better this interactive process works. In addition to
decoding, language comprehension is needed as part of the equation of the SVR in
the quest to acquire reading comprehension successfully.

**Language Comprehension**

Language comprehension (sometimes known as linguistic comprehension and
listening comprehension) is defined as the ability to derive meaning from spoken
words, parts of sentences or other discourse. Smaller components make up the
language component. These include phonology (the basic sound units of language
called phonemes), morphology (an awareness of parts of words including stems, root
words, prefixes, suffixes and parts of speech and the ways context can change a
word’s pronunciation and meaning), syntax (phrase and sentence structure, what
makes sense), semantics (the way language conveys meaning) and vocabulary
(Moats, 2010; Snow et al., 1998). Syntax and morphology are often investigated
using the SVR framework. Both syntax and morphology support language comprehension and are crucial for successful literacy acquisition because syntactic cues help determine the meanings of words (Hoover & Gough, 1990; Snowling & Hulme, 2005; Tunmer & Hoover, 1993; Vellutino et al., 2007). A good understanding of morphology helps children spell and pronounce words correctly. For example, the word “smallest” has two morphemes, “small” and “est” where each part has meaning. Applying morphological knowledge to words and morphemes assists children in acquiring more complex words, such as uncommon from un and common, and strangely from strange and ly (Cain, Oakhill, & Lemmon, 2004).

The more that is known about words and the interrelated components of words (i.e., their phonemes, orthographic patterns, semantic meanings, syntactic uses, and morphological awareness), the more efficiently words are decoded, retrieved, and comprehended (Cain et al., 2004). While language components are necessary, the focus for this study is solely on vocabulary as the main semantic measure.

**Vocabulary**

Vocabulary (otherwise referred to as oral vocabulary) is defined for the purpose of this study, as words used by a child in speaking and listening. This extends to an expandable, stored set of word meanings in the lexicon (Stahl & Nagy, 2005). Research related to the SVR suggests learning to read is closely connected to children’s underlying oral language skills and plays an important role in the development of word recognition and reading comprehension (Ricketts et al., 2007). Vocabulary is crucial for a reader to gain meaning from text. Consider what happens when a beginning reader comes across the word cat in a book. As the reader begins
to work out the sounds represented by the letters, familiar sounds are recognised if
the word has been heard and said many times previously. If the word has not been
previously heard and is not part of the speaking (oral) vocabulary, the reader will
experience difficulty in figuring out the word (Adams, 1994). Print reading vocabulary
contributes to vocabulary but first and foremost vocabulary comes from hearing
words used by others (Oakhill et al., 2014).

Poorly developed vocabulary knowledge results in difficulty identifying and allocating
the correct meanings to unknown printed words (especially partially decoded words,
irregularly spelled words, or words with complex spelling patterns) if the
corresponding spoken words are not in the listening vocabulary or are only weakly
represented phonologically in the lexicon (Nation & Snowling, 1998; Perfetti, 2007).
A correctly decoded unknown word is added to the reading vocabulary but the
semantic representation (the meaning) is unable to be connected to the word
because the word has not previously been heard (or has only been heard a few
times) and has not been stored in the lexicon.

Once children can decode (read) words, the focus shifts from ‘learning to read’ to
‘reading to learn’ and reading to learn requires comprehension (Oakhill et al., 2014).
Vocabulary is essential to reading comprehension as readers cannot understand
what is being read without knowing what most of the words mean (Westby, 2005).

**Reading Comprehension**

Reading comprehension, as reflected by the SVR, is the culmination of both word
recognition (as a measure of decoding) and language comprehension. Reading
comprehension can still occur if all the cognitive elements are not fully developed but it is less effective if some are not functioning well (Stuart et al., 2008; Tunmer et al., 2013). The multidimensional process involves the reader constructing a representation, by quickly and accurately locating individual words and sentences in the lexical memory where the intended meanings of individual words and sentences are determined (Kendeou, van den Broek, et al., 2009; Lorch & van den Broek, 1997; Oakhill et al., 2012; Vellutino et al., 2007). Quick and accurate access to words and word meanings is essential to comprehension. The reader is required to activate and collate word meanings within a fraction of a second to ensure connections are made with the semantic processor. Slow activation prevents word meanings connecting before the new words are encountered within the sentence. Knowing the meanings of lots of words will not be sufficient for comprehension if an excessive time is taken to activate those meanings (Oakhill et al., 2012).

A growing body of evidence related to the SVR supports the claim that vocabulary has an instrumental influence on word recognition and reading comprehension. Muter et al. (2004), found vocabulary was the most significant predictor of reading comprehension. Ricketts et al. (2007), and Sénéchal et al. (2006), also found vocabulary significantly predicts reading comprehension once word reading skills are more established. Ouellette (2006) concluded that the depth of vocabulary knowledge predicted reading comprehension and the relationship between decoding and vocabulary was primarily a function of the size of the receptive (phonological) lexicon. In Ouellette’s study, further measures were taken by dividing vocabulary into breadth (how many words one knows) and depth (how well one knows those words) which highlighted the importance of semantic representations and the size of the lexicon. Ouellette reported reading comprehension was explained by unique
contributions in both vocabulary breadth and depth, outside of the sizable contributions of word recognition in Grade 4 readers. Only receptive vocabulary breadth (words a person can comprehend even if they are unable to produce those words) predicted decoding performance, whereas expressive vocabulary breadth (the ability to use words and put words into sentences) predicted word recognition, depth of vocabulary knowledge predicted reading comprehension (Ouellette, 2006).

The indirect relation between vocabulary and reading comprehension found by the Language and Reading Research Consortium (2015) may be important because words form the basis of sentences and longer units of text and because vocabulary supports inferencing. However, the Language and Reading Research Consortium conceded other forms of knowledge also enable successful comprehension. Furthermore, Lepola, Lynch, Laakkonen, Silvén, and Niemi (2012) found children as young as 4-years-old generated inferences from picture books which predicted vocabulary knowledge 1-year later, and in turn predicted their language comprehension.

A reciprocal relationship exists between vocabulary knowledge and reading comprehension. Verhoeven, van Leeuwe, and Vermeer (2011) found vocabulary predicted early word decoding and reading comprehension. Their data provided support for the lexical quality hypothesis, namely, the knowledge of word forms and word meanings which predicted the development of reading comprehension. Seigneuric and Ehrlich (2005) found vocabulary skills in 7 and 8-year-olds were related to reading comprehension later, at age 9. This finding is consistent with other studies that highlight vocabulary as a strong predictor of reading comprehension in the early years (Bast & Reitsma, 1998; de Jong & van der Leij, 2002). Stanovich
(1986) argued if the development of vocabulary knowledge significantly facilitates reading comprehension, and if reading itself is a major mechanism to improve vocabulary knowledge, then reading should continue to initiate further vocabulary development. Therefore, those with good comprehension (or good vocabulary), will read more and go on to build and improve their vocabulary (or comprehension). The proposal suggests vocabulary is necessary for both word recognition and language comprehension (Oakhill et al., 2014).

Oakhill et al. (2014) investigated the way in which vocabulary contributes to the acquisition of reading comprehension. Their findings suggest the most obvious way was by comprehending the meanings of words in text because vocabulary is necessary to understand the message being conveyed. An unknown word meaning may result in a vague understanding along with further difficulty comprehending information on the same topic. However, it is not usually necessary to stop and look up all the unknown words because an approximate meaning of a new word can usually be worked out by making a reasonable inference from the context of the sentence (Oakhill et al., 2014).

Vocabulary knowledge is not static and changes across year levels. Children entering school with a large vocabulary have approximately 6,000 words and the average high school student knows about 45,000 words by Year 11 (Stahl & Nagy, 2005). Ouellette and Beers (2010) demonstrated the importance of vocabulary in word reading, decoding and reading comprehension with the variance occurring across different age groups. They, along with Muter et al. (2004), agree that vocabulary predicts reading comprehension but not word recognition in children below Grade 3. Ouellette and Beers (2010) and others (Protopapas et al., 2012;
Tunmer & Chapman, 2012) suggest the contribution of vocabulary knowledge to word recognition and reading comprehension has been shown to increase between about 7 and 10-years-old, which is beyond Grade 3.

The Language and Reading Research Consortium (2015) pinpoint Grade 2 as the transition point at which the contribution of vocabulary to word recognition occurs adding further support for a change in weightings of the two components of the SVR. The change in development occurs as the developing reader acquires faster and more automatic word recognition (Gough, Hoover, & Peterson, 1996). As the contribution of word recognition to reading comprehension reduces, language comprehension accounted for more of the individual difference in reading comprehension. As children become more skilled and fluent at decoding words, vocabulary becomes more important as a predictor of reading comprehension skill (Oakhill et al., 2014). An enhanced vocabulary allows the reader to concentrate on comprehension rather than having to stop and decode words (Ehri, 2005a). Another plausible reason for the increasing contribution of vocabulary knowledge is the occurring change in the instructional focus. Relatively predictable texts with easily decoded words are used at the onset of formal literacy instruction. Increasingly complex texts containing challenging vocabulary are used as children get older and become more competent readers (Oakhill et al., 2014).

The findings from the Language and Reading Research Consortium (2015) supported claims that vocabulary was the specific basis for the indirect influence on word recognition in two ways. First, vocabulary knowledge reflects consolidated knowledge about familiar individual word forms and second, an extensive vocabulary supports the processing of unfamiliar words through strategies such as reading by
analogy or identifying spelling patterns within the words (Ehri, 2014). According to the research, there appears to be little doubt that vocabulary influences reading comprehension. Words need to have been previously heard and the meanings need to be understood to comprehend what is being read. However, vocabulary is not as important for word recognition because isolated, unknown words can be read correctly without knowing the meaning of the word.

In concluding, quick and accurate access to words and meanings is necessary for comprehension (Sénéchal et al., 2006; Wasik, 2010). Effective orthographic and phonological cognitive processors are needed to automatically activate the semantic processor and the meanings of target words. Automatic activation of the meanings of words allows the reader to focus on comprehension instead of stopping to decode words which is crucial in successfully learning to read (Ehri, 2005b). Vocabulary is one aspect of language comprehension increasingly noted as being a significant contributor in this process. The reciprocal nature of vocabulary development facilitates reading comprehension and the skill of reading initiates growth in both vocabulary knowledge and understanding (Oakhill et al., 2014). The development is not static and the importance of vocabulary development increases with age as reading to learn becomes more prevalent. While vocabulary is increasingly becoming recognised as being a contributor to word recognition and reading comprehension, further evidence is needed to support the few studies which have been previously carried out. This highly interactive cognitive process of reading is too complex to be understood from directly observable behaviours. However, set for variability is a skill which is said to be instrumental in adding an observable measure of the cognitive reading process.
Set for Variability

Set for variability is a skill thought to be possessed by some beginning readers which enables them to generate approximate phonological representations of unknown words, that in turn, moves the reader close enough to the correct phonological form (Venezky, 1999). To foster set for variability, children learn to use their developing knowledge of spelling-to-sound relationships to produce approximate phonological representations, or partial decodings, for unknown words, especially those containing irregular, polyphonic, or orthographically complex spelling patterns. The phonological representations provide the basis for producing an alternate pronunciation of the target word until one is generated that matches a word in the child’s lexical memory and makes sense in the context in which it is being read. Therefore, if the child produces a word that does not sound like something already stored in the vocabulary then the child has to change at least one of the sound associations and attempt the word again (Venezky, 1999). This is the phonological recoding process in action as discussed previously.

A measure of the skill, set for variability, was introduced by Tunmer and Chapman (2012) in a study carried out in Aotearoa, New Zealand. This pertinent research investigated the role of vocabulary knowledge in the development of word recognition skills, and more specifically the possibility that this influence is mediated by set for variability.

Aotearoa, New Zealand Research: Tunmer and Chapman (2012)

Tunmer and Chapman (2012) conducted a 3-year longitudinal study in Aotearoa, New Zealand. Participants included one hundred and forty children, aged 4 years 11
months to 5 years 3 months. The findings indicated a moderately high correlation between vocabulary and set for variability which indicated that children will not be able to determine the correct pronunciation of mispronounced words if the target word is not in their listening vocabulary or is only weakly represented in the lexicon. Moderate relationships were evident between vocabulary and context-free word recognition. Tunmer and Chapman suggest the results are consistent with the hypothesis that set for variability has a more direct influence on the development of word recognition skills than does vocabulary. Furthermore, context-free word recognition made significant independent contributions to variance in reading comprehension. The findings reinforce the point that children need to accurately recognise the words of text and successfully decode unfamiliar words or they will be limited in their ability to comprehend text.

Tunmer and Chapman (2012) found Year 1 (new entrant) vocabulary directly influenced Year 3 reading comprehension and indirectly influenced Year 3 word recognition through Year 1 set for variability. They also found set for variability influenced reading comprehension indirectly through both decoding skill and word recognition. These results are consistent with the research findings from Sénéchal et al. (2006) that vocabulary is one of the best predictors of reading comprehension and highlights that vocabulary contributes to the development of word recognition skills indirectly through set for variability. Set for variability contributes to the development of both decoding and word recognition skills by enabling children to identify unknown words based on partial decodings. Additional letter-sound patterns can be induced through implicit learning which are activated, from the orthographic representation of the words stored in lexical memory.
Tunmer and Chapman (2012) suggest set for variability may play an important role in the development of word recognition skills but cautioned against drawing any firm conclusions. They recommend further research and training studies to determine if exposing children to activities designed to promote the development of set for variability can produce a significant increase in this skill which will have an effect on word recognition skills. In their study, the extent of additional variance accounted for by set for variability was relatively small, suggesting the results are spurious. Tunmer and Chapman also suggest the test for reading comprehension, Neale Reading Comprehension (Neale & Konza, 2001) may have influenced the results. Distortions have been evident in reading comprehension tests which use shorter passages but the Neale Reading Comprehension uses passages which are eight sentences in length except for the first passage. Tunmer and Chapman’s study added further evidence to support the claim that vocabulary influences word recognition and reading comprehension.

The Present Study

The role of word recognition (as a measure of decoding) and language comprehension has been firmly established in learning to read in previous investigations. Contemporary research has focused on the role of vocabulary in reading comprehension and, moreover, the role of word recognition. Most of the evidence gathered has involved older children of up to 13-years-old. A limited number of studies have provided evidence from children starting school. Evidence surrounding the influence of vocabulary on word recognition and reading comprehension from this younger age group is pertinent. Children with low scores identified at the emergent literacy learning stage can allow for specific teaching
strategies to be identified and implemented which could prevent the achievement gap from widening. The evidence from these international and national studies provides the foundation for this study.

The current study is a modified replication of the Tunmer and Chapman (2012) study but differs in terms of the number of participants, their age, and the times in which the measures are carried out. The main difference is the participants of this study are a year younger than the participants of the Tunmer and Chapman study at the outcome point. The aim of this study is to add further evidence to the growing body of existing research to determine if vocabulary knowledge directly influences both word recognition and reading comprehension, and if vocabulary knowledge indirectly influences word recognition through set for variability. These concepts will be explored through the following two questions:

1. *Does vocabulary knowledge at the beginning of the Year (Time 1) contribute to word recognition at the middle of Year 2 (Time 3) directly or indirectly through set for variability (Time 2)?*

2. *Does vocabulary knowledge at the beginning of Year (Time 1) contribute directly to reading comprehension (Time 3)?*
Chapter Three

Methodology

This chapter describes the methodology of the current study. The chapter begins with an explanation of the research context and the research design followed by a description of the recruitment and sampling procedures and a description of the research setting and the participants. These descriptions are followed by an outline of the procedures used for each assessment and the chapter concludes with the ethical considerations.

Research Design

The present study uses secondary data analysis to explore the relationships between vocabulary knowledge and reading comprehension in new entrant children. Data are drawn from a larger 3-year longitudinal study in progress in New Zealand examining teacher knowledge and practice in new entrant classrooms (Chapman et al., 2016). The comparison and intervention groups from the larger study were combined in the current study due to both groups from the original study performing at comparable levels at the end of the first year. The sole group used in this study was administered selected reading related measures at three time points over an eighteen month period. Data were used to examine two hypotheses about the relationships between vocabulary and word recognition and vocabulary and reading comprehension using regression analysis.
The first hypothesis states that the receptive vocabulary knowledge of a child, when they commence school, directly contributes to the child’s word recognition skills in the middle of the second year at school and indirectly through a mechanism known as set for variability. The second hypothesis states that the vocabulary knowledge of children when they commenced school predicts reading comprehension ability at the middle of Year 2 but is not mediated by set for variability.

**Sampling**

The sample for the current study was taken from a larger research design which selected participants as clusters within the school level using a random selection procedure. Cluster sampling is a probability technique which randomly selects and uses whole, naturally occurring groups (Ary, Jacobs, Razavieh, & Sorensen, 2014). In this study, the cluster is the school. The sample was randomly selected from the population of schools in the lower North Island, with the assumption that the sample is reasonably representative of that population.

The randomised selection procedure was carried out by ranking the entire population of lower North Island schools by the number of new entrants enrolling at the start of the school year in 2013. The schools were then grouped into clusters of decile level: 1-3, 4-7, 8-10. Within each decile cluster, schools which enrolled less than 14 children were deleted, with the exception of the decile 1-3 cluster. The number of schools within this cluster was lower than the set level of representation, therefore the cut-off for student enrolment was set to 10. Subsequently, all schools in the decile 1-3 cluster that met the criteria of 10 new entrants in term 1 of 2013 were included in the sample. For the other two decile clusters, a random sampling of a set number was used to select the specified percentage of schools.
The schools in each decile cluster were randomly selected as either intervention or comparison. This study combined the two groups into a sole group as both groups performed at comparable levels at the end of the first year of the Early Literacy Project study (Chapman et al., 2016). After the initial sample was conducted, further schools were excluded for specific reasons. These reasons included prior participation in other small scale longitudinal projects in which the researchers were associated (n = 15); the request of the Ministry of Education that a specific area not be included in the sample as many schools were already targeted in other programmes (n = 2); were part of a specific school programme that did not have English medium or mainstream teaching practices (e.g., Kura or Steiner; n = 3) or that they were geographically too isolated to be practical for the project team (n = 6).

Sampling fell into the following cluster areas, Hutt Valley-Wellington, Manawatu-Wanganui, Wairarapa, and Kapiti-Horowhenua. Cluster meetings, arranged by the Ministry of Education, provided information to school Principals about the project. A forum was arranged for Principals to indicate their willingness to participate in the research. The schools were then grouped into clusters according to the decile level: 1-3, 4-7, 8-10. Schools were then recruited into the wider project. The final sample included 38 schools, 31% in the 1-3 deciles, 44% in the 4-7 deciles, and 25% in the 8-10 deciles.
Setting and Participants

The children for the present study were enrolled in 38 urban schools from the mid to lower North Island and were located in a range of socioeconomic areas. The participants were drawn from a cross-section of school deciles and types using the procedure previously stated. The data set for this study comprised of 374 children.

Slightly more boys than girls were included in the sample. There were 197 boys (52.4%) and 177 girls (47.1%). The ethnicity breakdown was predominantly (see Table 1) New Zealand European/Pakeha (55.1%), followed by children who were identified as being Maori (22.3 %), children of Pacific Island decent (5.1%), children who identified as being Asian (2.9%), others (2.4%) who did not specify an ethnicity and the remainder (12.2%) of families did not wish to disclose their ethnicity.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>197</td>
<td>52.4</td>
</tr>
<tr>
<td>Female</td>
<td>177</td>
<td>47.1</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
<td>.5</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NZ European/Pakeha</td>
<td>207</td>
<td>55.1</td>
</tr>
<tr>
<td>Maori</td>
<td>84</td>
<td>22.3</td>
</tr>
<tr>
<td>Pasifika</td>
<td>19</td>
<td>5.1</td>
</tr>
<tr>
<td>Asian</td>
<td>11</td>
<td>2.9</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>374</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Data Collection Procedures

All participating children were individually assessed on four measures designed to assess vocabulary knowledge, word recognition, mispronunciation (set for variability) and reading comprehension. These measures were part of a battery of assessments given in a larger longitudinal study (Chapman et al., 2016). Research assistants carried out the assessments during each round of testing with the researcher carrying out some assessments on some occasions. Testing took place on three occasions over an eighteen month timeframe. The vocabulary measure was conducted upon commencement of school during February 2015. The mispronunciation task was carried out at the end of the child’s first year at school during November 2015, while the word recognition measure and reading comprehension measures were carried out in the middle of Year 2 during June/July 2016.

Data were collected at three different timeframes. The early literacy assessments were taken at valuable time points and analysed. The analysis of this data can assist with the identification of early literacy support and interventions. In turn, this could contribute to advanced reading achievement in the early years of school.

All testing was conducted on a one-to-one basis in a relatively quiet space outside the classroom environment but within each participating school. The tests were administered in random order but each individual assessment was completed in one attempt. While most children completed the assessments in one sitting, some assessments were conducted over two or more sittings, sometimes on separate days. The number of times children were seen for the assessments varied and was dependent upon the child’s ability to focus and concentrate or on how tired they were
and where the school breaks fell during the day. Some assessments were not completed due to ongoing student absentee which made it difficult to carry out the necessary assessments on subsequent occasions. Overall the entire battery of assessments took between 45 minutes and 1 hour 15 minutes for each child.

The four measures used to collect the data were not changed during the course of the study. However, different research assistants performed the assessments and collected the data with different children. The data collectors (research assistants) were trained and were provided with detailed guidelines explaining how to administer each assessment which reduced the possibility of each assessment being administered differently. Four measures were used to collect data on vocabulary knowledge, set for variability, word recognition and reading comprehension. These measures were namely the *British Picture Vocabulary Scale (BPVS III)* (Dunn, Dunn, Styles, & Sewell, 2009), the *Mispronunciation Task* (Tunmer & Chapman, 1998), the *Burt Word Reading Test* (Gilmore, 1981) and the *Neale Analysis of Reading Ability* (Neale & Konza, 2001).

**Measures**

**Vocabulary Knowledge - British Picture Vocabulary Scale (BPVS III)**

Vocabulary knowledge was measured using The British Picture Vocabulary Scale (BPVS III) Third Edition (Dunn et al., 2009). Administration of the BPVS III commenced with two training items using test plates. The child was shown a set of four pictures on a page and asked to point to the picture which depicted the word provided by the research assistant. For example, “Put your finger on cup.” The test items followed the same procedure as the training items and were conveyed in sets
containing 12 items. The test scores were based on the number of errors made by the student. Test items began at set 4 because this set targeted children who were 5 to 6-years of age. A basal set was established by finding the set in which the student made no more than one error in a set. If this occurred on the first set, then Set 4 was considered to be the basal set and the test was continued forward to find the ceiling set. If the student made more than one error in the first set, then the preceding set was administered. A ceiling set was obtained when the student made 8 or more errors within a set of 12 items. Raw scores were calculated by taking the number of the last item in the ceiling set and subtracting the total number of errors. The standardised score was found by converting the raw score using the provided conversion tables. The internal reliability estimate for this scale was .81.

Set for Variability - Mispronunciation Task

Set for variability, a child’s ability to determine the correct pronunciation of mispronounced spoken words, was measured using an adapted version of a task developed by Tunmer and Chapman (1998). The words were derived from regularised pronunciations of irregularly spelled words, the incorrect pronunciation of words containing polyphonic spelling patterns, and approximations to correct pronunciations based on the application of context-free spelling rules. The task comprised of 40 mispronounced words presented in isolation using a recording on an iPad. For example, wasp was pronounced /wăsp/. The task was presented in the form of a game where a ‘friend’ tries to say a word but the ‘friend’ says the words the “wrong way.” The child had to figure out what the ‘friend’ on the iPad recording was trying to say. The task was administered in one session but a break was allowed between each set of 20 words if necessary. Scoring was based on the number of words correctly identified. The set for variability task was related to the children’s
ability to go from a partial decoding generated from emerging decoding skills to word identification (Tunmer & Chapman, 2012). The internal reliability for set for variability (context) was .86.

*Context-free Word Recognition - The Burt Word Reading Test*

The Burt Word Reading Test, New Zealand Revision (Gilmore, 1981) was used to assess context-free word recognition at the middle of Year 2. The Burt was revised and standardised for use in New Zealand in 1981. Burt is a quick and easy assessment of word recognition and decoding skills and enables a broad estimate of a child’s reading achievement. The administration of the test involved the child being presented with a printed list of 110 words which decreased in font size and type while increasing in difficulty. The child was asked to look at each word carefully and read the word aloud. The test continued until 10 successive words were read incorrectly or were not attempted. Scoring is based on the number of words read correctly. The Burt Word Reading Test has a reliability coefficient of .97.

*Reading Comprehension - The Neale Analysis of Reading Ability*

The Neale Analysis of Reading Ability revised third edition (Neale & Konza, 2001) subtest, Form 1, was used to assess reading comprehension ability at the middle of Year 2. The children were asked to read aloud a series of narratives and detailed descriptions which were ordered in level of difficulty. All passages were eight sentences in length except for the Level 1 passage which contained four sentences. The words increased from 26 words for the Level 1 passage to 505 words for the Level 6 passage. Upon completion of each passage the children were asked a series of questions relating to the passage. Scoring procedures were used to
calculate the comprehension. One correctly answered question equated to 1 point. The internal reliability estimate was .89.

Ethics

Because a secondary data analysis was being carried out, separate Ethics Committee approval was not required for this study as ethics approval was granted for the primary study, the Early Literacy Project (Chapman et al., 2016). The data were identified by the project student identification numbers and the names of all participants were deleted. This ensured the anonymity of all participants.

Consent was sought from parents via schools for all children who turned 5-years-old and started school from the 1st November 2014 to the 28th February 2015. An information letter had been sent to parents for the larger project and a consent form was signed by those parents who gave permission for their children to participate in the study. Verbal assent was given by these children to carry out the assessments. Data were only used from children whose parents had given consent for data collection.

Summary

This chapter described the methodology of the current study. The aims of the study were, first to examine the independent contribution of vocabulary to word recognition, and whether the contribution was direct or indirect through set for variability. A second aim of the study was to determine if a direct relation exists between vocabulary knowledge and reading comprehension. The study involved a randomised design using one group. The schools, teachers and the parents of the
children (participants) within specific geographical areas were invited to participate in the study. The assessments took place in the school environment at three time points, over eighteen months.
This chapter presents the results of the analyses carried out to investigate the two hypotheses regarding the role of vocabulary knowledge in the development of word recognition skills.

The purpose of the study was to investigate if vocabulary accounts for a unique proportion of variance in both word recognition and reading comprehension for this sample of young readers. In addition, this study aimed to investigate if set for variability mediates the relationship between vocabulary and word recognition. Finally, the nature of the relationship with reading comprehension, over and above the effect on word recognition was examined. The independent variables were receptive vocabulary, and the mispronunciation (set for variability) task. The dependent variables were word recognition (as a measure of decoding) and reading comprehension. Word recognition was also an independent variable when using reading comprehension as the dependent variable. Initial descriptive results and correlations are presented first, followed by the regression analyses.

**Descriptive Results**

Table 2 provides the correlations, means and standard deviations for all tests administered to the children. The vocabulary test was administered to children aged 5, in February 2015. Set for variability was administered in the form of a mispronunciation test at the end of the children’s first year at school, in November
The word recognition test and the reading comprehension tests were administered in the middle of the children’s second year at school, in June/July 2016.

Table 2 Correlations, Means and Standard Deviations for all Measures

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vocabulary T1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Set For Variability T2</td>
<td>.348**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Word Recognition T3</td>
<td>.397**</td>
<td>.511**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Reading Comprehension</td>
<td>.416**</td>
<td>.466**</td>
<td>.819**</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>98.65</td>
<td>5.75</td>
<td>26.99</td>
<td>6.04</td>
</tr>
<tr>
<td>SD</td>
<td>11.647</td>
<td>4.310</td>
<td>16.456</td>
<td>5.086</td>
</tr>
<tr>
<td>Maximum Score</td>
<td>175</td>
<td>40</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>349</td>
<td>286</td>
<td>290</td>
<td>275</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

All measures were significantly correlated, and the magnitudes of the correlation coefficients ranged from moderate to high. There is a significantly high positive correlation between word recognition and reading comprehension. This finding indicates that the ability to correctly pronounce words is essential for gaining meaning when reading. Moderate relationships were found between set for variability and word recognition, set for variability and reading comprehension and vocabulary and reading comprehension. The correlations between vocabulary and word recognition and vocabulary and set for variability were lower, but were also significant.
The correlations were lower than were to be expected. The mean score for vocabulary ($M = 98.65$, $SD = 11.647$) was slightly above the mid-range. The set for variability scores ($M = 5.75$, $SD = 4.310$) and the word recognition scores were low ($M = 26.99$, $SD = 16.456$). The reading comprehension scores were also low ($M = 6.04$, $SD = 5.086$).

Among the results there were some varied and unexpected findings. While some children had low vocabulary scores, low set for variability scores, low word recognition scores and low reading comprehension scores other children had varied scores between each of the measures. Some children with low vocabulary scores and low set for variability scores, recorded high word recognition scores, but had low reading comprehension scores. Some children with moderate vocabulary scores had high set for variability scores, high word recognition scores and high reading comprehension scores. Similar findings were evident in children with high vocabulary scores. These children generally had high set for variability scores, high word recognition and reading comprehension scores but interestingly some children had low reading comprehension scores.

Children from different backgrounds start school with very different-sized vocabularies. Beginning readers with smaller vocabularies and limited vocabulary knowledge continue to have limited vocabulary knowledge and are disadvantaged if the poor vocabulary cannot be easily compensated for (Oakhill et al., 2014). This was evident in this study where the children with low vocabulary scores on entry to school, continued to have low reading comprehension after eighteen months at school. This finding is consistent with the finding from other researchers (Nation,
Clarke, Marshall, & Durand, 2004) where children with poor reading comprehension displayed low levels of vocabulary knowledge.

The stronger relationship between set for variability and word recognition indicates set for variability has a stronger influence on the development of word recognition skills than vocabulary. This pattern of results is consistent with the hypothesis that set for variability has a more direct influence on the development of word recognition skills than vocabulary does. The longitudinal nature of this research means that tentative causal explanations can be explored through correlational analyses. In this study regression analyses are used to explore the relationships in more detail.

**Contributions to Word Recognition**

The first hypothesis states vocabulary knowledge at the beginning of Year 1 (Time 1) will contribute to word recognition at the middle of Year 2 (Time 3) both directly and indirectly through set for variability. A multiple regression was carried out to explore the direct and indirect contribution of set for variability (mispronunciation task) to determine which explained a greater share of the variance than the direct contribution. At the first step, vocabulary was entered alone, and set for variability was added at the second step.

The analysis was run to determine if vocabulary knowledge at the beginning of the year (Time 1) contributed directly to word recognition at the middle of Year 2 (Time 3) and if vocabulary knowledge at the beginning of the year (Time 1) indirectly contributed to word recognition at the middle of Year 2 (Time 3) indirectly through set for variability. Table 3 presents the model summary and the coefficients for each of
the variables. The Table provides information about the ability of the independent variables, vocabulary and set for variability, to account for the total variation in the dependent variable, word recognition.

Table 3 Model Summary and Coefficients

<table>
<thead>
<tr>
<th>Step</th>
<th>R</th>
<th>R²</th>
<th>Beta</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Vocabulary</td>
<td>.374&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.140</td>
<td>.374</td>
</tr>
<tr>
<td>2.</td>
<td>Vocabulary</td>
<td>.204</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Set for</td>
<td>.541&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.293</td>
<td>.426</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), T1BPVSstandard  
b. Predictors: (Constant), T1BPVSstandard, T2mispronunciation  
Dependent Variable: T3Burtwordreading

At Step 1, vocabulary at Time 1 explained 14% of variance in Time 3, word recognition when entered separately. As indicated in the correlations (see Table 2) this variance is statistically significant. At Step 2 there is a greater amount of variance in word recognition explained, although the contribution of vocabulary decreases. Set for variability, when entered after vocabulary explains more variance in word recognition. This finding illustrates that vocabulary at school entry directly contributes to word recognition but it also indirectly contributes to word recognition through set for variability.

The multiple regression analysis in model 2, ($R^2$ equals 0.293) displays a higher proportion of variance and indicates that 29.3% of variance in word recognition can be explained by vocabulary and mispronunciation jointly. A further 15.3% additional variance accounted for in word recognition when set for variability
(mispronunciation), assessed mid-year, is added in. A child using the mechanism, set for variability, will have a 15.3% greater chance of correctly recognising a word from a partial or close attempt at an unknown word using the vocabulary stored in the lexicon than will a child who only attempts to decode a word using the bank of words stored in the lexicon.

Vocabulary made an independent contribution to word recognition (as indicated by the beta value of .374). The addition of mispronunciation added a significant contribution to word recognition (as indicated by the beta value of .630). Therefore, the addition of mispronunciation greatly increases the accuracy in prediction of children's scores in word recognition, over and above what vocabulary can predict on its own.

**Contributions to Reading Comprehension**

The second hypothesis states that vocabulary knowledge at the beginning of Year 1 (Time 1) will contribute directly to reading comprehension in the middle of Year 2 (Time 3) but is not mediated by set for variability. A multiple regression was carried out to explain the additional variance in the scores when mispronunciation was added as a mediator. Finally, word recognition was entered to examine the role of vocabulary on reading comprehension when word recognition is entered into the model.

To test the second question, a second set of regression analyses using three models were carried out to determine if vocabulary knowledge at the beginning of Year 1 (Time 1) contributed directly to reading comprehension or indirectly through set for
variability (see Table 4). A simple linear regression was carried out to determine if there is a direct contribution which vocabulary makes to reading comprehension. Two multiple regressions were then carried out to explain the variance in scores on reading comprehension and vocabulary when set for variability was added and to determine the variance that vocabulary and word recognition have jointly on reading comprehension when set for variability is added.

Table 4 Model Summary and Coefficients

<table>
<thead>
<tr>
<th>Step</th>
<th>R</th>
<th>R²</th>
<th>Beta</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.404a</td>
<td>.163</td>
<td>.404</td>
<td>.000</td>
</tr>
</tbody>
</table>

2. Vocabulary

<table>
<thead>
<tr>
<th>Set for</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variability</td>
<td>.519b</td>
<td>.269</td>
<td>.354</td>
<td>.000</td>
</tr>
</tbody>
</table>

3. Vocabulary

<table>
<thead>
<tr>
<th>Set For</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variability</td>
<td>.404a</td>
<td>.163</td>
<td>.130</td>
<td>.002</td>
</tr>
<tr>
<td>Word</td>
<td>.826c</td>
<td>.682</td>
<td>.760</td>
<td>.000</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), T1BPVStandard
b. Predictors: (Constant), T1BPVStandard, T2mispronunciation
c. Predictors: (Constant), T1BPVStandard, T2mispronunciation, T3Burtwordreading
Dependent Variable: T3Reading Comprehension

At Step 1 the direct relationship was explored between vocabulary and reading comprehension using a simple linear regression. The simple linear regression analysis, at Step 1 ($R^2 = 0.163$) means that 16% of variance in reading comprehension can be explained by vocabulary. The multiple regression analysis at Step 2 ($R^2 = 0.269$) means that 26% of variance in reading comprehension can be
explained by vocabulary and mispronunciation jointly. A further 10% additional variance accounted for in reading comprehension when set for variability (mispronunciation), assessed mid-year, is added. A child possessing the skill of set for variability will have a 10% greater chance of understanding what they read than if they only drew on their vocabulary.

The multiple regression analysis at Step 3 ($R^2 = 0.682$) means that 68% of variance in reading comprehension can be explained by vocabulary, word recognition and set for variability jointly. This is a significantly higher proportion of variance together than when considered individually. However, set for variability no longer explains significant amounts of shared variance in reading comprehension.

Vocabulary made an independent contribution to reading comprehension (as indicated by the beta value of .404). The addition of mispronunciation only slightly increased the contribution to reading comprehension (as indicated by the beta value of .618). Despite being small, the addition of mispronunciation still slightly increases the accuracy in prediction of children’s scores in reading comprehension, over and above what vocabulary can predict on its own. Set for variability does not directly contribute to reading comprehension when word recognition is added to the model (as indicated by the beta value of 0.020). Therefore, the addition of mispronunciation did not increase the accuracy in the prediction of children’s scores in reading comprehension over and above what vocabulary predicted on its own.

Each variable contributed significant amounts of shared variance when entered individually, as indicated by the significant $R^2$ change. However, in terms of significance when all were entered, single word reading subsumed all the variance
contributed by the mispronunciation task. This suggests that although set for variability contributes to reading comprehension, it is mediated through word recognition, rather than a direct contributor. The standardised beta scores indicate that both vocabulary and word recognition independently contribute to reading comprehension in the middle of Year 2 for these children. Word recognition explains more of the variance than vocabulary.

Summary

Word recognition and vocabulary both directly contribute to reading comprehension for these children in the middle of Year 2. In keeping with previous research findings, word recognition explains a greater amount of the shared variance of reading comprehension at this point in schooling as they are still trying to read a range of unfamiliar words. Set for variability does not directly contribute to reading comprehension when word recognition is added to the model. This suggests that set for variability mediates the relationship between vocabulary and word recognition but not for overall reading comprehension.
The aim of this study was to investigate the influence of vocabulary in the development of word recognition skills and reading comprehension. More specifically, the study examined if the role of vocabulary on the growth of word recognition skills is mediated by set for variability, which is deemed to be the ability to generate an approximate phonological representation of an unknown word (Venezky, 1999).

Data to explore the relationships between vocabulary knowledge and reading comprehension in new entrant children were drawn from a larger 3-year long longitudinal study in Aotearoa, New Zealand examining teacher knowledge and practice in new entrant classrooms (Chapman et al., 2016). Secondary data analyses were used to explore the relationships between vocabulary knowledge and word recognition and reading comprehension in new entrant children.

Two hypotheses were investigated in this study. It was first hypothesised that the receptive vocabulary knowledge of a child, when they commence school, directly contributes to the child’s word recognition skills in the middle of the second year at school and indirectly through a mechanism known as set for variability. The second hypothesis was that the vocabulary knowledge of children when they commenced school predicts reading comprehension ability at the middle of Year 2 but is not mediated by set for variability. Both hypotheses were supported by the findings.
The Direct and Indirect Influence of Vocabulary on Word Recognition

Regarding the first hypothesis investigated in the current study, vocabulary directly contributed to word recognition for these children in the middle of their second year of school. In addition, vocabulary contributes to the development of word recognition skills through the variable called set for variability, the ability to determine the correct pronunciations of approximation to spoken English words (Venezky, 1999). The current findings are consistent with those of Ricketts et al. (2007) who found that vocabulary plays an important role in the development of word recognition. They argued that this is because printed word reading is closely connected to children’s underlying oral language skills.

The support for the importance of vocabulary in word recognition found in this study is also consistent with other developmental studies (Muter et al., 2004; Ouellette, 2006; Ouellette & Beers, 2010; Sénéchal et al., 2006; Tunmer & Chapman, 2012). Additionally, while Muter et al. (2004), Ouellette (2006) and Ouellette and Beers (2010) failed to find a significant contribution of vocabulary breadth to decoding in Grade 1 (6-7-year-olds), they found that vocabulary did predict irregular word recognition and regular word decoding in older children. The word reading measure used in the current study contained a mix of regular and irregular words, suggesting that vocabulary contributed to the reading of both types.

Whitehurst and Lonigan (2002) have argued that within the lexicon, the word and spelling sound relationships are stored, processed and retrieved when a word is sounded out. A correct phonological representation of a word is made when the letters in the word are decoded and the word is pronounced correctly. Children become more and more sensitive to the ways in which letters are linked to sounds
and the sounds are assembled into a whole word. Sounds assembled into whole words are matched accordingly and read as sight words.

This study also found set for variability had a stronger influence on the development of word recognition skills than vocabulary which is also consistent with the findings from Tunmer and Chapman (2012). When encountering an unknown word, the reader does not always make a correct attempt. An incorrect decoding attempt can cause the reader to either stop or try a partial decoding. Venezky (1999) claims, set for variability is a mechanism which can support beginning readers. Effectively, if beginning readers use set for variability to support the decoding of words, they implicitly (using phonological recoding) draw upon the spelling sound relationships stored in the lexicon. A partial decoding of an unfamiliar printed word, close to the phonological form and using the context (not used in this study) of the sentence or paragraph, and for which the corresponding spoken version of the printed word is stored in lexical memory, is likely to result in an accurate decoding of the word using set for variability. A greater chance of this exists if there are a high number of phonological entries in the lexicon. The number of phonological representations in the lexicon is most important in terms of facilitating implicit learning of letter-sound associations and identifying unknown words (the number of semantic representations) (Ouellette, 2006). Hence, vocabulary in isolation cannot support word recognition. The reader needs to be able to make a partial attempt by starting to sound out the word to make use of the vocabulary stored in the lexicon. Which is why set for variability explains more variance.

As Nation and Snowling (1998) and Perfetti (2007) argue, children with poorly developed vocabulary knowledge will have difficulty identifying and allocating the
correct meanings to unknown printed words (especially partially decoded words, irregularly spelled words, or words with complex spelling patterns) if the corresponding spoken words are not in their listening vocabulary or are only weakly represented phonologically in the lexicon. The limited number of semantic representations stored in the lexicon limits decoding skills and hinders the ability to attach word meanings to the semantic representations because the additional spelling-sound relationships cannot be induced from the stored orthographic representations of words which have previously been correctly identified (Turner & Chapman, 2012). Weakened decoding skills can lead to poorly developed lexical representations and ongoing difficulties in identifying printed words. These findings are supported by (Braze et al., 2016) and colleagues (Ouellette & Beers, 2010; Ricketts et al., 2007; Verhoeven et al., 2011) who found vocabulary to be more influential in older children.

**The Direct Influence of Vocabulary on Reading Comprehension**

The results supported the second hypothesis. Both word reading and vocabulary have a direct influence on reading comprehension but vocabulary is not mediated by set for variability. Word recognition explains more variance in reading comprehension than vocabulary at this point in children’s’ reading development. However, vocabulary does explain more shared variance in reading comprehension than it did for word reading, which is similar to earlier findings with Dutch children from Verhoeven et al. (2011). Key studies have shown that word recognition explains a greater proportion of variance in reading comprehension in the early grades compared with listening comprehension (Catts & Kamhi, 2005; Language and Reading Research Consortium, 2015; Oakhill et al., 2012). Listening comprehension
is used in such studies as a proxy for vocabulary and therefore, this study supports such studies in that word recognition explains more of reading comprehension than vocabulary at the early levels of reading development.

In much the same way that vocabulary influences individual words, vocabulary influences sentences and sections of text. The available evidence suggests the link from vocabulary to comprehension is derived from the detailed knowledge of a word’s meaning. This involves the ability to activate relevant aspects of a word’s meaning (and meaning-related words) and the use of the information to support comprehension. The reader draws upon the representations of individual words stored in lexical memory when assigning a meaning to the word being read. The context of the sentence or paragraph in which the word is located is used to assist this process further (Kendeou, Savage, et al., 2009; Lorch & van den Broek, 1997; Oakhill et al., 2012; Vellutino et al., 2007). The meanings of words must have been previously heard and understood to comprehend what is being read. Accurate word recognition is also needed for reading comprehension because if a specific word was unknown or could not be worked out, the reader would only have a vague understanding of the paragraph being read. Further difficulty would also be experienced in understanding additional information on the same topic (Oakhill et al., 2014).

Vocabulary and word knowledge are ultimately responsible for reading comprehension. More specifically, the development of reading comprehension is dependent upon the addition of successfully decoded words to a reader’s lexicon. The information (word meanings, spelling sound associations) of successfully decoded words are established in the lexicon to build an individual’s vocabulary. The
growing vocabulary strengthens as words connect with other words to provide a depth of knowledge which the reader draws upon when new material is encountered (Oakhill et al., 2014).

Reciprocal benefits are applied to both vocabulary and reading comprehension. An extensive vocabulary allows an individual to read, and reading in turn, improves vocabulary (Oakhill et al., 2014). That is to say, new vocabulary entries are learned, and existing vocabulary is refined, through reading (even in adulthood). Once children become fluent readers, written text is a major source of new vocabulary (Nagy & Scott, 2000). Associations between limited vocabulary knowledge and reading comprehension difficulties have been found to be evident. Cain and Oakhill (2011) demonstrated that children with specific reading comprehension difficulties have slower rates of vocabulary growth than peers of the same-age with good reading comprehension.

This study is a replication study of that undertaken by Tunmer and Chapman (2012). While Tunmer and Chapman found support for the view that vocabulary has a direct influence on word recognition, and an indirect influence was evident when set for variability was added, the correlations from the Tunmer and Chapman study indicated a slightly higher influence than that of the current study. The final assessments for the Tunmer and Chapman study for word recognition and reading comprehension were taken at the end of Year 3. The period of time is eighteen months later than the final assessments for word recognition and reading comprehension for the current study, but the findings suggest the developmental relationships are the same at both times.
In summary, the results of this study demonstrate how vocabulary knowledge influences reading development in beginning readers and therefore aligns with the Simple View. Vocabulary knowledge is related to word recognition in that it reflects the reader’s knowledge about familiar words and how they are formed. A broad vocabulary also supports the processing of unfamiliar words by using strategies involving similarities of words stored in the lexicon. This study also found vocabulary is related to reading comprehension because it reflects an individual language competence which, in turn, influences reading development. Vocabulary is also influential because words form the basis of sentences and longer passages of text and vocabulary enables the integration and inferencing of text.

**Limitations**

There were limitations which may have influenced the findings of this study and the validity of the results. These include the limitations of data collection and the selection and use of assessment tools. Bias in data has the potential to skew the results and conclusions. Any potential bias from those running the Early Literacy Project was minimised in this study (Chapman et al., 2016).

While the measures used were useful for predicting reading outcomes, Tunmer and Chapman (2012) questioned the accuracy of the reading measure used, the Neale Analysis of Reading Ability (Neale & Konza, 2001). Recent research indicates that commonly used reading comprehension tests vary in the component skills (decoding vs. oral language comprehension) which they assess (Keenan, Betjemann, & Olson, 2008) and the variance can make differential demands on two aspects of oral language comprehension: vocabulary knowledge and sentence-processing abilities.
(Cutting & Scarborough, 2006). Tunmer and Chapman (2012) propose the test of reading comprehension could have influenced the results in their study but further suggest, based on available research, the distortions which occur appear to be mainly due to using reading comprehension tests that involve one or two sentence passages. The identified contribution of vocabulary to reading comprehension is reduced in assessments using shorter passages due to the smaller number of words in comparison with longer passages and where the reader is required to attribute their understanding to the passage (Cutting & Scarborough, 2006; Keenan et al., 2008). While most of the passages in the Neale Reading Comprehension Test are eight sentences in length, the first passage (the lowest level) had fewer sentences. Many of the children did not reach beyond this first passage therefore this could be considered as being a limitation (see Methodology section).

The measure used for set for variability in this study was the correction of mispronounced words out of context. Although Venezky (1999) conjectured the ability to identify approximations to spoken words operates in combination with the use of contexts, the use of decontextualized words meant that vocabulary, and not syntactic knowledge was the key construct targeted.

**Future Research**

The present results have important implications for future research. Suggestions for future research include, expanding the Simple View, further longitudinal studies with younger children and further studies measuring both vocabulary breadth and vocabulary depth.
Recently, a number of researchers (Catts, Adlof, & Weismer, 2006; Tunmer & Chapman, 2012) have proposed that the framework of the Simple View of Reading does not show the complexity very well in terms of the two components, word recognition and language comprehension. As such, researchers have argued additional components should be included. One such component is vocabulary and how the influence of this construct changes predicted reading comprehension over time. This study, and a study by Tunmer and Chapman (2012), added an independent assessment of vocabulary to the Simple View. Further research should be undertaken in which vocabulary is included as an independent factor in the model. This study, along with that of Tunmer and Chapman, found that the inclusion of vocabulary as an independent variable was beneficial in being able to explain the additional variance in reading comprehension over and above word reading in young children. However, further research is needed to support this evidence.

The influence of vocabulary on reading comprehension changes over time and as such, further longitudinal studies with younger children should be carried out. Further research is needed to find out the reasons behind this expected developmental change and to further pinpoint where that change occurs. While patterns showing listening comprehension can account for more variance in reading comprehension than word recognition in older age groups, few studies have empirically confirmed these observations. The Language and Reading Research Consortium (2015) found listening comprehension influences reading comprehension during the earliest stages of reading and identified Grade 2, (7-8-years-old) as the transition point at which listening comprehension becomes more prominent. Further longitudinal research needs to be undertaken from the onset of formal literacy instruction to add evidence to support the observations and identify where the
influence of vocabulary on word recognition and reading comprehension increases and to what extent. Children from different backgrounds start school with very different-sized vocabularies. It would be important to verify if those children with smaller vocabularies, when they start to read, continue to have poor vocabulary knowledge over time and if these children continue to be disadvantaged, as limited vocabulary cannot always be easily compensated for.

Questions around the extent of the influence of vocabulary depth have arisen as a result of the current study. While the influence of vocabulary depth on reading comprehension is currently implied, further studies measuring vocabulary at deeper levels are needed to explain and verify the likelihood of a causal link between vocabulary depth and comprehension difficulties. This is an important future study because shallow levels of vocabulary breadth (measured in this study), is not causally linked to poor reading comprehension. Ouellette and Beers (2010) recommend assessing both breadth and depth of vocabulary and further suggest this could provide a clearer indication about the extent and role of the lexical size and semantic representations and how each support word reading and reading comprehension.

The growing vocabulary strengthens as words connect with other words to provide a depth of knowledge which the reader draws upon when new material is encountered (Oakhill et al., 2014). Ouellette (2006) concluded the depth of vocabulary knowledge which predicts reading comprehension is due to the number of words stored in the lexicon and for which meanings are understood. To emphasise, a relatively deep understanding of words is necessary as it is easier for the reader to understand the text when more is known about the key words. Studies which measure both
vocabulary breadth and depth will go some way in determining which construct is more influential.

**Implications for Practical Applications**

Reading encompasses decoding, word recognition, comprehension and vocabulary. Therefore, children need to be taught strategies for decoding, strategies for identifying words correctly or strategies which enable the correct pronunciation of words using the meaning and context of a text.

Accordingly, teaching should place an emphasis on a combination of phonics and vocabulary enrichment in addition to teaching word recognition skills. Generating phonological representations can be enhanced through phonics instruction. According to Venezky (1999), one main function of phonics instruction is to provide beginning readers with a method for generating approximate phonological representations of unknown words, which are close to a correct pronunciation using context, so the unknown word can be correctly identified. With this in mind, beginning readers should be encouraged to become active problem solvers with regard to visual information in the text using set for variability, the skill Venezky claims is essential for learning to read in English. In acquiring this skill, children learn to use their developing knowledge of spelling-to-sound relationships to produce approximate phonological representations, or partial decodings, for unknown words, especially those containing irregular, polyphonic, or orthographically complex spelling patterns.

The phonological representations provide the basis for producing alternative pronunciations of target words until one is generated that matches a word in the
child's lexical memory and makes sense in the context in which it is being read. Therefore, if the child produces a word that does not sound like something already stored in the listening vocabulary then the child has to change at least one of the sound associations and attempt the word again. The sound association which usually needs to be altered is a vowel (Venezky, 1999). Successful instructional programmes explicitly emphasise the interrelations between the orthographic, phonological, morphological, semantic, and syntactic aspects of reading. The basic principle is that the more one knows about a word (i.e., its phonemes, orthographic patterns, semantic meanings, syntactic uses, and morphological roots and affixes), the more efficiently the word is decoded, retrieved, and comprehended.

Vocabulary development involving both phonological and semantic growth is relevant to word reading and reading comprehension processes as reported in these findings. As children develop better vocabularies, reading development is naturally fostered. Children with limited vocabularies or cognitive skills upon school entry should therefore, be given rich opportunities to strengthen these skills prior to formal reading instruction.

Vocabulary instruction can help both comprehension and word recognition and a broad vocabulary is necessary for skilled reading comprehension. Growth in word recognition is hampered when the printed words a child attempts to sound out are not in the child’s vocabulary (Juel, 2006). Vocabulary should be taught directly, particularly key words and terms, before reading a text and indirectly to enhance the reader’s ability to inference and process word meanings from a text. A mental model of the content of the text is more easily built when the key words are known. Other useful keywords within a text targeted for direct teaching are frequently encountered
words and words which are likely to be met by reader’s at higher levels (Medo & Ryder, 1993).

Opportunities to detect and use new words will enhance vocabulary learning, e.g., during dialogues with the teacher (Coyne, McCoach, & Kapp, 2007). The incidental learning of vocabulary can be improved by teaching children how to derive meanings from context. This includes how to search the context for clues about the unknown word’s category (“what sort of thing is it?”), for defining characteristics (“how can you describe it?”), and for likes and opposites (“do you know of something similar or the opposite?”). Repeating new words is also supportive to learning (Stahl & Fairbanks, 1986).

Helping children to learn the meanings of specific words and to become better at figuring out meanings of new words through independent reading are both important educational goals. Oakhill et al. (2014) suggests once children can read independently, they learn new vocabulary through reading and not through direct teaching of meanings. Therefore, the amount of reading undertaken by children in the early school years is a crucial determinant of their vocabulary development and educators should provide lots of opportunities to read. The evidence suggests that readers engaged in high volumes of reading, have higher levels of reading performance. A causal relationship between high volumes of reading and reading performance is due to skilled readers having larger vocabularies (Martin-Chang & Gould, 2008). Every opportunity to read adds to the reading proficiencies and skills possessed by a reader including the self-teaching strategy of phonological recoding, a potentially powerful factor in skilled reading. The reciprocal relationship which exists between vocabulary knowledge and reading comprehension where vocabulary
supports reading comprehension, and reading (with good comprehension) supports vocabulary growth, occurs across development (Beck & McKeown, 1983).

In terms of practical teaching, the Simple View of Reading acknowledges the different foundations of word recognition processes and comprehension processes and, therefore, both language comprehension abilities should be developed alongside the development of word recognition skills.

**Conclusion**

The aim of this study was to investigate the influence of vocabulary in the development of word recognition skills and reading comprehension with children starting school. The evidence from this study adds to the limited number of studies surrounding early literacy in New Zealand Schools.

In support of the first hypothesis, the receptive vocabulary knowledge of a child, when commencing school, was found to directly contribute to the child’s word recognition skills in the middle of the second year at school. Furthermore, this study found vocabulary indirectly contributes to the development of word recognition skills through set for variability. That is, vocabulary contributes to the determination of the correct pronunciations of spoken English words (Venezky, 1999). The findings indicated that set for variability had a stronger influence on the development of word recognition skills than vocabulary. Effectively, vocabulary in isolation is unable to support word recognition. The beginning reader needs to be able to make a partial attempt of a word by sounding the word out and make use of the vocabulary stored in
the lexicon to make a successful attempt. It is for this reason that set for variability explains more of the variance when added to the model.

In support of the second hypothesis, the findings from this study indicated the vocabulary knowledge of children, when they commenced school, predicted reading comprehension ability at the middle of Year 2. Set for variability does not directly contribute to reading comprehension when word recognition is added to the model suggesting that set for variability mediates the relationship between vocabulary and word recognition but not for overall reading comprehension.

While both word reading and vocabulary were found to have a direct influence on reading comprehension, (but vocabulary is not mediated by set for variability), word recognition explains more variance in reading comprehension than vocabulary at this point in children’s’ reading development which is the middle of Year 2. However, vocabulary explains more shared variance in reading comprehension than it did for word reading. The available evidence suggests the influence which vocabulary has on reading comprehension arises from ability of the emergent reader to activate the detailed knowledge of a word’s meaning and use the information to support comprehension. For this to occur, the meanings of words must have been previously heard and understood to comprehend what is being read. The representations stored in the lexicon assist in the process of assigning a meaning to the word being read using the context of the sentence or paragraph in which the word is located.

In summary, the results of this study demonstrate how vocabulary knowledge influences reading development in emergent readers and therefore aligns with the Simple View of Reading. Vocabulary knowledge is related to word recognition in that
it reflects the reader’s knowledge about familiar words and how they are formed. A broad vocabulary also supports the processing of unfamiliar words by using strategies involving similarities of words stored in the lexicon. This study also found vocabulary is related to reading comprehension because it reflects an individual language competence which, in turn, influences reading development.

Acquiring reading comprehension is the fundamental product of the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990). The present research demonstrated the importance of vocabulary to both reading comprehension and word recognition which suggests the Simple View of Reading can account for the relations between decoding and language comprehension across development. Both word recognition and language comprehension are essential and necessary at all levels of reading because both components accommodate the nature and operation of cognitive and linguistic processes. This study could have implications for instructional practices in New Zealand primary schools. Knowing about these key components and how reading development occurs is important in terms of providing timely support. Children identified as having limited vocabulary knowledge and/or experience difficulty with word recognition and reading comprehension at the emergent literacy learning stage can allow specific teaching strategies to be recognised and implemented. Such approaches may prevent the current achievement gap from widening and go some way towards improving the early literacy levels of children in Aotearoa, New Zealand.
References


Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading


