Improving Reading Comprehension via Instruction in Morpheme Meanings with EAL and L1E Primary Level Students in the United Kingdom

Kathryn L. Ownby

Doctor of Philosophy

University of York
Education

October 2021

Abstract

This thesis explored the connection between morphology instruction and reading comprehension with primary aged children. It focused on an intervention that was designed to teach the meaning of common morphemes with the aim of encouraging students to utilize the meaning carrying aspect of morphology as a method to understand vocabulary. The main questions were: Is there a difference between EAL and L1 children in the effectiveness of teaching the semantic meaning of derivational morphemes? Is there an effect of individual differences in gender, attendance and age in the effectiveness of teaching the meaning of morphemes? The second question is What is the effect of measures of reading ability on derivational morphology knowledge? These questions were redesigned to better reflect the measures utilized in the study.

The study employed a double baseline design with participants being their own control group. The participants (n=22, EAL= 11) were recruited from schools in Yorkshire, UK. They were tested on three occasions with ten weeks of instruction between tests. This provided a baseline measure of growth which was then compared to the growth as measured in the same participants after intervention instruction. The measures employed were both standardized measures and researcher constructed measures.

The results indicated that there was a significant improvement in morpheme knowledge and comprehension, particularly with the EAL participants. There were inconclusive results for the L1E participants on the morphology measures. However, for this latter group, there was significant improvement on reading comprehension. The mixed models showed that there were no significant impacts from additional factors except for attendance. There were no significant differences between EAL and L1E participants on overarching components of reading that were measured. While there was growth it was not significant over the intervention period alone, while it was significant over the total period of the study.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This thesis would not have been possible without support from many people and institutions. I would like to thank my three supervisors who have helped me shape and form this work over my time at York. First, Claudine Bowyer-Crane for her support at the beginning of this whole process with shaping the research plan and recruitment, along with making supervisions a safe place to debate issues and questions. Secondly, I would like to thank Danijelja Trenkic for her help with the data analysis. Thirdly, I need to thank Leah Roberts for her help with the writing up phase and support with the struggles of getting words on the page. Additionally, I need to thank Cylcia Bolibaugh for her help and advice in TAP meetings.

Next, I would like to thank the schools and students who welcomed me and my ideas and participated in this research. They were always curious and flexible when it came to allowing me to work with them. None of this would have been possible without their support and participation. I also would be remiss, if I didn't thank my former colleagues and students who over the years have been the inspiration behind this work.

None of this would have been possible without the support of my family. My parents have generously supported this quest and have been there through all of the ups and downs of the process. My mother who was more than willing to listen to me excited when things were going right and encouraged me when things went wrong. My father, who was willing to proofread sections when I had spent so much time on them that I no longer could see where there might have been issues. My older brother, David who was willing to debate issues of higher education and my Uncle Ray, who seemed to know just when to send an email encouraging me to just get it done. I appreciate all of your unwavering support and belief that I could finish this.

I additionally need to thank my fellow PhD students as we supported each other from the Berrick Saul days to Derwent, to Alcuin and back to Derwent. Your help with piloting measures, debating issues, along with advice about surviving a PhD, I am incredibly grateful for. An extra round of thanks for the PhD Peeps chat group that kept me sane and somewhat productive throughout the pandemic. I am eternally grateful for all the rides, help with moving, meals, and sounding boards that you were for me. No one can complete this marathon on their own and my village has been an amazing support system.

Author's Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1. Introduction

There is a growing body of research evidence that morphological knowledge (conscious awareness of the internal structure of words) plays a vital role in reading comprehension and that this may begin in the primary years and increases in importance through to higher education. This has been supported by a number of studies within the morphological systems framework (Goodwin et al., 2017; Goodwin, Gilbert, et al., 2013a; Levesque et al., 2020; Taft, 2004) and leads to questions about how morphological knowledge can be operationalized and utilized in a classroom: a question that has yet to be addressed, in particular with regards to how this knowledge by should employed by educators to help students who might be struggling with reading.

Specifically, while governments and curricula (Department for Education, 2015; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) have stated that morphemes should be taught for varying purposes, the question of how this should be carried out and at what level might it start to make a noticeable impact has also not yet been fully addressed. Part of this gap is what this current research project is focuses on. Its goals are two-fold. Firstly, to explore if instruction in derivational morphemes can affect reading comprehension for upper primary students. Secondly, to see if a potential intervention is successful at improving knowledge of morpheme meanings with participants. Another aim of the current study is to examine if there are differences in the impact made by the instruction for English as an Additional Language (EAL) students or first language English (L1E) students as they potentially could have different needs in vocabulary acquisition and reading comprehension. The intervention may support teachers who are looking for "best practices" and "research based" practices to use in their teaching. Many educational systems have pushed over recent decades for teachers to base their work in methods that are supported by research evidence (e.g., National Governors Association Center for Best Practices and Council of Chief State School Officers, 2009). To support this a selection of methods that have been demonstrated as successful and others that are commonly used in classrooms were combined to provide a cohesive program.

Work that has been conducted into vocabulary instruction has been split over two viewpoints or 'camps', both finding substantial support. The first is the' shallow but wide' camp which posits that the most useful method of instruction for students to acquire new vocabulary would be for them to learn three to six new

roots words daily resulting in acquisition of over 230 new words annually (Biemiller & Boote, 2006) The second viewpoint is the 'deep but narrow' camp, which supports instruction in fewer words but with a more complete picture of each word being instructed (i.e., meaning, pronunciation, and being able to fully utilize the new word across different contexts), and is conducted through multiple exposure and repeated usage of the target word (e.g., Beck et al., 2013) Since there is strong evidence that supports both approaches, this study adopted parts of each for its intervention. It applies the concept of multiple exposures and usage along with instruction in approximately four target morphemes weekly, from the ideas put forth by Biemiller and Boote (2006). The use of four morphemes for instruction each week was a modification from the original plan to use six as this would mean instruction would include two morphemes in each session. However, when the number of sessions each week was reduced to two this resulted in a reduction in morphemes that was still within the guidance given. The use of multiple exposures and more "depth" (in this case, meaning, usage, recognition in contexts) was taken from the work of Beck et. al. (2013). Both of these researchers' work will be explored in depth in the Literature Review. It should be noted that Beck has tended to focus on primary grades reading and Biemiller and Boote have tended to utilize slightly older populations that depending on the country where the work is being conducted could be considered middle grades or secondary school. Thus, slightly more was drawn from Beck's work than Biemiller and Boote's. This is a different method to what has been applied previously but combines key parts of the established work in order to use methods familiar with teachers and supported by research in vocabulary but applying it to the concept of derivational morphemes (morphemes that change the part of speech or the meaning of a word, e.g., happy-*ness*).

The instruction used for the intervention is a cycle of four lessons, including a review lesson, over a two-week period. This cycle is repeated five times for an intervention lasting ten weeks total. Each week includes morphemes whose meanings are grouped by topic. These morphemes are introduced through reading passages that are targeted at a level that should be accessible to students beginning at a year three reading level. The additional activities include word maps, where students have an opportunity to morphologically decompose morphologically complex words and then use the morpheme knowledge, they possess to define the original word after understanding what meaning each component morpheme carries.

This activity, combined with word sorts, are intended to help students develop skills in recognizing morphemes and being able to group words by similar component morphemes or similar meanings. Finally, as part of the review sessions, students select a word that contains one of the target morphemes covered in the previous three sessions to use in a sentence. This is to aide them in applying the knowledge they have acquired to a context they are familiar with and to practice using the new words. As a whole, this cycle of instruction supports the current work into vocabulary acquisition while extending it to morphemes. The development of skills in both identifying morphemes in morphologically complex words and being able to recall their meaning theoretically should support an increase in reading comprehension abilities. A detailed description of this intervention and its design is set out in Section 3.2

The methodology used in this study is different than the traditional intervention and control group with pre and post-test measures. As recruitment of participants was difficult, an experimental method was tried using the participants as their own control group. This is accomplished by the use of a double baseline design and an intervention consisting of twenty sessions over ten weeks. The intervention is evaluated using a pre-test/ post-test design where participants are tested at a ten-week interval before the intervention and then post intervention. A comparison of the change over both periods is then made. The intervention is designed using activities that are commonly used in classroom instruction but have been modified to focus on morphology. Word maps, word sorts, short passage readings and sentence writing are all commonly used in vocabulary instruction with success. The modification of these activities to focus on morpheme knowledge is a newer direction. The measures employed in this work are a blend of previously published and standardized measures and a new measure of root morpheme meanings. They are all designed for use in the upper primary grades or years. The measures included the York Assessment of Reading Comprehension (YARC), the British Picture Vocabulary Scale v3 (BPVS), and the meaning section of the Word Part levels test (WPLT). The new measure of root morpheme meanings was designed by the researcher specifically for this study.

Part of the initial inspiration for this work comes from the researcher's personal experience in teaching EAL students in multiple contexts over her 14-year teaching career. The contexts range from blended classrooms in English dominant countries to EAL only classrooms in countries where English is not spoken widely but is part of the government curriculum. She noticed a growing pattern of student

performance on reading and vocabulary assessments when students were instructed in morphemes and how they carried meaning in vocabulary. This led to an exploration as to why teachers were or were not teaching it as part of her Masters in TESOL dissertation (Ownby, 2017). The primary reason that was given was an unfamiliarity with morphemes and their meanings by both experienced and new teachers. This finding is not just supported by the researchers own work as other studies have found that teachers are lacking knowledge about morphemes. One such study found that when 383 students who were training to be teachers were asked to define the terms word, prefix, suffix and stem, less than half were able to(in order respectively, 47%, 42%, 40%, 8%(Sangster et al., 2013). This led to thoughts of how teachers could be assisted to encourage the teaching of morphemes and their meanings along with questions as to whether this would potentially be beneficial to the students. It has been shown that when teachers learn about morphological awareness, it affects their instructional practices to increase usage of morphological concepts(Newton, 2018). This personal interest along with the ever-increasing number of EAL students in primary classrooms, encouraged this work to be completed. The following paragraphs will detail some of the current and recent statistics about EAL students in English dominant countries to demonstrate the need for teaching methods that are suitable for both EAL and L1E students and can be implemented easily for teachers.

According to the recently released statistics for the 2018-2019 school year there are approximately 4.7 million students attending mainstream primary level schools in England. Of these 21.2% do not speak English as a first language and 23 % have been eligible for free school meals at any time in the previous six years. This means that there are about 1 million students who are attending school for whom English is not their first language and about 1.08 million who were eligible for free school meals. The data from the government does not give an indication of how much overlap there is, however a slightly older report does provide some interesting analysis of English as an Additional Language (EAL) and achievement data. It states that the population of EAL students in primary and secondary schools in England doubled from 7.6% in 1997 to 16.2% in 2013 with just over a million pupils being classified as EAL(Strand et al., 2015). This is comparable to a similar increase in English Learners (EL) in the United States over a similar period. In school year 2000-2001 there were 8.1% of all students or approximately 3.8 million ELs enrolled in public schools from kindergarten to grade 12. By school year 2016-17 it was 9.6% of the total enrolment or 4.9 million students in public schools this number includes all fifty states and the

District of Colombia. Public schools in the United States are funded by the federal and state governments. In a recent report form the National Center for Education statistics in the fall of 2018 10.2% or 5.0 million students in the United States were ELLs though there was a significant drop by the fall of 2019 to 9.2% or 4.5 million. If this is a new trend or the result of political issues that were present in the United States under the government that was in place from 2016-2020 is yet to be determined. What can be drawn from this data is that there has been and is likely to continue to be a significant increase in EAL or ELL students in classrooms in both the United Kingdom and the United States. This increasing percentage reflects an increasing need for teaching pedagogy that works for both EAL and L1E students.

Using 2013 assessment results, a report was commissioned by three charitable groups that focus on EAL students (Strand et al., 2015), the aim of the report was to identify the most at-risk groups of EAL learners and identify programs or interventions that could show potential based on evidence. What has been identified in this report in England is that EAL pupils tend to have lower odds at achieving a good level of development at the end of reception than their L1E peers. In this report it was calculated that 44% of EAL pupils reached this mark in comparison to 54% of L1E pupils. This would mean that the odds for an EAL pupil to achieve a good level of development are 33% lower than they are for L1E pupils. Interestingly the mathematics scores are always higher than for reading at every age for EAL pupils with the gap between EAL and L1E is almost eliminated by age 11 on mathematics assessments. EAL pupils who are entitled to free or subsidised meals tend to be three months behind EAL pupils who are not entitled. This is surprising when compared with L1E students who are five months behind L1E students who are not entitled to free or subsidised meals. This is important to note as it could be connected to the work on vocabulary gaps due to exposure to spoken language from a young age rather than home language. It has also been noted that if EAL students arrive during Key Stage 2 (between 7-11 years of age) they tend to score 12 months behind EAL students who were present during Key Stage 1. This is relevant to this study as all of the participants attended school in England for all of Key Stage 1 so there was not the gap that seems to be present with students who arrive post Key Stage 1 on reading assessments. This was planned as part of the criteria for participants to avoid this issue and in an attempt to have participants who had similar educational experiences, even if they had differing language backgrounds.

Before introducing the structure of this thesis there are numerous terms that should be clarified as there are several with overlapping meanings used in morphology and language instruction. The terms were chosen to be respectful to participants and in an attempt to be concise as to what is being described in this present work. Where possible definitions commonly used in published research have been selected. The next section will cover primarily terms being used for participants and morphological constructs.

1.1. Terms used in this thesis

The choice of terms employed in this thesis was made with a concerned eye toward being as respectful of participants' experience and cognisant of the use of terms in literature as possible. The decision to use English as an Additional Language (EAL) instead of English as a Foreign Language (EFL) or English as a Second Language (ESL) was made to recognize that the participants might speak more than two languages and the order of acquisition of said languages could be varied. Also, English is not a foreign language in the context where this work took place. English is the common and dominant language of the United Kingdom. EAL is currently the term used most often in the UK when discussing students who speak more than one language. EAL is used by the Office of Statistics when tracking this population as well and it also was the classification the participating schools utilized for the participants categorized as EAL. This lends itself to being a rational and easily accessed term. The emergence of the term Language Minority Youth (LMY) in the United States is interesting as it is defined as "individuals from homes where a language other than a national language is actively used, who therefore have had an opportunity to develop some level of proficiency in a language other than a national language" (Shanahan & August, 2008, p. 2) This is interesting as the United States has no national official language. Thus, the use of EAL avoids these quandaries and potential marks of nationalism. While additionally respecting the linguistic diversity of the population being studied. Many of the terms discussed carry value judgements and imply a level of value for the linguistic knowledge populations might possess. In cases where previous work is being discussed the terminology of the study being discussed will be used as there could be slight differences in definition.

The next term and acronym that is widely used in this thesis is First Language English speaker (L1E), these are participants who have English as both their home language and language of instruction at

school. This abbreviation is a modification of the more widely used L1 and L2 that are commonly used to denote first and second language. The decision to use this abbreviation instead of the more widely used Native Speaker (NS), English Speaker (ES) or, L1 is to try to value all participants linguistic backgrounds equally (for an overview of this debate see: (Dewaele, 2018)). The connotations that are carried by the NS term of it being valued above other linguistic backgrounds particularly in language teaching contexts should not apply in this work, as all of the participants, L1E and EAL, have had similar educational backgrounds in English schools from Reception to year 3. This should mean that all participants have had similar exposure to the English language in educational settings. The difference being the language use outside of the school context. There has been a great deal of debate about the label of 'native speaker', including how to define it when there are bilinguals who are exposed to two languages from birth. To try and avoid any confusion or potential for bias from the terms being used, the concept of L1 as first language is being used with the additional descriptor of E for English. This is simply an attempt to identify a subgroup clearly while positioning it as close to equal as possible.

There are several terms concerning morphology which are used in this thesis that should be clearly identified as to how they are being used. This section will give a short overview of the most important and commonly used in this thesis. There will be an additional and more in-depth discussion in Section 2.3 at the start of the Morphology literature review.

First is *morpheme,* meaning the smallest meaning carrying unit of a word that cannot be broken down further. It may be bound or unbound. *Bound* meaning that it must be combined with an additional morpheme to make a complete word (e.g., un-, unless). *Unbound* morphemes can stand alone as a word (e.g., book). There are additionally two types of morpheme: *inflectional* and *derivational*. An inflectional morpheme is an affixation that specifies features such as case, tense, gender, or number but does not affect the syntactic category of the word (work-*ed;* girl-*s*; eat-*en*). Derivational affixation can either or both change the syntactic category (happy/happy-*ness*) or add a new meaning to a base (*un*happy) (S. H. Deacon et al., 2010) This thesis is mainly concerned with the meaning part of derivational morphology. It looks to examine the ability of students to learn meanings of morphemes and thus potentially understand additional words containing these morphemes.

Morphological knowledge comprises three parts: relational, syntactic and distributional. These three parts relate to the sharing of common morphemes, knowledge of grammatical usage, and the rules about how they can be combined to form new words. Morphological knowledge can be further divided into *awareness* and *analysis*.

Morphological awareness is "the ability to reflect upon and manipulate morphemes and employ word formation rules in one's language" (Kuo & Anderson, 2006, p. 161)(e.g. book is a noun, and we know it can be combined with worm to make bookworm). Morphological awareness is knowing how to deconstruct and combine morphemes to form words.

Next is *morphological analysis*: "to break an unfamiliar word into parts and then recombine the parts into a meaningful whole"(White, Power, et al., 1989) This meaning is examined in more depth at the beginning of the literature review on morphology. It is similar to the term being used frequently in this thesis, morphological decomposition, which is the breaking apart of complex words into morphemes and then recognizing known morphemes to aid in the identification of the word as a whole. Having now briefly defined some of the key terms the next section presents an outline of the thesis with descriptions of what each sections contains.

1.2. Structure of this Thesis

The second chapter of this thesis will be the literature review that will begin to examine what it means to read, then progress to two theories of reading development that explore how the skill of reading develops and its connection to morphemes. Then we will progress to looking at current and past theoretical frameworks of reading that illustrate potential influences of morphological knowledge on reading comprehension. These will be followed by a summary of work that has been conducted into reading comprehension with EAL students of varying ages and L1 English students. These sections together should give a clear overview of the current status of the field of reading comprehension.

The next area of the literature review is the narrower focus of vocabulary, including vocabulary knowledge with respect to both size and depth of vocabulary and how they can be measured and the choice of this work for focus on size rather than depth for measures. Then we will progress to the differences between receptive and productive vocabulary. Additional consideration is given to academic

vocabulary and its importance for students and the substantial component of academic vocabulary that is related to morphological knowledge. This is followed by an exploration of the idea of a 'vocabulary gap' between both high and low SES students and thus potentially EAL and L1E students and the effects of this gap on learning and how it theoretically expands over time due to the Matthew effect.(Stanovich, 1986b) Following on from this is an examination of current methods of vocabulary instruction and a discussion about the advantages and disadvantages of both the 'shallow but wide' and the 'deep but narrow' camps. The last part of the vocabulary section concerns what governmental guidance has been given and how it is incorporated in curricula, as one of the goals of the intervention included in this work is to support the curricula goals and student learning.

The final main section of the literature review examines specifically morphology and current models of morphological processing. This then leads to a survey of the connections between morphology and reading and prior studies that have looked at morphological interventions. It concludes with a brief study of work on Greek and Latin roots.

The third chapter introduces the research questions and the methodology selected for answering them. There is also detailed description of the intervention sessions and what the scope of the intervention is. It then follows with an explanation of the measures that were selected and designed specifically for this work. The measures selected are the York Assessment of Reading Comprehension (YARC), the British Picture Vocabulary Scale v3 (BPVS) and the meaning section of the Word Part levels test (WPLT) together with a measure of root morpheme meaning that was designed by the researcher specifically for this study.

Chapter four focuses on the successful pilot study that was conducted in the spring of 2019 to determine the feasibility of the study and measures in the selected context of a primary school in Yorkshire, England. It describes in detail the participants, materials, procedure, analysis of results and reflections and modifications that were needed to conduct the main study over the 2019-2020 academic year. Which leads to Chapter five covering similar topics addressing the main study which was carried out from September of 2019 to March of 2020 when it was cut short by the COVID-19 pandemic and the subsequent closure of schools in England for the remainder of the school year.

Chapter six is a presentation of the data analysis it starts with an examination of the data collected from all participants collectively and then as subgroups consisting of EAL and L1E participants. The data for each measure is presented and analysed individually. The Receptive vocabulary measure (BPVS3) is analysed first followed by the subtests of the York Assessment of Reading Comprehension (YARC), these include a single word reading test (SWRT), Reading rate (YARC RR), Reading Accuracy measure (YARC A), and Reading Comprehension (YARC RC). This is then followed by the two morphology measures the modified WPLT or Morphology measure for prefixes and suffixes and the morphology measure roots. These two were additionally analysed as a whole group for differences between morphemes that were taught in the intervention and morphemes that were not targets of instruction in the intervention. The second main section of data analysis where the results of the EAL and L1E participants are examined follows this same pattern excluding the differences between taught and untaught morphemes as there were not enough participants to draw reasonable conclusions from the subgroups. The third and final section of the data analysis is a short examination of interviews conducted with the teachers of participants post intervention. These were conducted to examine if the intervention had any effects on the participants that were beyond the scope of the measures, such as increased confidence or application of the skills learned in classroom discussions. These were planned to be combined with focus group interviews with participants that were lost due to school closures. They are being utilized to support the quantitative data that was gathered.

Chapter seven is an in-depth discussion of the findings with focuses on the results as a whole and with regards to each subsection and subgroup. It also addresses the answers to the research questions considering what was found. It additionally includes an evaluation of the intervention specifically and its future applicability to additional contexts. Which leads to chapter eight in which the final conclusions of this study are presented, and chapter nine which address applications of the findings from this work for instructional settings and how the materials and activities designed could be applied to current teaching. The final chapters discuss the limitations of these findings in detail as there are several that were beyond the control of the researcher and need to be considered thoughtfully and addressed in any future work. Directions for future work and modifications from the work discussed here are in chapter eleven. Following these are the references and appendixes that support the work contained in this thesis.

The findings of this work indicate that the method employed in the intervention has a strong potential to be a feasible and successful method for instructing upper primary aged students in morphemes. There are also some possible additional positive effects of the intervention upon the participants' reading comprehension skills. This was supported by both the quantitative measures that were the focus and the qualitative teacher interviews that occurred post intervention.

2. Review of Literature Concerning Reading Comprehension, Vocabulary and Morphology

The current study focuses on an intervention designed to teach morpheme meanings with the aim to improve reading comprehension amongst primary aged students with both EAL and English only linguistic backgrounds. This literature review presents an overview of the current theories and research in reading comprehension, vocabulary instruction, and morphology relevant to primary level education. This review focuses on educational settings where English is the primary language of instruction. This background was necessary to develop both an educational intervention and the methodology to examine the intervention's effectiveness. The original goal of the interventional study was to evaluate a method for improving reading comprehension of EAL students by instructing them in morphology. The evaluation was also designed to evaluate the effect of the intervention on the reading comprehension of L1E students. The combined evaluation of the intervention on EAL and L1E students was thought to be applicable to current UK and US educational settings where EAL students are often integrated to a L1E environment as a result of immigration.

Morphology, or the study of morphemes, the smallest meaning carrying unit in a word, has emerged in reading research as an important component of reading comprehension, and its importance is also reinforced by the recent inclusion of the teaching of morphology in primary level curricula (Levesque et al., 2017; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Despite the inclusion in curricula, primary classroom instruction in morphology has been hampered by a lack of teacher knowledge and educational pedagogy (Hammond, 2015; Hedrick et al., 2004; Newton, 2018; Ownby, 2017; Pittman et al., 2022; Sangster et al., 2013). If the proposed intervention were successful, it would provide both additional evidence for the impact that the teaching of morphology can have on reading success and initial guidance for developing instructional methods.

Morphology breaks down into two main areas, inflectional and derivational morphology. Inflectional morphology is where affixation is utilized to change the grammatical structure of a word (e.g., walk, walk-*s*, walk-*ing*). Derivational morphology relates to the meanings carried by morphemes and how they can be combined to change the meaning or part of speech of a word (e.g., use, *mis*-use, *un*-used; happy vs.

happy-*ness*) (Kuo & Anderson, 2006). As this study is focused on participants learning the meanings of morphemes, it is mainly concerned with derivational morphology.

The review starts by examining the concept of reading comprehension and then narrows the focus to vocabulary and then finally morphology. Reading comprehension will be examined looking at theories and important frameworks of reading comprehension and reading development. Then we will move on to explore vocabulary and how it is currently being taught. Finally, we will consider morphology, paying specific attention to the research that has been conducted into how morphology connects to both reading comprehension, reading development and vocabulary knowledge.

2.1. Review of Reading Comprehension Research for Primary Level Students

As background to the current study, this section of the literature review will examine the concepts and theories of reading comprehension that underlie the design of the intervention and the ideas both theoretical and practical behind it. The review will begin with investigating how others have conceptualized reading comprehension while recognizing that reading comprehension involves so many components that no single theory can adequately cover the entire process of reading. This review will examine some major theories of reading such as the Simple View of Reading (Hoover & Gough, 1990), Reading Systems framework (Perfetti & Stafura, 2014) , the Lexical Quality Hypothesis (Perfetti, 2007), and the recent Morphological Pathways Framework (Levesque et al., 2020). It will then look at skills needed to become a successful reader including phonological and morphological awareness, orthographic knowledge and lexical processing, often measured via naming speed. It will conclude with a look at recent research into reading comprehension directed toward English as an Additional Language (EAL) learners.

2.1.1. What does it mean to read?

Reading comprehension is defined as "a process through which a reader constructs a mental schema, or representation, by integrating the information presented in a text with her own prior world knowledge—of content and language," according to Galloway and Lesaux (2015, p.158). If we examine this definition, it

includes mental images, text, and prior knowledge of both content and language. A student's mental schema is key to reading comprehension. In psychology and education theory a mental schema is a mental process or behaviour allowing the organization of concepts (words) and the interrelationships between concepts. The mental processing of these webs of organized concepts and their interrelationships finally crystalizes into and understanding of what the words are intended to convey. One of the purposes of the intervention upon which this study was designed was to expose students to knowledge and content in the form of morphology, with the intention of nurturing their mental schema for when they encounter morphemes in the future.

2.1.2. A Tale of Two Theories of Reading Development

This section will examine two theories of reading development and how they relate to the choice of participant age in the study. It is important as there are underlying skills that would need to be developed prior to a child being able to recognize morphemes and decompose them from morphologically complex words. The two theories selected for examination are first Ehri and McCormick's phases of word learning (Ehri & McCormick, 1998) and secondly, Seymour and Duncan's dual foundation model (Seymour & Duncan, 2001). These have been selected as they both stress morphology as an important component of reading specifically. Additionally, Ehri and McCormick's model has been widely tested, while Seymour and Duncan's is newer but appears to build on what has been established and refined.

According to Ehri and McCormick's phases of word learning (Ehri & McCormick, 1998) there are five phases to learning to read. These phases may overlap and thus a child could be already developing skills for the next phase as they are completing the previous phase. The five phases are pre-alphabetic, partial alphabetic, full- alphabetic, consolidated alphabetic and finally the automatic alphabetic phase. What follows will be a short overview of each phase and where the participants in this study are likely to fall.

The first phase is the pre-alphabetic phase where alphabetic knowledge is not used to read words. This is where children might use other visual cues to read a word like a commonly occurring graphic on a sign (i.e., the golden arches for McDonalds). This phase typically is present in pre-school and kindergarten aged children.

The second phase is the partial alphabetic phase where children are weak at decoding words, but they can use guessing strategies to read words. This stage typically is present in kindergarten and year one students. Students with good sight word memory are more effective readers at this phase.

The third phase is the full-alphabetic phase. This is where children acquire and are able to use sound-letter correspondences to read words. This level needs to be mastered for students to be able to effectively be successful in the final two stages. This is where students have a working knowledge of grapheme-phoneme correspondences and can utilize them for conventional spellings of words.

The fourth phase is the consolidated alphabetic phase. This phase usually begins to appear in year 2 of school. This phase tends to overlap with the full alphabetic phase. This is where students begin to store larger chunks of letters that occur frequently in words, including affixes, roots and syllables (Henry, 1989). Here, children will begin to be less dependent on grapheme- phoneme connections and use more multi-letter combinations.

The fifth and final phase is the automatic phase. This was deemed as such as children have a highly developed automaticity and speed in identifying words both familiar and unfamiliar. This is also where, if a child has a difficulty when encountering a word, they have multiple strategies that they are able to call upon to identify it.

The final two phases are likely to be where the participants of this study are located. They have started to store larger chunks and while they may not be fully automatic at reading words, they are making progress in that direction. This would correspond with students in years 3 and 4 of the English school system. Which was the target years for recruitment. Children in the earlier phases would not be suitable for this study as they would not be able to successfully employ morphemes. The lack of skill and ability developed to recognize groupings of letters would make it very difficult for them to be successful. Whereas students who have reached these phases where they can and do store larger chunks and morphemes would be able to use the information given and build upon the skills they have already developed.

While Ehri and McCormick's phases are applicable, an even more applicable model is arguably the one by Seymour and Duncan (Seymour & Duncan, 2001). Their model includes morphology explicitly. It also contains 4 developmental periods versus the five in Ehri and McCormick's. The periods are pre-literacy, dual foundation: logographic and alphabetic, orthographic and, morphographic. These will be discussed in the following paragraphs in more detail.

Pre-literacy refers to children who have not yet begun to read. They might recognize a few letters and they have a beginning grasp of linguistic awareness, for instance, children can acknowledge that words can be broken into smaller units be they syllabic, phonemic, or morphemes. They do not need to recognize all the different units but so long as they can recognise that they exist, it is claimed that this will kick-start the use of the linguistic awareness system in reading. This most likely occurs with children who have not started school and are approximately 3-4 years of age. During this phase knowledge of sounds letter correspondence develops and interacts with linguistic awareness.

The next phase is Dual Foundation. This comprises two separate areas of development: the Logographic and the Alphabetic that function together. The logographic foundation is "a process whereby lexical representations are established in memory at the onset of learning to read." (Seymour & Duncan, 2001, p.291) This would be similar to the term 'sight words' used by other researchers. Alphabetic foundation is "a simple decoding procedure by which individual letters are converted to sounds and the sounds are synthesised to for a pronunciation." (Seymour & Duncan, 2001, p.292). This dual route would also connect to a mixed method of teaching reading using a balance of whole language and phonics. While they posit that there could be potential delay in development of the logographic system if the focus is on phonemic instruction, they additionally state that the reverse could be true if there is a focus on whole language instruction and therefore, they advocate for a more balanced approach. This phase most likely occurs during the early years of schooling.

The Orthographic phase seems to parallel the consolidation phase from Ehri and McCormick's work. The focus in this phase is on the orthographic framework where the spelling structure and is organized logically. It involves a two-way connection with metalinguistic awareness and connects to the

logographic foundation and alphabetic processes. Evidence indicates that this is occurring around the age of 7 years.

The final phase is the Morphographic literacy. This is where combined syllabic units or morphemes are recognized and utilized in reading. This is built on previously developed epi-awareness of morphology. This is the stage where students can recognize morphemes and have begun to connect them to meanings and possible ways of constructing words. This would likely occur at the end of Year 2 of schooling and the beginning of year 3. Using participants at the midpoint of year 3 will mean that this knowledge should be in place and accessible but not necessarily automatic.

Both theories are based on decades of research (e.g., Ehri & McCormick, 1998; Ehri & Wilce, 1987; Seymour & Duncan, 2001; Uhry & Ehri, 1999) . Although similar in a number of ways, Ehri and McCormick's work has been widely accepted and employed in research, while Seyour and Dunncan's is not as well cited. By using both theories, the selection of mid to late primary grade participants is logical, as they should have developed their reading skills to an acceptable point where they can recognize morphemes and decompose them as well. This connects to work by Berninger, Abbott, Nagy, & Carlisle (2010) which is discussed in more detail in the sections on morphology. They state the morphological awareness begins to develop in the early primary grades at a rapid pace then continues less rapidly in the mid to late primary years.

Having discussed two theories of reading development and how similar they are. We will now turn to theories of reading comprehension. We will start with the simple view of reading and then move to the reading systems framework and finally the morphological pathways framework.

2.1.3. The Basis for Reading Theory: The Simple View of Reading

The Simple View of Reading developed by Hoover and Gough (1990), proposes that reading comprehension is the product of decoding and linguistic comprehension. Decoding involves word recognition from print input and translation or matching to entries in an individual's mental lexicon (the mental system which contains all the information a person knows about words). The difficulty for beginning readers is that they must connect the phonological representation or sound representation in the mental lexicon to the printed form, where the printed form has never been seen before. Linguistic

comprehension refers to the ability to take in semantic information (in reading comprehension, taken in visually via words on the page), and then interpret it at the sentence or word level.

According to the above theory, improvement in either decoding or linguistic comprehension has been theorised to improve reading comprehension as a whole. If this theory holds, a beginning understanding of morphology could improve reading comprehension for students by positively influencing both decoding and general comprehension skills. It would have an impact on decoding as it could reduce the effort required to read as morphemes are stable units which do not change how they are formed or spelt. It could potentially reduce the effort required to comprehend as a reader would not need to think about if the word was a change in form but carried the same meaning. This theory is the groundwork for many other theories. While the simple view of reading has been argued to lack depth (e.g., Lesaux et al., 2010; Levesque, Kieffer, & Deacon, 2017), it is a starting point for many other concepts. As this review progresses it will illustrate how more detail has been added, thus giving a clearer picture as to how reading is thought to occur as a process.

2.1.4. Evolution of Theoretical Frameworks of Reading Comprehension that Explicitly Include Morphology

There have been several frameworks developed to explain the process of reading comprehension. Perfetti and Stafura's (2014) the Reading Systems Framework, which was adapted from Perfetti's earlier Blueprint Theory (1999) is highly relevant to the current study. In this framework, the importance of word-level processes, including morphology, are highlighted. That is, in contrast to the Blueprint Theory, the lexicon in the Reading Systems Framework is situated centrally between the comprehension system and the word-identification system and is seen as a key factor in how words and their meanings are stored in and retrieved from the mental lexicon as part of the reading comprehension process. This framework also involves the types of knowledge: linguistic and orthographic (writing system) and general/world knowledge. The main components of the reading comprehension processes in this theory are: visual input (visual recognition of characters as a word), word identification, and meaning and form selection. In this framework it is posited that visual input is parsed into both orthographic units and phonological units at the same time which leads to word identification and meaning form selection. Together these processes are being carried by the lexicon and comprehension processes. While this is going on additional

knowledge of the linguistic, writing systems and general knowledge support these processes as they feed to comprehension.

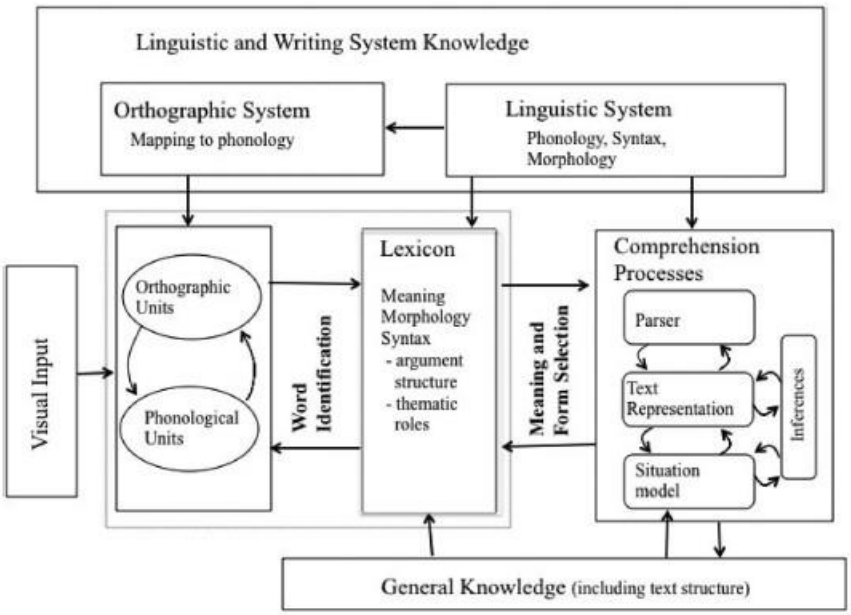

Figure 1: Reading Systems Framework (Perfetti &Stafura, 2014a, p.24)

Perfetti and Stafura (2014) state that their model allows for the hypotheses to be generated as regards the sources of possible reading comprehension difficulties. One particular key aspect for successful reading comprehension is a high-quality lexical representation (see also the 'lexical quality hypothesis', (Perfetti & Hart, 2002)). They define a high-quality lexical representation as having three components: orthography, phonology and semantics and if one of these areas is weak, the result will be a difficulty in comprehension by the reader. The importance of the quality of lexical representations has explored by others with supportive results that indicate that the quality of lexical representation that a reader possesses can impact reading comprehension. For instance, a study with 292 grade four (9-10 years old) Dutch students found that the measures of lexical quality significantly explained 30% of the variance in reading comprehension measures (Swart et al., 2017). A Canadian study of 272 students, with 178 EAL students tracked from grade 1 to grade 6 (O'Connor et al., 2019)found that there were two groups within the students, those that were good comprehenders and poor comprehenders. Both EAL and L1E students

could fit into either of these categories. Good comprehenders tended to have strong listening comprehension and semantic knowledge while poor comprehenders struggled with phonemic awareness. (e.g., “the ability to focus on and manipulate phonemes in spoken words.” (National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.)., 2000, pp. 2–1)) and orthographic knowledge. Comparisons of the mean scores showed that there was no significant difference in decoding skills between the EAL and the L1E participants. Notably this study’s EAL participants had been in Canada and schooled in English since grade one, so had undertaken approximately four years of instruction in English. What is important about this study is that it shows that students with what could be termed as ‘a quality lexical representation of words’ are stronger comprehenders of reading. Additionally, there were no differences in decoding skills between monolinguals and bilinguals. This suggests that if students are EAL, but they start schooling in an English environment, their EAL status may not necessarily have an impact on their decoding skills, and further, it suggests that the observed differences may be due to some other cause than the children’s L1. Given effects of quality of lexical representations on reading comprehension, it can be argued that quality lexical representations would be supported by knowledge of morphology as can potentially be built upon by both semantics and orthography.

A more recent framework of the reading comprehension process is the morphological pathways framework recently developed by Levesque (Levesque et al., 2020). It is clearly based on the Reading Systems Framework but provides much more detail in particular illustrating how morphology significantly impacts reading (see Figure 2).

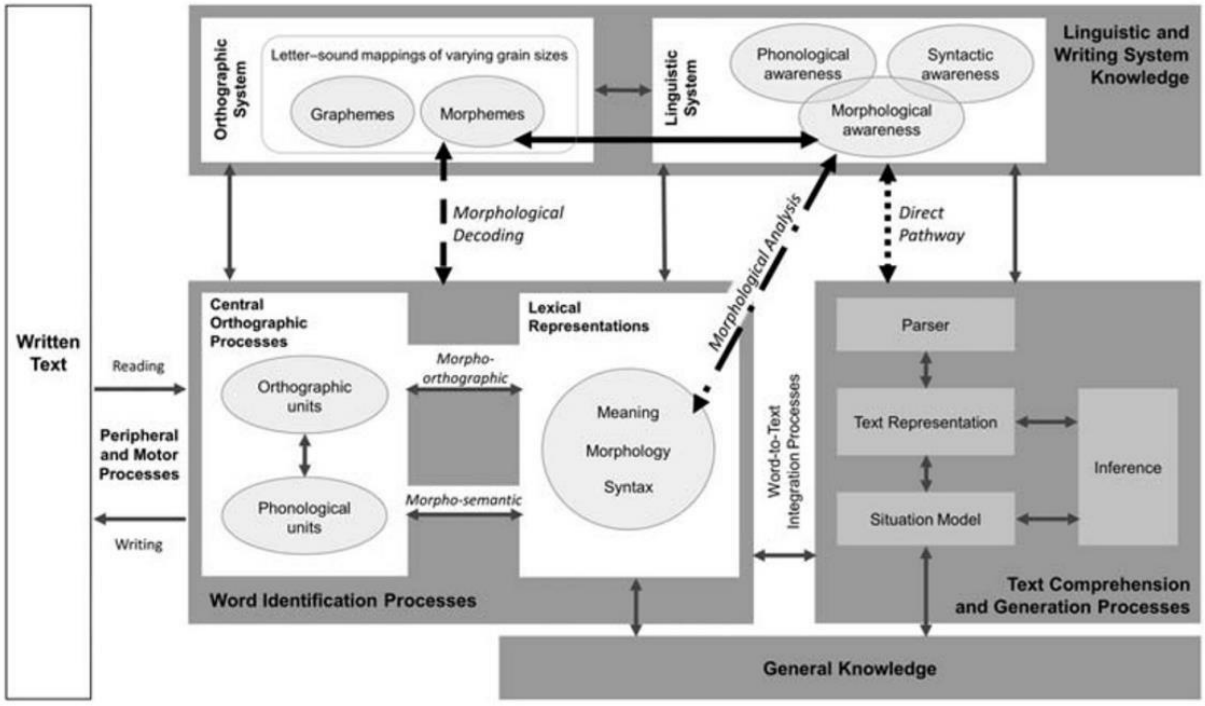

Figure 2:Morphological Pathways Framework (Levesque et al., 2020 p.3)

In this framework there is a critical differentiation between morphological analysis and morphological decoding. Morphological analysis is the use of individual morpheme meanings to attempt to understand morphologically complex words (e.g., a student uses their knowledge of the morphemes *un- hate* and *-ful* to understand the meaning of the word unhateful). Morphological decoding is the use of morphemes in spoken word reading using decomposition of morphologically complex words to come to a correct pronunciation of a written word. Connections should then be formed between the orthographic system and word identification processes and lexical representations in the linguistic and writing system, respectively. Thus, morphological decoding and morphological analysis are key processes in reading comprehension in this model and they are key because they for allow the identification and connection of morphemes to students' prior knowledge of words and morphemes in their reading. This allows students to establish connections within their existing knowledge to enhance reading comprehension. The work of Deacon and colleagues support this aspect of the model. In in multiple longitudinal and cross-sectional

studies using L1E primary aged students, they have examined and established the foundations of the relationship between morphological awareness and morphological decoding with reading comprehension.

(e.g., S. H. Deacon et al., 2014, 2017). For example, Deacon et al. (2014) followed 124 students from grade 1 through to grade 4 and report that the children's early morphological awareness partially explained their gains in reading comprehension, and that this was a reciprocal effect, in that their gains in reading comprehension explaining their gains in morphological awareness. Investigating a further 101 children from grades 3 and 5, they found that 8% of the variance in reading comprehension could be accounted for by morphological structure awareness, morphological decoding, or morphological analysis. This was after controlling for differences in word identification, phonological awareness, age, and non-verbal ability. The use of multiple aspects of morphology is unique and provides evidence for the importance of morphology in reading comprehension. This evidence along with her work recently with Levesque (e.g., Levesque et al., 2017; Levesque, Kieffer, & Deacon, 2018; MacKay, Levesque, & Deacon, 2017) supports the Morphological Pathways Framework and it demonstrates a growing trend that morphology has a significant impact on both the development of reading comprehension and reading comprehension in general for primary aged students.

A second new model has been proposed that is more expansive and combines multiple other versions which has potential to give a different perspective on reading and the varying influences and how they connect. This model is in the process of being tested and researched is the Direct and Indirect Effects Model of Reading (DIER) by Kim (Y. S. G. Kim, 2020, 2023; Y.-S. Kim, 2023; Y.-S. G. Kim, 2020) This work, and model are still in the early stages. However, the evidence that has been produced by the research so far seems to suggest that that it has potential. This model combines the Simple View of Reading(Hoover & Gough, 1990), the triangle model(Adams, 1990), construction -integration model(Kintsch & Intsch, 1988), the reading systems framework(Perfetti & Stafura, 2014), and automaticity theory(Laberge & Samuels, 1974) into one unified model that utilizes both skills and knowledge to attempt to explain how these components work together as a whole to produce reading comprehension. The illustration of the model is in Figure 1.

This blending of models and theories gives an interesting perspective on reading as it is not just focused on comprehension it looks at the broader picture. This could give a more well-rounded picture of

interactions that are possible. As part of Kim's work in validating this model studies have been run, the most recently published included a cohort of 372 children in grade 1 (mean age of 6.79 years) in the South-eastern United States, with 52% on free or reduced school lunch from 29 schools. However, this study only had two students who were classified as Limited English Proficiency, which means that the findings may or may not be transferable to EAL students. This study looked at how executive functions such as working memory and attentional control and morphological awareness explain variance between word reading and listening comprehension. They found that morphological awareness and executive function significantly explain the differences between listening comprehension and word reading with English speaking beginning readers.

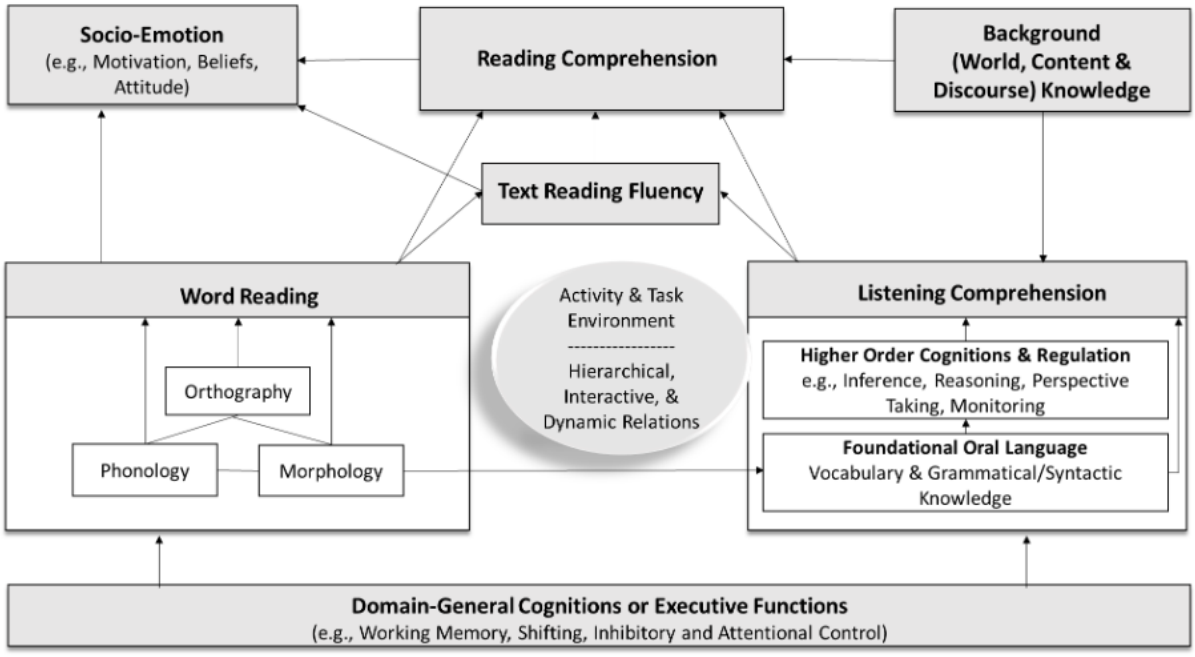

Figure 3: The Direct and Indirect Effects Model of Reading from Kim (Y. S. G. Kim, 2023, p. 8)

An earlier study with 201 native Korean speakers in grade 1 (6.84 years old), "supported the hierarchical relations hypothesis and revealed multichanneled direct and indirect effects of component skills."(Y. S. G. Kim, 2020, p. 469) This is support for the concept that as more complex skills develop they shift the dependencies on lower level skills. This additionally advocates the concept that for some of the skills the relationships are multi directional. This work has additionally been examined in a longitudinal form

beginning with a sample of 165 students in grade 2 (aged 7.5 years) from the South-eastern United States, 67% of the participants were eligible for free or reduced school lunch. There was no data given about home languages but, it was mentioned that 34% were from Hispanic backgrounds and 6% were from Asian backgrounds. This could mean that there is potential for 40% to have a different home language to English but, this is not confirmed. The participants were also tested in grade 4. The notable change that was found between the two times was that the direct effects for word reading and listening comprehension to reading comprehension were similar in grade 2 by grade four there was a change of +.22 for listening comprehension while it decreased for word reading by -.22. This could be indicative of the shift in skills as reading develops that is described in the models discussed earlier. Where there is a shift from reading individual words to comprehending passages more as a whole. This transition of skills could be illustrative of the transitions in reading development discussed by both Ehri and Seyomr and Dunncan. The connections between these ideas are not directly relevant to the original questions examined in this thesis but, they could be indicated by the data is gathered. Thus, the potential for applying this model needs to be considered as it is a potential explanation for findings that could be found.

Now having looked at a brief overview of the reading systems framework, the morphological pathways framework and, DIER model we need to move on to look at the research that informed the formation of these frameworks particularly the work done using EAL students and L1 English students in similar contexts.

This section has examined some of the theories that support this study. Starting with how the simple view of reading and its concept of improvement in either decoding or linguistic comprehension can improve reading comprehension. This supports the idea that if the intervention improves participants' knowledge of morphemes, it should in turn positively impact their reading comprehension ability. It then illustrated how morphology is viewed as an important component of reading comprehension, through the simple view of reading, the lexical quality hypothesis, reading systems framework, the morphological pathways framework and DIER. These theories illustrate how morphological analysis and morphological decoding skills could be an important piece to improving reading comprehension skills.

The Simple View of Reading has, however, been criticised for its assumption of the simplicity of the reading process. Its lack of detail has pushed others to flesh out additional frameworks that more explicitly illustrate where different reading skills fit and work together. The Reading Systems Framework (Perfetti & Stafura, 2014) promoted the idea of multiple subsystems working together and that breakdowns in any of these subsystems would lead to reading difficulties. The work done into how morphology fits into these frameworks led to the new morphological pathways framework which explicitly lays out in detail the different aspects of morphology and how they connect and interact with the subsystems in the models. This additional detail based in the research examined above gives insight to how strengthening morphological knowledge and skills could affect comprehension.

It is now important for us to consider the work that had already been done with regards to reading comprehension with both EAL and L1E participants. There has been considerable work that has been used to develop and refine the models discussed in this section. We will cover work led by Van Geldrern, Sparks, and Lesaux (Kieffer & Lesaux, 2010; Sparks et al., 2008; Van Gelderen et al., 2004)

2.1.5. An overview of some important studies of reading comprehension with EAL participants and L1 English participants

There have been several studies that involve the idea that L1 reading comprehension is a good predictor of L2 reading comprehension. For example, the large study of 397 Dutch L1 and English L2 secondary students conducted by Van Gelderen (Van Gelderen et al., 2004) found that there are similar patterns of performance for L1 and L2 reading, including strong positive correlations between vocabulary knowledge, grammar knowledge, and meta cognitive knowledge. This is similar to findings by Sparks that show that skills in an L1 correlate with skills in a L2 (Sparks et al., 2008). Sparks et al. (2008) studied 54 students from primary school through secondary school over 10 years, focusing on the development of their L1 skills and found positive correlations with the development of skills as they started to study a L2. The study also showed that there is evidence of the influences of weaknesses in an L1 when functioning in an L2 academic environment. The authors argue that L1

weaknesses may be related to struggles that students who are functioning in an L2 academic environment on a regular basis face daily. That is, for students who do not fully develop their L1 language academically and then are moved to a new country to study an L2, they may also have major struggles with their new (L2) academic language (see also the work of Cummins, e.g., Cummins, 2001). This idea is supported by work done by Lesaux, Crosson, et al. (2010) who looked at 87 native Spanish speakers who were developing English literacy skills in primary grades. The participants' word reading skills in either L1 or L2 were not significantly related to their English reading comprehension. However, the participants' oral language was a strong predictor of reading comprehension in both L1 and L2. The disconnect in this study between word reading and reading comprehension is notable as it is unusual. The suggested reasoning is the type of reading comprehension measure used, for example if it is a cloze procedure or measured by questions and a passage of text. Oral language is a stronger and more consistent predictor for reading comprehension than word reading (Lesaux et al., 2010). The participants had average word reading skills, but their comprehension skills were significantly below average. Their oral language skills both in listening comprehension and vocabulary were a strong predictor of reading comprehension. This could be due to the emphasis placed on phonics instruction in early years. It is possible that a focus on phonics instruction can lead to the development of an ability to read words aloud but may result in children missing the knowledge of what the word actually means (Darnell et al., 2017; Davis, 2013b, 2013a). There is also some concern that the results of this study could have been confounded by what is known as the '4th grade slump' or the period of time where students shift from learning to read to reading to learn (Chall & Jacobs, 2003). Teachers anecdotally will discuss this slump occurring anytime from midway through third grade to the end of fifth grade as students are making the transition from learning and developing basic reading skills to applying those skills automatically when reading. This connects with the theories of reading development by both Ehri and McCormick and Seymour and Duncan, discussed earlier, as there is a predicted shift towards the end of year 3, where students should have completed the phases of reading development and be moving toward reading to learn new information and not having to focus on the process of learning to reading as much (Ehri & McCormick, 1998; Seymour & Duncan, 2001). This is important for this research as it will involve both L1 English and EAL participants who have studied in the UK from at least

reception to year 3. These students were selected as they should be developmentally equivalent, and they should have been exposed to at least three full years of schooling in an English dominant environment. The issue that relates to the studies examined in this section is that EAL participants will not have developed their home or L1 to the same level as their English. This could cause significant issues with their reading comprehension (Cummins, 2001). Also, the participants are at a pivotal time in their reading development with the potential to be making the change from learning to read to reading to learn.

2.1.6. Shift from decoding and phonemes being keys in Reading Comprehension for EAL and L1E Students to morphemes and whole words.

An important consideration in both EAL and L1 English students is the changing importance of different factors as their reading skills develop. While in the initial years decoding and phonemic awareness are of critical importance, as students mature, a gradual shift occurs from reading words individually to reading more automatically, reading sentences as whole thoughts instead of as comprising individual words. This is consistent with the models of reading development by both Ehri and McCormick (1998) and Seymour and Duncan (2001) discussed earlier.

It has been repeatedly found that the reading comprehension of EALs is significantly less well developed than that of their monolingual counterparts (e.g., H. Deacon & Cain, 2011; Geva & Farnia, 2012). There are some other significant components that identify where there are weaknesses and strengths within populations of EAL students. It is important to note that EAL students are not at a disadvantage when it comes to accuracy, speed of grade appropriate word reading skills or, grade appropriate text reading in comparison to their monolingual peers (Geva & Farnia, 2012). Where the weaknesses seem to cause issues for EAL students is in vocabulary knowledge, syntactic knowledge and the comprehension of spoken language. These are all factors in the models of reading comprehension discussed previously. Support for this comes from the findings of a study conducted with 390 ELLs and 149 monolingual L1E students in Canada who were tested in Grade 2 and again in Grade 5 (Geva & Farnia, 2012). They hypothesised that the demands for reading text fluently, which does not necessarily require any

understanding of the meaning of the (words in) the text, are not as great as the demands to comprehend text. This would demonstrate that ELL students are able over time to "catch-up" with their monolingual peers on word level reading but, may still struggle on complex tasks that are multidimensional, automatic and involve actual comprehension.

2.1.7. Review of Reading Comprehension Studies conducted with L1E students.

When looking at studies of reading comprehension that are done with majority English as a first language participants, we find quite a few large-scale studies especially based in the United States. One of the major groups that has been involved in this research recently is the Language and Reading Research Consortium (Consortium, 2015). In a study with 325 first, second and third graders, they confirmed that the Simple View of Reading is a valid broad explanation of reading comprehension, even though, as noted above, there has been criticism of its being rather too simplistic a model (Perfetti & Stafura, 2014). One of their significant pieces of evidence for this is that word recognition and listening comprehension combined, explained almost 90% of the variance in reading comprehension at each of the three grades. Additionally, they confirmed that with the increasing automaticity of word recognition as students' progress from grade 1 to 3, there is a shift to listening comprehension as a better predictor of reading comprehension than word recognition. Another large-scale study, using 757 children in grades three through to five, confirmed that the simple view of reading is also still valid for slightly older primary age students (Lonigan et al., 2018). Specifically, the authors found that the contributions of decoding and linguistic comprehension shift as students' progress, with linguistic comprehension again becoming more important for older children. This shift appears to occur between 2nd and 4th grades. This change could affect the results from the current study as the participants are in this age range and might also be encountering this shift from decoding being of primary importance to linguistic comprehension taking more of a focus. The same idea was also supported by Storch and Whitehurst (2002 who found in a study of 626 children from preschool through to 4th grade, that the ability to decode became a less important factor for reading comprehension as students progressed through the primary grades. These studies together support the

idea that decoding becomes less influential as students' progress, meaning that linguistic comprehension becomes much more important for students as they start to read more complex texts. This is not to say that students stop decoding, but rather that they appear to rely more on linguistic skill. It has also been found in an analysis of studies that by the eighth grade reading comprehension and listening comprehensions form a unitary construct. (Hogan et al., 2016) One such study followed 604 students from second grade through eighth. While in the second grade almost all the variance in reading comprehension could be explained by word recognition accuracy, by eighth grade reading comprehension and listening comprehension were "indistinguishable constructs". (Adlof et al., 2006; Catts et al., 2006)This is additional support that there is a shift in focus for students as they progress from learning to read to using reading as a mode of learning. This is increasingly relevant as the current study is using participants who potentially are in this transition between the two. Additionally, this connects back to the models of reading development discussed, in particular to the dual pathway model by Seymour and Duncan, where students consolidate the recognition of letters to larger units such as morphemes.

2.1.8. Summary of Studies of Reading Comprehension

This section has attempted to provide a brief summary of the literature relating to reading and specifically reading comprehension with primary aged students. Two theories of reading development have been reviewed and discussed as to how they support the selection of upper primary learners for this study as they should have acquired or be in the process of acquiring the necessary skills to recognize and utilize the meanings of morphemes. The concept of reading has been defined and explored as it is a complex process. We then reviewed several accepted and one recently developed framework for reading comprehension and how these are being used in the current study. In particular, the work by Levensque indicates the potential for morphological analysis and decoding to support the comprehension of text. There was also a short discussion about the shift of importance as indicated by both theories of reading development and research to a growing importance of linguistic comprehension as students progress through schooling. Finally, there was an examination of research with L1E students and how it has been

established that there is a change in what skills become the primary focus for students as they develop as readers.

In the next section we will look at vocabulary knowledge and reading comprehension. It is well established that word identification skills and lexical comprehension skills are mutually supportive. It is also well known that knowledge of words and specifically word meanings is vital to support reading comprehension (e.g., Biemiller & Slonim, 2001; Hsueh-Chao & Nation, 2000; W. E. Nagy et al., 1984).

2.2. Review of Research in Vocabulary

Vocabulary is arguably the foundation to learning a language. There has been considerable research completed about acquiring vocabulary as part of both first language instruction in reading and in second language learning. The investigation in this dissertation focuses on students learning how to break apart vocabulary into morphemes and recognize that morphemes carry meaning to see whether this can aid in understanding new words and thus lead to the development of larger vocabularies.

This chapter/section will be a review of the closest concepts in vocabulary acquisition to the work conducted in this study. It will discuss both L1E and EAL research to illustrate the theoretical underpinnings of the current study and how the methodology employed was developed based upon the current work being done in vocabulary acquisition. It will begin with the concept of what it means to know a word and then progress to cover vocabulary size, vocabulary depth, the concept of vocabulary gap, current pedagogy, and concluding with an overview of guidance from governments about instruction.

2.2.1. What does it mean to know a word?

There has been a significant amount of research conducted into vocabulary knowledge and acquisition. At the heart of this is the concept of what it means to know a word. Knowing a word clearly means more than knowing how to spell it, but rather it requires understanding the meaning of the word and being able to use it effectively when speaking or writing. Nation asserted "learners think of knowing a word as knowing what the word sounds like (its spoken form) or how it looks (its written form) and its meaning. But learners not only need to know the form of a word and its meaning, but they also need to be able to connect the two." (2001, p. 23). Additionally, he noted there are nine important aspects encompassed within the idea of knowing a word (See Table 1 for details). We can see that he divides morphological knowledge into two word-part components that are "recognizable parts" (receptive skill) and parts that relate to "expressing meaning" (productive skill). The idea of recognizable parts being used to understand the meaning of a word is part of the idea behind the teaching of derivational morphology with derivational morphemes being affixes that carry or change meaning of another morpheme (e.g., unspoken: un- and spoken with un- meaning not, thus changing the meaning to not spoken). Among the receptive skills listed in the table is "association"; that is, what words does this word make us think of? How is it related

to (networks of) other words? An important concept within this skill, especially for students, is the ability to connect the meaning of a word to other words by recognizing other words containing the same morpheme (e.g., tran*scrib*e, *scrib*ble, pre*scrib*e, and de*scrib*e). The concepts of receptive and productive skills will be examined later in more detail, but it is valuable to recognize that the differences in these concepts are related to both derivational and inflectional morphology, the latter being morphemes that grammatically change words (e.g., the affixes -ing, -ly in talking or quietly). Derivational morphology is the aspect taught in the intervention examined in this thesis, as the concept emphasized is that morphemes carry meaning and when you know those meanings you can understand new words that also contain those morphemes. Nation's aspects of word knowledge also highlight the idea of a mental schema for vocabulary and how students themselves might perceive vocabulary.

Table 1: Aspects of Word Knowledge from (Nation, 2001, p. 25)

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word spelled?
	Word Parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the meaning?
Meaning	Form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?

Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	Collocations	R	What types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use (register, frequency…)	R	When, where, and how often would we expect to meet this word?
		P	When, where, and how often can we use this word?

R=Receptive Skill P= Productive Skill

Vermeer provides an additional perspective of word knowledge or mental schema for vocabulary. He proposes that knowledge of words "is multidimensional and encompasses various types of knowledge" (2001, p. 218). Thus, words can be thought of as nodes in a network; these nodes may be interconnected within different dimensions such as: "thematically (table–chair–couch), phonologically (table–fable–label), morphologically (similitude–gratitude–altitude), conceptually (cup–glass–mug), or sociolinguistically…" (Vermeer, 2001). These nodes can also be built up as part of a growing mental schema. Both Nation's and Vermeer's analyses of vocabulary knowledge are important to this research as they emphasise that the concept of vocabulary knowledge requires the consideration of conceptual aspects of vocabulary using several different perspectives. The perspectives of Nation and Vermeer are valued differently depending on the purpose of the discussion. While Nation is more focused on parts of word knowledge to be taught to students, Vermeer is concerned with how words connect in memory and recall. This connection on a morpheme level and the meaning of morphemes is what is being employed in theory with the aim of helping participants to deduce meaning when encountering multimorphemic words. Thus, word parts, concepts and referents, and associations from Nation are the areas of word knowledge that are being focused on in the current study. The study reported in this thesis encompasses concepts from these different perspectives depending on the specific question being examined. These differing perspectives also highlight why various aspects of vocabulary knowledge may develop in different ways or at different paces. These developmental variations may have a domino effect to other literacy skills such as

comprehension. Having considered the concept of what it means to know a word, the next section will look at vocabulary size and how many words students should know.

2.2.2. Why is Vocabulary Size Important for Reading?

The size of one's vocabulary influences the ability to read. As will be illustrated below, it has been shown that students need adequate vocabulary to understand what they read. When the vocabulary is not sufficient there are breakdowns in comprehension. For this thesis, vocabulary size will be defined as the number of words for which a participant has knowledge of both meaning and form. This definition is chosen because it is both a common method of defining this term and relatively straight forward to access (Schmitt, 2014). Vocabulary depth is different and will be discussed in the next section. The term word families is being used to describe a group of words consisting of a base word and all of its inflected and derived forms that can be understood without having to be learned separately (e.g. talk, talks, talking, and talked) (Bauer & Nation, 1993). Additionally, the terms root word and base word (e.g., *known* in unknown, or *angle* in triangle) are used interchangeably in this section dependent upon the original research usage. Both are being used as: a basic part of a word which may appear on its own with no additional morphemes in English (e.g., man, base, plan).

While this study is examining both EAL and L1E participants, the majority of the research into vocabulary size and growth of vocabulary has been done only with L1E participants, when focused on participants who would be of the age to attend primary school (e.g., Duff & Brydon, 2020; W. E. Nagy et al., 1984) . For example, the average adult L1E speaker knows about 20,000-word families, which they have acquired over the first twenty years of life at a rate of about 1,000 new word families per year (Nation, 2001). When looking at research with children, estimates have been made that at grade 2 (7 years old) students would have a vocabulary of between 5,175- 6,157 root words (Biemiller & Slonim, 2001). A more recent study with primary age students has determined that the average vocabulary size for fifth to eighth grade students is approximately 7,300-9,100-word families (Duff & Brydon, 2020). Research into university level EAP students' vocabulary, estimates that after 800-1,500 hours of instruction a learner's vocabulary range is approximately 1,200 to 2,000-word families (Schmitt, 2008). There were no strong

studies of vocabulary size for students who have been living in an L1E environment for the majority of their school career but who are classified as EAL found by this researcher.

An alternative way to look at the size of an adequate vocabulary for both EAL and L1E students is to look at the research done on how much coverage is needed for students to comprehend a piece of writing, which is to say, how many words per hundred words of text do students need to know for them to be able to understand a text. At various levels of schooling, it has been established that students need to be able to understand 48-49 words out of every 50 in a text to be able to understand the text (Carver, 1994; Hsueh-Chao & Nation, 2000). This would equate to 96-98% of the words in a text excerpt. A study by Hsueh-Chao and Nation (2000) was conducted with sixty-six adults in a pre-university English course in an English-speaking country where they were preparing to study of either a graduate or undergraduate degree. A piece of text was selected which included varying degrees of nonsense words replacing the original text. Participants read the text and then were given two comprehension tests, a multiple-choice measure and a cued recall test. Based on the results they determined that 98% coverage would be "adequate for unassisted comprehension" (Hsueh-Chao & Nation, 2000, p.419). A study by Carver (1994) was completed using L1E subjects who were in grades 3,4,5, and 6 and employed 100-word passages from textbooks and trade books (paperback books commonly used in for self-selected reading in classrooms that are levelled to make selection easier for students and teachers to know what should be accessible to students.) The children were asked to underline the words that they did not know. It was found that passages were considered easy when 0% are unknown basic words and relatively hard with 2% or more being unknown words.(Carver, 1994) This is related to the research reported in this thesis, it illustrates why students need to have access to skills that would help them to understand unknown words because of the evidence that even a very small number in a passage can make it difficult for them to understand. Carver's study can be argued to have employed a very simple methodology; however, it was a beginning to work that has developed more complex methodologies. The level of coverage in terms of word knowledge/vocabulary size has been found repeatedly with varying age groups and languages (Laufer & Ravenhorst-Kalovski, 2010; Schmitt et al., 2011; Webb, 2008; Webb & Rodgers, 2009). For instance, Laufer and Ravenhorst-Kalovski (2010) found with 745 college students in Israel who had studied EFL for at least eight years in high school, that there were two potential levels of vocabulary size that are needed. They report the optimal being knowledge of 8,000-word families or about 98% coverage,

and the minimum being 4,000-5,000-word families or 95% coverage on a standardized national test. This emphasises the importance of vocabulary knowledge for EAL students when reading. It also illustrates that a small increase in coverage could take knowledge of between 3-4,000-word families. Schmitt et.al. (Schmitt et al., 2011) asked 661 participants across eight countries to read passages and to complete a reading comprehension measure based on them. Together with a task a vocabulary measure, the authors found that 98% was a reasonable coverage target for readers of academic texts. Even though these studies have used a variety of methods and participant groups they have reported a similar figure; that is, in order to understand a passage, one must know about 95%-98% of the words contained in it.

Additionally, there has been some research on how many word families are used in printed school texts. This is important as students need to be able to access the texts that are used in schools. If a student does not have an adequate vocabulary to access the text, they are not going to be able to learn the material. Also, if word families are used as a unit of measure, students need to be able to recognize that those words are related, and they also need to recognize that their meanings are related. One widely cited estimate is that there are 88,533 distinct word families present in printed school texts for grades 3-9 in the United States (W. E. Nagy et al., 1984). While the Nagy & Anderson (1984) study is dated and many things have no doubt changed in texts since then, this figure is still regularly cited (Beck et al., 2013; Duff & Brydon, 2020; Stanovich, 1986b). This researcher has found no more recent, published, studies of the number of word families used in school textbooks written in English.

If we look at all of this as a whole, L1E students in grades 3-9 (9-15 years) would have a vocabulary of 915,000-word families which is just a small fraction (between 10-16.9%) of what would be needed to comprehend text. This leads to questions about what is being done to bridge this gap between what is known and what needs to be known for students to comprehend what they are reading in school. This gap could mean the difference between students being able to understand what its being read and not. If students were to be given the tools through an intervention or lessons on morphology to make informed deductions about what an unknown word means this could help reduce the distance between what they know and what is needed to understand school textbooks or printed work. Having explored vocabulary size, we will now turn to work vocabulary depth.

2.2.3. Depth of Vocabulary Knowledge versus Size of Vocabulary

The concept of vocabulary depth relates to the possession of multiple meanings and senses of words and/or having rich networks of thought associations for a word. For instance, Proctor, Silverman, Harring et al. state "A student who has depth of word knowledge for the word table, however, is aware that one sits at a table, but in reading a science text or conducting an experiment, one also creates, reads, and/or interprets a table. Further, while table is a noun, its morphological derivation, tabulate, is a verb, and is thus used in a grammatically and syntactically different context" (2012, p. 1636). While this concept was not measured in the current study it is an important component of vocabulary knowledge and should be measured in future work providing there are reliable and accessible measurement methods. An interesting hypothesis is that the connections and associations between words via their morphemes could potentially have an impact on students' vocabulary and thus lead to better comprehension of what they read. There have been strong correlations (r=.98) found between measures of vocabulary depth and breadth among young (4–7-year-old) Dutch children, which suggests that there are no significant conceptual differences when these two concepts when they are measured in either monolingual or bilingual children (Vermeer, 2001). However, there some researchers argue that the current measures of vocabulary breadth and depth are measuring extremely similar dimensions or a single common dimension leading to the high correlations between the measurements (Kieffer & Lesaux, 2012c) This very unsettled area of study will not be further addressed in the present work.

2.2.4. Differences in Receptive and Productive Vocabulary

Receptive vocabulary knowledge is the ability to recognize a word when it is presented, and to be able to connect it to its correct meaning. Productive vocabulary knowledge is the ability to produce a word when asked for a word based on some characteristic, such as meaning. Receptive vocabulary is usually viewed as more accessible and developing earlier, whereas productive vocabulary is more difficult to acquire fully, and it therefore develops later and more slowly than receptive vocabulary.

Significant gaps between the two have been observed when both receptive and productive vocabulary are measured and/or tested. Productive vocabulary has been found to be between 50 to 80% of the size of receptive vocabulary, when examined in university level EAL students (Fan, 2000; Ozturk, 2015; Schmitt & Schmitt, 2020). Research has found significant differences between being able to recognize a word and being able to use it when measures of receptive and productive vocabulary are compared. These gaps between receptive and productive vocabulary are arguably to be expected. This is potentially because it is more difficult to retrieve a word from memory and use it appropriately than to connect a word to its meaning when the word is encountered.

There has been some work into how students best learn receptive and/or productive vocabulary. It is generally accepted that if you learn vocabulary productively it will increase your productive vocabulary, and if you learn it receptively your receptive vocabulary will increase. For instance, Griffin & Harley's (1996) study with 110 students between the ages of 11-13 years old in state schools in the United Kingdom found that receptive learning was more effective for receptive knowledge based on results with English L1 students of French. Similar findings are reported in work with Japanese university students when learning vocabulary without context from word cards (Waring, 1997). In this study 60 students had two sets of 15 words to learn, one receptively and one productively. However, there have been examples where this clear link between instruction type (productive/receptive) and word knowledge has not been found. In a study by Webb with Japanese university students when given two different treatments to learn 20 nonsense words which were a mix of nouns and verbs. The first treatment was where they were instructed to write sentences using the nonsense words after being given example sentences for meanings in order to test productive knowledge. The second treatment was where they were given the example sentences and words and told to learn them. On ten separate measures of word knowledge, including orthography, association, syntax, grammatical function, form and meaning, that included both productive and receptive measures for each, it was found that participants who completed the writing treatment scored better on all ten measures than participants who received the reading treatment.

These studies are relevant to the present work as part of a number of considerations for the design of the intervention and selection of measures. As it is commonly accepted but not completely proven that

instruction in receptive vocabulary or productive vocabulary yields better results for each type of knowledge, the morphemes are instructed and practiced in both receptive and productive manners but due to timing constraints and for consistency across measures they are only tested receptively. The current study used the British Picture Vocabulary Test 3 (Dunn & Dunn, 2009), a measure of receptive vocabulary, as it was readily available and standardized for the appropriate age range. This is the British version of the widely used Peabody picture vocabulary test. It is frequently used in research because it is easy to administer (e.g., Hatcher et al., 2006; Lawrie et al., 2005; Oakhill et al., 2003). It is however standardized for a majority L1 population, so the standards may not be as equitable for EAL students. These issues will be further discussed in later sections in the thesis. Additionally, as discussed above, receptive vocabulary tends to be larger and more accessible than productive vocabulary. As activities in the intervention are both receptive (e.g., reading passages to learn meanings) and productive (e.g., writing sentences with newly learned words containing target morphemes), selecting if the measures for vocabulary should be productive or receptive is not a simple decision. However, as the study by Warring (1997) illustrated that if vocabulary was learned productively, it could be tested both productively and receptively, but the inverse was not the case as vocabulary learned receptively was not as easily accessed productively. It is logical that using a receptive measure would best illustrate if there were any changes in vocabulary size due to the intervention.

As well as being differences between receptive and productive vocabulary sizes in individuals, there are additionally documented gaps of vocabulary size within student populations. The size of the gap along with possible causes and the Matthew effect are explored in more depth in section 2.2.6. Next will be a short discussion of academic vocabulary as it will play a role in studies discussed in upcoming sections.

2.2.5. What is Academic Vocabulary and Why is it important?

Academic vocabulary is closely defined and intertwined with the most common definitions of academic language. In this dissertation academic vocabulary will be defined using the parts of the characteristics of academic language by Nagy, Townsend, Lesaux and Schmitt (2012) that relate specifically to vocabulary. These characteristics are: Latin or Greek vocabulary, morphologically complex words, informational

density and abstractness, since the two parts that are focused on in the intervention developed and employed in this dissertation are Latin and Greek vocabulary and morphologically complex words. Latin and Greek vocabulary are still regarded as more prestigious than words with Germanic etymologies. (Nagy et al., 2012). Many words on the widely utilized Academic Vocabulary List do not have an obvious Germanic equivalent whereas the majority have visible Latinate or Greek origins (Coxhead, 2000; Lubliner & Hiebert, 2011). The Academic vocabulary list was developed by looking at corpus data from common academic texts and looking at the most commonly occurring word families (Coxhead, 2000). In the following section it will be demonstrated that academic vocabulary occurs in texts utilized in primary grades as well as at university levels so developing an awareness of how to decompose morphologically complex words and interpret meaning from them can arguably only increase in usefulness for students as they progress through their time in school/in academic learning. Morphologically complex words are those that have more than one morpheme and academic vocabulary tends to have longer words due to affixation, sometimes multiple levels of affixes (e.g., Predisposition: pre + dis + position) (Nagy & Anderson, 1984). There have been calls from higher education in the United States to start instruction in academic language early, as it is a necessary skill for success in both academia and business (Scarcella, 2003). Included in this call is a framework of components that should be included in academic language which explicitly includes a lexical component that is defined as "knowledge of the forms and meanings of words that are used across academic disciplines (as well as in everyday situations outside of academic settings); knowledge of the ways academic words are formed with prefixes, roots, and suffixes, the parts of speech of academic words, and the grammatical constraints governing academic words" (Scarcella, 2003, p. 12). This framework of knowledge that needs to be acquired to truly understand academic vocabulary additionally includes, components such as phonological (e.g., knowledge of sound, stress and intonation patterns), grammatical component (i.e. knowledge of grammatical features associated with academic language), sociolinguistic component (e.g. knowledge of a variety of genres including making requests and argumentative text) and finally discourse component (e.g., introduce topics and keep the talk going and for beginning and ending informal types of writing). This knowledge that is being called for will need extensive exposure and time to be developed. This leads to the premise that it should be started early in formal education, even potentially in the primary grades, so that students are not barred from

progressing and attaining a potential career goal simply on the basis of lack of knowledge of academic language.

There have been interventions focusing on academic vocabulary that have been moderately successful. The ones that are closest to the work in this dissertation have been done with slightly older students. The interventions have also been for longer periods of time, typically between 16-24 weeks. They were considerably longer than the present intervention that was scheduled for 10 weeks but only lasted nine (see Methodology section 3 for a detailed description). Lesaux et al. (2010) conducted a large-scale teacher-implemented intervention with Language Minority Learners in the United States which was focused on academic vocabulary instruction. It involved 21 classes of sixth grade students with 346 language minority learners and 130 native English speakers. The program lasted 18 weeks, with 16 weeks of instruction and two weeks of review, one at the midpoint of the program. The program resulted in significant positive effects for meanings of taught words, and also for morphological awareness. In an article looking at just the morphology component of this study, they reported significant positive impacts on both relation and syntactic aspects of morphological awareness ((Kieffer & Lesaux, 2012b). The authors additionally estimated that the intervention resulted in about four months' worth of additional growth in decomposition development beyond what would be expected due to typical schooling and maturation of students.

There have been questions about teaching academic vocabulary to primary aged students as it is seen as a skill that would be more needed in higher levels of education. However, it is reasonable to consider introducing academic vocabulary at younger ages so that it builds a foundation for later exposure and use (Galloway & Lesaux, 2015). This could additionally contribute to the background knowledge of students, which would lead to a lower cognitive load when encountering new related concepts. This is additionally supported by the models of reading development discussed in section 2.1.2. In particular, the final phases of both models where consolidation has taken place and processing of larger groups of letters or morphemes occurs. Along these same lines of building foundations for future learning have been

suggestions that if students develop an early interest in words, this word interest will encourage effective word usage, continued vocabulary growth and broader word exploration in general (Zucker et al., 2021) .

Academic vocabulary size increases in importance as students' progress to higher levels as the complexity of vocabulary used in instruction also increases, so early access could support later learning. While research about this is not straightforward, as noted above, there have been moves by governments to encourage academic vocabulary instruction from the early primary grades (see also: National Governors Association Center for Best Practices and Council of Chief State School Officers, 2009) Part of this move has come from the work by the National Reading Panel (Panel, 2000) which advocates for a focus on vocabulary instruction as a regular part of reading instruction. This is based on an analysis of current research into reading instruction (National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.)., 2000). While this is likely to support future students, there are questions as regards the usefulness with populations that are already behind. This leads us to the concept of vocabulary gap or the difference between the vocabulary known by a student and the vocabulary used in curriculum.

2.2.6. The Vocabulary Gap between students and curriculum

One concept that is regularly discussed among educators is the vocabulary gap or the difference between what students know and what they need to know in order to access the curriculum. The term vocabulary gap has been used to describe multiple different facets, including the differences between different subgroups (e.g., EALs/ L1E, SES) of students the same age in vocabulary size (e.g.: summary of this concept in Blachowicz et al., 2006, p. 562). It is being used in this dissertation to describe the difference between the vocabulary that students already know and what is needed for them to understand what is being taught and read in the classroom at their grade level. This is important as students need to be able to access material presented in their classroom and if a student is not able to do so, this could have a significant impact on their future academic success.

In this section we will examine the size of the differences in student vocabulary at the same age, one possible theory as to its effects and several possible reasons for this difference in vocabulary size.

2.2.6.1. How large is the gap for primary students? Why are there differences in the size of the gap for students?

Students who have a vocabulary gap of one standard deviation (SD) (from the average score for their age), on a standard measure of receptive vocabulary such as the Peabody Picture Vocabulary Test (PPVT) will have a vocabulary that is approximately 1,246 words smaller than a student who has attained the mean score for that grade (Dawna Duff & Brydon, 2020).Which is a significant hurdle to overcome if students generally acquire between 8-12 new words a week during the school year as estimated by Biemiller and Boote(2006).

An additional line of thought is that there is difference in the development of word recognition skills. This was reported in a study by Duncan and Seymour (2000) with 160 students from nursery to primary year 3 in Scotland where there were equal groups of L1E speakers from both high and low SES tested on measures of letter knowledge, content word reading, non-word reading, the British Ability Scales Word Reading Test, The Wechsler Preschool and Primary scale of Intelligence or Wechsler Intelligence Scale for Children (dependent upon participant age), and a meta phonological awareness task. They found that low SES students performed at the same level as the high SES students who were a year younger on measures of word recognition. They also found that the students' logographic and alphabetic abilities developed in parallel. Finally, they established that low SES reading develops along the same pathway as High SES reading but with delayed development. This could indicate that while there is a notable delay, it is not necessarily due to any specific learning issues in the low SES population. However, due to the delay in development of these skills, the movement of children through the different phases of reading development might be delayed overall, as while the phases have some overlap, they are built upon each other. Methods that could assist in speeding up the development of these processes could theoretically be useful in helping students acquire skills and reduce the distance in reading ability between students who seemingly start at a disadvantage, whatever the reason for that disadvantage is. The earlier students can build up these skills, reducing the gap distance they would need to overcome, if as theorized by the Matthew effect, the gap widens over time.

Thus, there is a need for methods that will assist teachers in accelerating vocabulary growth and reading skills for students who are lagging. If this gap widens as theorized by the Matthew effect, there is an even

greater need to reduce the gap at the earliest possible stage. This leads to the next section providing details on the Matthew Effect.

2.2.6.2. What is the Matthew Effect? How is it relevant for vocabulary?

The Mathew effect is based on a verse from the biblical gospel of Mathew that states 'For to all those who have, more will be given' (Matthew 25:29). This academic term was originally coined by Robert K. Merton while explaining discrepancies in recognition received by scientists doing similar work (Merton, 1968). It was later adopted by Stanovich to explain differences in the acquisition of literacy(Stanovich, 1986a). In terms of vocabulary knowledge and vocabulary growth, the Matthew Effect describes the state of affairs when those who already have a large vocabulary are likely to be exposed to more words while reading more complex texts thus progressively expanding their vocabulary. In contrast, those who have smaller vocabularies to start with presumably read fewer complex texts and this in turn reduces their likelihood of encountering new words resulting in a reduction in their acquisition of new vocabulary compared to those with larger vocabularies to begin with. Related to this is the possibility that those with larger vocabularies will more fully comprehend anything they read and through better comprehension are more likely learn new words through their use in context.

The Mathew effect has not been conclusively proven or disproven. In a meta-analysis of reading research over 25 years no definitive conclusion about possible relationships with the Matthew effect and reading achievement over time was found (Pfost et al., 2014). There have, however, been studies that have found a positive relationship between strong early word reading skills and later vocabulary growth and size (Duff et al., 2015). Duff, Tomblin and Catts (2015) investigated 485 English monolingual students. This study collected data at kindergarten, fourth, eighth and tenth grades to model the rate of vocabulary growth after fourth grade controlling for the vocabulary level at kindergarten. The authors report that word reading in fourth grade accounted for 8% of the total variance in growth between the fourth and tenth grades. They additionally found that fourth grade vocabulary size had the same impact on vocabulary growth rate as maternal education levels during the same period. On the basis of these results, they hypothesize that it is not the case that only vocabulary size influences vocabulary growth, but that vocabulary size may be an indicating variable of other reading related activities that students would be

involved in. Therefore, those who had stronger word reading at an early age will have had more access to new vocabulary and thus more chances to gain vocabulary which would be compounded over time and finally leading to a larger vocabulary in the later school years.

The Matthew effect is a popular, but unproven theory. It certainly is an area deserving more longitudinal research. There are additionally theories for vocabulary gaps, that have not had as much exploration, and these will now be addressed.

2.2.6.3. Potential reasons for Vocabulary Gaps among primary students

There has been a wealth of research into vocabulary gaps and the long-term impacts of these gaps on students. One highly cited reason for vocabulary gaps in monolingual students without documented learning disabilities, is the educational level of a student's parents. More highly educated parents speak to their children more and use more unique words starting from infancy (Hart & Risley, 1989). For example, professional families spoke 32 million more words to their children between birth and 48 months than parents who were on welfare (Hart & Risley, 1989). However, the methodology of this work has been questioned along with the small sample size of 42 families. When similar work was completed using all ambient speech, not just primary parent speech, and through corpus research, the results have not been as large (Sperry et al., 2019). Hart and Risley (1989) found that professional families spoke on average 2,153 words per hour to the child, while middle class families spoke 1,400 words, and families on welfare just 616 words. The more recent study by Sperry et al. (2019), using corpus from completed studies, found that middle class families spoke 1,491 words per hour, while working class spoke between 1,351-1,048 words. They noted that the number of words per hour depended on the location of where the corpus was collected. There was also a significant discrepancy between studies of two groups of poorer African Americans, one at 1,838 words and the other at 1,061 words per hour (Sperry et al., 2019). This illustrates that there could be gaps not just based on socioeconomic status but geographic location and possible other variables as well. Other work completed using recording devices attached to children's clothing and recording for 12-hour blocks starting when the child woke up one day a month for six months or 38 months depending on which phase of the study they participated in (Gilkerson et al., 2017). In this study,

the recordings were then analysed using LENA technology to determine occurrences of adult and child speech. Using this much larger and varied time of day sample they were able to confirm that there is a large gap between the amount of speech children are exposed to based on the mother's level of education. However, this was only observed when they compared the 98th percentile to the bottom 2% of families. They then come up with a gap of 29.95 million words, not quite the 32 million that Hart and Risley came up with but still a significant difference. The authors do caution, however, that there was wide variation within each SES group. This may mean that differences in adult speech around the child are not solely based on mother's educational level. This study was conducted with a much larger sample of 329 monolingual families that was matched to the US census data at the time the study was conducted.

A recent study with children starting at 18 months from monolingual English-speaking families has shown that there are differences in the development of spoken vocabulary use (Fernald et al., 2013). The sample comprised 48 families, with 23 being categorised as lower SES and 25 as High SES based on the Hollingshead four factor index of socioeconomic status which is based on both parents' education, and occupation. Participant responses to both visual and verbal cues along with an inventory completed by the parent were used to determine vocabulary size. It was found that while all participants gained on average of 225 vocabulary words over the period between 18-24 months, the differences between the two groups were significant. The lower SES group started with half the group only being reported with a vocabulary of fewer than 50 words, whereas most of the higher group had a vocabulary of more at the start. At 24 months the higher SES group reported vocabularies of 450 words and the lower group produced about 150 words less. This illustrates that the potential for significant differences by SES groups start very early and then continue to grow. This study, while focusing on very young participants, does give an indication that differences appear very early. The use of a parental report could influence the results, however, as parents are likely to think their child is the best and potentially knows more than they actually do. Though with the use of participant response to stimuli this issue could be balanced out.

There has been some work done with monolingual middle and high SES groups with two-year-old children and via recording observations ten weeks apart to examine both maternal language use and child language use (Hoff, 2003). The high SES group comprised of 33 mothers and their offspring with representation of multiple birth orders as well. The middle SES group comprised 30 mother and offspring

pairs with a similar representation of birth order. When birth order, and child vocabulary at time 1 were accounted for, SES accounted for 5% of the variance in child vocabulary at Time 2. More interestingly and potentially more easily changed, is that the quality of the maternal speech was the mediating factor. Children who were observed as producing more varied lexis and more complex structures were exposed to more complex lexis and structures from their maternal caregivers (Hoff, 2003).

Work has been carried out by Buac, Gross and Kaushanskaya (2014), with 58 Spanish and English bilingual children aged between 5 and 7 years old in a laboratory environment testing vocabulary of the children and that of their primary caregivers. Receptive vocabulary and productive vocabulary were tested in both English and Spanish along with non-verbal IQ. The authors found that caregivers' vocabulary knowledge, the amount of exposure the children had to each language and SES were all strong predictors of children's English, but not Spanish, vocabulary skills. It was also found that higher SES was positively associated with more English use in the home.

In a newer study by Cangelosi et al., (2023), with 22 children from age 9-11 years in Italy, when looking at how SES might affect reading comprehension and vocabulary found that SES plays a key role in affecting literacy skills for both monolinguals and bilinguals. They found that only SES was able to account for differences in reading comprehension skills and vocabulary knowledge, independently of bilingualism. They also propose that bilingualism no longer affects literacy after 4 to 6 years of schooling in a L2. This overall, contributes more to the growing field that SES is a key factor in the development and growth of reading and vocabulary in more than just L1E contexts.

These studies have been used to promote the idea that lower SES students have been exposed to smaller vocabularies, which drives the vocabulary inequality between socioeconomically disparate groups. Based on methodological concerns within the studies alone, the conclusion should not be as widely accepted as it is. This area needs more research and for the research to be conducted in a way that is culturally unbiased. What we can reasonably assume is that there is likely to be some difference between groups based on exposures at home. If we then think about the lack of in-home English exposure for students who are EAL, who live with one language at home and another at school, this gap would be substantially wider than for students living in English speaking homes. Such EAL students, who use their L1 within their family, would miss out on the likelihood of having parents available to, for instance, explain the

meaning of English words that are new to the child. The effects of diminished home background knowledge of English would also reduce a student's comfort when attempting to read or speak English at school.

The vocabulary gap between students in a single classroom also may be vast, especially as EAL students enter English speaking schools at different grade levels. Widely varying vocabulary gaps indicate a need for teachers to have methods for detecting and tools to remediating vocabulary gaps as rapidly as possible, since such gaps will potentially expand over time, as outlined in the research above. One suggestion for doing this, is to expose students to academic vocabulary from an earlier age. If exposure at an earlier age is not possible, then techniques or methods for accelerating vocabulary development would be valuable. In this study the plan is to give students tools to help understand academic vocabulary explicitly and earlier so that they would have more practice with the vocabulary before encountering similar, morphologically complex words, in future higher-level texts. It was also hoped that these techniques would enhance or accelerate the acquisition of vocabulary.

This leads to the next section on instruction and current research on classroom practices. This area is not settled and has multiple methods that are currently being explored.

2.2.7. Two Current Methods of Vocabulary Instruction

Vocabulary instruction, particularly for aiding reading comprehension, has previously tended to fall into two camps, the shallow but wide camp helmed by Biemiller (Biemiller & Boote, 2006; Biemiller & Slonim, 2001) and the contrasting narrow but rich approach led by Beck (Beck et al., 2013; Beck & Mckeown, 1983; Beck & McKeown, 2007). Both approaches will be addressed in the following sections in more depth.

2.2.7.1. Method One: The Shallow but Wide Approach

The results of a study by Biemiller and Boote (2006) suggest that students can learn at an average rate of 3-6 root words (a basic part of a word which may appear on its own with no additional morphemes in English e.g., man, base, plan) per day with brief explanations as they are encountered in context. His

study used classroom instruction with a teacher reading the text and asking comprehension questions at the end, there was also variation by groups on repeated readings, depending on whether the teacher gave explanations of target words or not. Using an intervention involving repeated readings combined with explanations of a significant number of words and reviews of those words, students learned between eight- and twelve-word meanings a week, in a population that was 50% ELL students. This method could therefore add between 288-432 new words to students' vocabulary over the course of a typical 180-day school year or between 1.6-2.4 words per day. This is just a small portion of the total 2,500 root word per year average that is acquired during the primary years. The testing used in this study involved students being read a sentence and then asked to explain a target word from the sentence either orally for younger students or in written form for older students. When students were tested using the same measures but without intervention in a different study by Biemiller and Slonim (2001), the investigators found that students in grades 3-5 gained on average 2.9 root words per day. This illustrates that while there is potential value to the intervention, the impact between non-intervention of 2.9 words and 2.4 words with intervention suggests that the intervention might have reduced the rate of word acquisition, an unlikely but, concerning possibility. Such a possibility indicates the need for additional research in the area. This of particular interest to the design of the intervention as it was influential to determining the number of affixes that could be taught in a session.

2.2.7.2. Method Two: The Deep but Narrow Approach

Beck, McKeown and Kucan (2013) have posited that vocabulary instruction will only affect comprehension if it is tied to deep processing and multiple varied experiences, such as the use of the new vocabulary in collocations, word associations, and in the student's own writing. They have called for robust instruction in vocabulary for students that can potentially drive an interest and awareness of words to keep students engaged with new words until they "own" them as a fully functional part of their vocabulary. They spend well over 200 pages in their book "Bringing Words to Life: Robust Vocabulary Instruction" discussing the details of how and why such instruction is grounded in research but explained in practical terms. In a study conducted with kindergarten students, using both a repetition intervention and an interactive intervention, it was found that students participating in both performed better than the

control students but, students in the interactive intervention performed better than those in the repetition intervention on all outcome measures (McKeown & Beck, 2014). The improvement that was found with the use of both types of instruction combined being more significant is part of the rational for the instruction methods selected for the intervention that was designed for the study in this thesis.

2.2.7.3. Which words to focus on in instruction

When approaching the question of which words should be focused on for instruction, 'the shallow but wide' and 'the deep but narrow' camps largely agree. They both posit words that students would not pick up though regular exposure and words that are very domain specific should not be the focus of instruction. Biemiller says that words that between 20-70% of students have had prior exposure to, should be the focus of instruction as they already have some knowledge of the word (Baumann & Kameenui, 2004). While Beck, McKeown and Kucan have a more specific three-tier system to use to categorize words into groups based on instructional value. The three-tier system is made up of tier one words which are typically found in oral language. Tier two comprises words that are common in written language, but not as common in conversation. Tier three involves words that are very rare, and therefore encountered by few readers (Beck et al., 2013). They advocate that the majority of instructional time for vocabulary should focus on the tier two words because they would be less likely to be encountered in conversation and picked up incidentally and are more useful for reading than words that would only rarely be used in specific contexts. Tier one words will be learned in conversation and if tier three words are encountered, they can be quickly explained or learned from a dictionary.

This leads to instruction using words that is focused on acquisition of vocabulary used in written language to which students are likely to have repeated exposures. This selection aligns with Beck et al.'s (2013) proposed methodology of intervention, involving repeated exposures and opportunities to work with and repeatedly use target vocabulary in various contexts.

2.2.8. Governmental Guidance about Vocabulary Instruction

Both the United States and United Kingdom have produced documents outlining the importance of vocabulary instruction in primary grades and to support reading. In 2000 the National Reading Panel

(NRP) released its report and recommendations for instruction in the United States (Panel, 2000). In 2015 the UK Department for Education released a white paper 'Reading: the next steps' (Department for Education, 2015). Both documents were aiming to guide classroom instruction. Both recommend that students be given multiple encounters with a word in different contexts to enhance student's ability to develop their understanding and knowledge of a word.

These documents also state that background knowledge has an impact on what a student can understand in a classroom. This means that if a student arrives with little knowledge of a subject, it will be harder for them to understand what is being taught and thus they will be slower to learn (as already discussed above). This principle will also impact vocabulary size and knowledge. It is akin to showing a student a picture of a stool and telling them to build it without them ever having seen or held a hammer before. In contrast, other students arrive at class having already used a hammer with their family thus understanding the hammer's use. These gaps in knowledge between students, as previously discussed, can compound over time and lead to increasing deficits for affected students. These knowledge deficits are in addition to the possible restraining psychological effects of feeling less prepared or less intelligent among students with larger gaps, subjects beyond the scope of this thesis.

The NRP report examined research, from 250 studies, and proposed that while no single method has shown optimal results for student vocabulary growth, there are methods that have shown promise and therefore should be employed in the classroom. These recommended methods include multiple encounters with words, opportunities to use words in varied contexts, active engagement in learning tasks, and utilization of a variety of methods since no single method will result in peak learning for all students. These four principles were considered fundamental in the design of the intervention used in the research presented in this thesis. Specifically, students were given multiple encounters in varying contexts with key morphemes. There was active student engagement in the process using a variety of methods as they deduced meanings, mapped words, sorted words, and used in them in new sentences of their own creation. All of this is explained in more detail in the methodology section 3.2 (Organization of the intervention and how it was implemented).

2.2.9. Summary of Vocabulary research

This section has provided a brief summary of the literature relating to vocabulary in relation to what is required for successful reading comprehension. While a significant body of work exists about the general size of vocabulary and the amount of knowledge a student would need to comprehend text at different grade levels, there are still many gaps in our understanding as to how to optimize vocabulary learning for EAL students who are studying in an English dominant environment. With the increase in immigration in both the UK and US and the subsequent rise in students who speak a non-English language at home, much more research is needed in this area. Current work with vocabulary instruction in classrooms and the choice of words to be instructed have been discussed. Existing research literature and government agency reports suggest that multiple methods and exposures should be employed as they are the strongest methods of intervention to enhance vocabulary growth (and in turn, reading comprehension). Words for instruction should be selected based on utility for students and possibilities for repeated encounters in varying contexts.

The next section will examine the development and current work relating to morphology, reading and vocabulary. This will continue the narrowing of focus for this literature review from reading comprehension to vocabulary and finally morphology.

2.3. Review of Research into Morphology

Morphology is the study of morphemes and how they are combined to form words. A morpheme is the smallest unit of meaning in a word; it can be a prefix, suffix or root. For example, the word unkindness, has a prefix-un, root -kind, and suffix- ness. Roots can be both bound and unbound; a bound root is one that cannot stand alone, that is, it is not a word when it is not connected to an additional morpheme, for example the derivational morpheme /-ness/ must be attached to another morpheme (e.g., fond-*ness*). An unbound root can stand as a word on its own; for example, /fond/ is an unbound morpheme as it does not need any additional morphemes to make it a word.

Many terms have been used to describe different aspects of morphology. As such a short review of the main terms and how they will be used in this thesis is presented here.

Morphological knowledge comprises three parts: relational, syntactic and distributional. Relational knowledge refers to being able to see that words may share common morphemes (i.e., how dis**rupt** is related to inter**rupt** but, not **dis**play.) Syntactic knowledge is knowing that *walk* is a verb and *walker* is a noun. Distributional knowledge is knowing about constraints on the concatenation of stems and affixes (e.g., the rules of how roots and affixes can be combined knowing that im- can be added to mature but the prefixes in- and un- cannot) (Tyler & Nagy, 1989). Morphological Knowledge can be further divided into Awareness and Analysis.

Morphological awareness is, "the ability to reflect upon and manipulate morphemes and employ word formation rules in one's language" (Kuo & Anderson, 2006, p. 161) e.g., walk is a verb, when you add the suffix -er it becomes the noun: walk**er** or, pig**let** means little pig so if one adds -let to owl: ow**let** would mean little owl. Put simply, morphological awareness is knowing that morphemes are part of a word and being able to use them to both construct and deconstruct words.

Morphological analysis is "to break an unfamiliar word into parts and then recombine the parts into a meaningful whole" (White, Power, et al., 1989, p. 283). Others define it in a more complex manner i.e., inferring meaning of an unfamiliar morphologically complex word based on its morphemic constituents.

(Baumann et al., 2003; S. H. Deacon et al., 2017; Levesque et al., 2017). Earlier work by Anglin (1993) termed it 'morphological problem solving'. Put simply, morphological analysis involves being able to find and understand the individual morphemes in a morphologically complex word to enable better understanding of that word. For example, when a student sees the word *unstoppable* and they recognize the prefix /un-/ means not, the root /stop/ they already know, and they understand the suffix /-able/ to mean *capable of*, they then combine the parts to create the meaning "not capable of being stopped".

Having defined what is meant by morphological analysis, morphological awareness, and morphological knowledge, it is important to note that the focus of the intervention in this study is morphological analysis. In order to understand the empirical foundation of this study, we should first look at why morphological analysis might have an impact on a student's reading and understanding. White, Power, and White (1989) looked at how many words in school texts are analysable prefixed words. They found that in grade 3, 230 words per million words of text are analysable. This number increases when you look at grade 5 to 631 words per million words of text. They also found that the most familiar meaning of a root word was relevant for understanding the meaning of the target word in 60% of the cases. As such, if a student encounters a word that they are able to deconstruct morphologically, then there is a likelihood that they could use their knowledge of morphemes to understand the meaning of the word. They only looked at words formed with four major prefixes (e.g., un-, re-, dis-, and in- (meaning not)) and the roots using the most common meaning. This limits the scope of the study as there are many other prefixes in English that are used regularly. This could indicate that the number of analysable words is actually much higher than stated. Suffixes are used and examined if they are derivational and not inflectional, 9.8% of the words had multiple suffixes. They also computed that 90% of the terminal suffixes were inflectional or neutral derivational. Additionally, this study is rather dated, it is hoped that school texts have evolved in the thirty years since this was conducted. As it is the only study that is widely used for this illustration it is a starting point for estimates.

A pair of recent studies have looked at the prevalence of morphologically complex words in core vocabulary and in widely used curricula. The first is by Hiebert, Goodwin and Cervetti, who examined the presence of morphologically complex words in the core vocabulary of school texts and found that

2,452 morphological families and that the 11,298 words in those families encompass 58% of the 19,500 most common words in written English (Hiebert et al., 2018). It is also interesting to note that 85% of these families have been encountered by the end of the elementary grades based on exemplar texts that accompany the Common Core State Curriculum. This indicates that a significant exposure to morphologically complex words is expected for students, even during the elementary years. The 2,451word families also accounted for 74.9% of words in text in the final two years of formal schooling in the

US grades 11 through college (ages 16-18 years old typically). A more recent study by Kearns and Hiebert (2021) that used three widely adopted literacy series for its base corpora, found that polymorphemic words are present in significant quantities even in first grade texts. They observed that it is likely that one in every five or six words will be inflected or will be a compound. By the third grade this number of morphologically complex words has risen to one in every fourth word. The type of morphological complexity also shifts with more compound words in first and more derived words in third grade. This is additional evidence that students need to be able to access the meanings of morphologically complex words and the need for these skills increases as they progress through their school years.

Interestingly, White et al. (1989) also investigated what percentage of these morphologically complex words were known by students, they found 3rd grade students knew about 50% and 4th grade students knew about 60% of the prefixes in the sample without any explicit instruction. Considering this finding together with the information about reading development and how morphological analysis impacts reading comprehension, the large proportion of analysable words would indicate that instruction in morphemes and their meanings may indeed improve the percentage of understood words in texts. If we apply this to the improvements shown in the research reported in the previous sections above, the implications seem to be that instruction will only raise these percentages of understood words in text. This is particularly important for students as studies have shown that textbooks become more complex and use more academic vocabulary as throughout schooling (Fang et al., 2006; W. E. Nagy et al., 1984). The increased complexity found in school texts has been studied by looking at nominalization (that is, the formulation of nouns from other parts of speech, e.g., the abstract noun *comparison* created from the verb *to compare*), which is a characteristic of academic language along with morphologically complex words. It was found by analysing three texts usually used in three grade ranges. The first text was a fable used in

grades K-3, the second was a science text used in grade 4-8 and finally a history textbook passage used in grades 9-12. Using a method developed by Halliday (1993), the authors found that the texts had lexical densities of 53%, 57% and 59% respectively, thus showing an increase the density of information carried by the text and making it more complex and difficult to read. This is evidence of how there is an increasing level of difficultly in school texts and thus a need for students to be able to access the skill needed to access these more complex texts. We can hypothesize that if the lexical density increases, the occurrence of morphologically complex words does as well. This would lead to a greater need for students to possess skills to understand or analyse these morphologically complex words.

2.3.1. A Model of Morphological Processing

One widely supported model of morphological processing is proposed by Schreuder and Baayen (1995) that views morphological processing (in real time) as a matter of decomposition of a morphologically complex word into its constituent morphemes and then activation of concept nodes connected to each morpheme occurs in order to construct meaning. They also hypothesised that frequently processed complex words would develop their own independent nodes to allow quicker access. This means that with sufficient automatization, a morphologically complex word can also be accessed as a whole, without the need for decomposition (although this is also still possible). This dual route of both whole and part processing is supported by a substantial amount of research including some recent work by Goodwin and colleagues). Goodwin et al. (2014; 2013b) focused on the investigation of the roles of lexical representation and root word knowledge with 172 seventh and eighth graders from a suburban middle school in the United States. They found that lexical representations are multidimensional, and readers develop these dimensions in no set order. This work illustrated that a reader might have different levels of knowledge related to word reading, self-reported knowledge of meaning and derived word spelling that can all be correlated. This lends support to the notion that concept nodes can branch into multiple different representations if there is enough exposure to derived words. This web of nodes could themselves have nodes with multiple characteristics or just one characteristic. This has also been supported by work that has examined how readers access morphologically complex words (e.g., Jared & Seidenberg, 1990; Jarmulowicz & Taran, 2007; Taft, 2004). The work by Goodwin et al. (2013; 2014)

additionally found that if a student could read a root word correctly, they were significantly more likely to be able to read and spell related, derived words correctly. A self-reported measure of knowledge also showed a larger positive response about word knowledge if the participants were able to read and spell the root word correctly. This supports the idea that morphological decomposition and analysis are utilized in reading when encountering unknown morphologically complex words and that readers can and do use knowledge of morphemes to determine the meaning of an unknown whole. This supports the intervention developed and used in this thesis, as a starting point for students to intentionally build concept nodes and start to access meaning information carried by morphemes in morphologically complex words to better understand them. While more frequently encountered morphemes might have a stronger network of nodes related to them and thus are a good starting point, adding additional nodes to the network could indeed potentially support a larger network of information for a reader to access.

This same evidence supports work by Reichle & Perfetti (2003), who in their episodic theory of word identification, view each experience with a word as a memory trace that is broken into three parts, meaning, orthography and phonology. The knowledge of these three parts deepens and develops as more encounters with the words occur. This can be reflected in two aspects of availability and familiarity. This is similar to what is proposed by Schreuder and Baayen (1995) who claim that familiarity builds lexical quality and thus increases ease of access or retrieval of the knowledge, although their evidence comes from simulations based on corpora data. This also connects and supports on a more granular level the Lexical Quality Hypothesis (Perfetti & Hart, 2002) which states that high quality lexical representations are needed to support reading comprehension. They define a high-quality lexical representation as having three constituents which are orthography, phonology and semantics. They state that if one of these areas is weak, the result will be a difficulty in comprehension by the reader. The similarity is obvious as the only constituent that is different from the work of Reichle and Perfetti (2003) is semantics which is replaced by meaning in morphology.

The main ideas supported by all of these theories and on a more macro level by the morphological pathways framework is that the more frequent the interaction with a morpheme and or the more encounters with it the stronger the representation and easier to access it will be. This also underpins the idea that starting instruction with more commonly encountered morphemes will help with building additional knowledge, as they will connect to currently established nodes. This will give readers a better

quality of lexical representation and will support their understanding of both monomorphemic and multimorphemic words. Based on this work it would make sense theoretically for both researchers and teachers to build instruction in morpheme meaning into interventions and classroom instruction particularly when working with morphologically complex vocabulary.

2.3.2. Connections between Morphology and Reading

There has been a great deal of research investigating the connections between reading and morphology. If we return to the work discussed earlier about the morphological pathways framework (Levesque et al., 2020), we can see how the current evidence is being incorporated into the frameworks for reading that have been developed. The inclusion of multiple aspects of morphology to the process of reading would indicate that it is an important component. Thus, the ability to use morphological analysis and decoding to aid in the identification of words in text is a critical skill.

These critical skills are additionally included in the models of reading development discussed previously. Both Ehri (Ehri & McCormick, 1998) and Seymour and Duncan (Seymour & Duncan, 2001) explicitly mention the development of the ability to recognize morphemes and to process them as a unit rather than as individual phonemes or letters is an important phase for a reader. It will then take less cognitive effort for the processing of text.

As noted above, some researchers are beginning to suggest that morphology should be to be broken into multiple facets such as morphological knowledge, morphological awareness, morphological analysis. For example, Goodwin and colleagues suggested that the results of their work discussed previously indicate that it should be considered a multidimensional construct with a general knowledge factor and specific factors that can be assessed by different tasks (Goodwin et al., 2017). In this study, the authors tested 371 participants who were in the seventh and eighth grades in the United States and were from relatively affluent areas. The students were assessed based on seven written morphological measures (e.g., suffix choice, generation of morphologically related words, morphological word reading processing, morphological spelling processing and morphological meaning processing) and two standardized reading vocabulary and comprehension measures. Their modelling suggested that there is a general construct but that there are differences in performance that go beyond processing and analysis skills (Goodwin et al.,

2017). Additionally, which of these specific dimensions or facets is being looked at should be carefully examined and stated in research but, this is not the current practice as of yet.

The reading frameworks and earlier described morphological processes combine to demonstrate that there are important significant contributions to reading comprehension and word reading from being able to understand morphemes when they are encountered in text. Starting from the simple view of reading (Hoover & Gough, 1990) as a broad view of reading, followed by the Reading systems framework (Perfetti & Stafura, 2014) and then the morphological pathways framework (Levesque et al., 2020), the models become more specific and provide more detailed views of the reading process with the morphological processing model (Schreuder & Baayen, 1995) being a very detailed examination of a specific part of the process. Together they illustrate that knowledge of morphemes and morphology combine to support the reading process and build knowledge based on exposure to morphemes. These processes and support have been examined utilizing various methods and with different populations over the years. What follows is a brief summary of some of the most relevant studies to the present work.

Many cross-sectional studies in the field are undertaken with students who are L1E. For example, Carlisle (2000) looked at 34 third graders and 26 fifth graders to explore whether there were correlations between the awareness of the structure of words and knowledge of meanings of morphologically complex words. Significant positive correlations were found for grade 3 ($r = 0.46$, $p < 0.01$), and for grade 5 ($r = 0.64$, $p < 0.001$) between the results for a test of morphological structure and a definition task where participants were asked to define a word and use it in a sentence correctly. They also found that the three measures of morphology were significantly positively correlated with reading achievement for the fifth-grade group, while only one measure was significantly positively related to reading comprehension for the third-grade students. This could show an increasing correlation between reading comprehension and morphology as students progress through school. Or it could be additional evidence that supports the idea that morphological analysis is multifaceted and that the facets develop at different times and the third-grade participants had yet to develop those that were specifically measured in this study. It could also indicate that there would be a greater effect of morphological interventions, if the interventions are conducted with

older students. This would be logical if we connect this idea to the dual pathways framework for reading development where the hypothesis is that morphological analysis is one of the final components or phases of reading to develop. It should also be noted that the students who participated in the Carlisle (2000) study were from a private day school in the United States in a middle to upper class suburb. This could also affect the results and the participants could have acquired certain linguistic skills earlier than if this study had been conducted in an urban setting with low SES students. This would be supported based upon the findings from studies reviewed earlier in which gaps in development of vocabulary size were observed which were linked to the size of vocabulary at school entrance.

If we turn to a study with students in grade 3 and 4, we find additional support for effects of morphological awareness on reading comprehension. Deacon, Kieffer and Laroche (Deacon et al., 2014) found morphological awareness had a direct positive effect on reading comprehension as well as an indirect effect through word reading skills for both Grades 3 and 4. They tested 100 L1E students from working class backgrounds in Canada, repeatedly in both third and fourth grades, twelve months apart and as part of a much larger longitudinal study. They suggest that these routes help directly with reading comprehension, and by supporting individual word reading which also benefits reading comprehension (Deacon et al., 2014). One thing to note is that while it does give more support to the claim that morphological awareness positively impacts reading comprehension ability, there was only one measure of morphology used in this study and it looked at both inflectional and derivational morphology. This could have affected any specific skill change if inflectional and derivational morphological awareness develop separately. The measure also used analogy pairs, which could mean that participants would need to know the pattern used to be able to correctly complete the second pair (e.g., luck: lucky:: curl: *curly*, wide: width:: deep: *depth*) (Deacon et al., 2014, p. 451). Seven out of the twenty-one items were inflectional and 14 were stated to examine derivational knowledge. Even with taking all of this into account it is still additional evidence that morphological awareness has a positive impact upon reading comprehension.

Gilbert, Goodwin, Compton and Kerns (2014)(propose that morphological awareness contributes to better reading comprehension particularly for participants with poor multisyllabic word reading. In their

study they looked at 164 L1 fifth graders with a majority receiving free or reduced lunch (a measure of low SES in the US, families' income must be below the poverty line to qualify). The authors found that there was a significant positive relationship between morphological awareness and reading comprehension even when they controlled for general vocabulary and academic knowledge. They used a composite score for morphological awareness consisting of scores from four measures, however three of the measures were suffix choice tasks the other was similar to the analogy measure used by Deacon et al. (2014). While this is a positive that multiple measures were used, it also indicates the lack of measures that have been used to solely examine the meaning carrying function of morphemes. They also suggest that this is possible supporting evidence for Pefetti's Lexical Quality Hypothesis, as participants with better word reading ability were less likely to be dependent on auxiliary information such as roots and suffixes because they already had high quality representations of the multisyllabic words.

Turning to a longitudinal study, Kirby, Deacon and Bowers (2012) followed students from kindergarten to grade 3. They found in grade three that morphological awareness was better at predicting reading comprehension ability than in other grades. They found effects in the range of 3-9%; the most significant being on two measures of reading comprehension. One possible reason they gave for this is, is that "it is possible that the relationship between morphological awareness and reading actually does increase with grade, associated with an increase in reading for meaning." (Kirby et al., 2012, p. 405). This is in line with research by Levesque, Kieffer and Deacon that, "morphological awareness contributed to children's ability to analyse the meaning of unfamiliar derived words, which in turn supported their reading comprehension" (Levesque et al., 2017, p. 16). This claim is based on findings from a study with 221 L1 children in grade 3. Their results also showed that Morphological awareness had a direct positive effect on reading comprehension even when taking into account the effect of the four potential mediators between reading comprehension and morphological awareness: morphological decoding, morphological analysis, word reading, and vocabulary. Additionally, two additional paths of similar strength were encountered: a morphological decoding pathway and a morphological analysis pathway. (Levesque et al., 2017). This is part of the supporting evidence that was used to develop the Morphological Pathways Framework discussed earlier (See section 2.14**)** The authors measured morphological awareness using two measures already described the test of morphological structure from Carlisle (2000) and the word analogy test from Kirby et al. (2012). For morphological analysis they used a measure where students

selected definitions for infrequent morphologically derived words. This had previously been utilized in (Deacon et al., 2017). This continues the recent trend of employing multiple measures of morphology and it further supports the concept that morphology is a multifaceted concept. When Levesque et al. (2017) followed up this study, they found that morphological analysis predicted gains in students' reading comprehension abilities, but morphological awareness did not have the same predictive effect. Overall, they suggest that instruction in morphological analysis would have a strong impact on reading comprehension (Levesque et al., 2018). These findings are supported by the work of Carlisle and Fleming (2003) who found that early elementary lexical analysis of complex words is related to vocabulary and reading comprehension in the later years of elementary schooling (Carlisle & Fleming, 2003). This could be an example of the Matthew effect, where students who find reading easier because they can comprehend what they are reading easily are more likely to read more and thus be exposed to a wider variety of words and build larger and more varied vocabulary because of this.

Overall, these studies support the view that morphological awareness has a positive impact on reading comprehension in the primary grades and that the impact of morphological awareness seems to increase as students develop their reading skills. This is additionally reflected in the work by Seymour & Duncan (2001) who theorised that as reading skills develop there is a shift from readers reading individual letters to processing morphemes and even whole words. The idea that students who are not as skilled at morphological awareness are also more likely to struggle with reading comprehension has been raised and there is some initial support for the idea (e.g., Zhang & Shulley, 2017). However, while most of these studies use the term morphological awareness, caution should be used here, as greater understanding of morphological awareness leads us to believe it is a multidimensional construct. This takes us to the next point: that there are different concepts contained in knowledge.

Thus far it has been shown that there are significant connections between reading comprehension and morphological awareness with L1 students, particularly as they develop as readers and shift into the consolidated alphabetic or morphographic literacy phases. We have seen that morphological awareness in lower grades is a predictor of reading comprehension in higher grades, though morphological analysis is a better predictor when awareness is split into more specific factors. We have also seen that while this relationship starts in lower primary grades it continues and increases in upper grades. However, we also

need to consider whether similar findings occur when participants are learning English as an additional language.

Zhang and Shulley (2017) looked at 81 fourth and fifth grade students who were both poor comprehenders and typical readers; both groups were composed of students who are NS and EAL from the United States. They analysed a smaller sample from a larger study on measures of working and short-term memory, incidental word learning vocabulary and morphological awareness. They then grouped participants as poor comprehenders if they scored above the 40th percentile for word reading but below the 25th percentile on reading comprehension measures. The participants identified as typical readers had scores on those measures between the 40th and 84th percentiles. This then gave them a group of poor comprehenders that consisted of 21 EAL and 16 NS and typical reader group that consisted of 25 EAL and 19 NS participants. The morphological awareness measures comprised a set of three measures including: suffix choice, compound production and extract the base. Typical readers performed significantly better on all three morphological awareness measures than the group of poor comprehenders. They found that students who were poor comprehenders regardless of language background were less able to use morphological analysis in order to infer unknown meanings while reading than the typical readers when controlling for morphological awareness and vocabulary knowledge. Another interesting finding, they had was that the same patterns were found in the results for both language groups. Since this was not an intervention study, they suggest a fruitful area of research might be investigating whether poor comprehenders who are both English only and English Language Learners (ELL) would both benefit from explicit instruction in morphological analysis as a word learning strategy. Similar suggestions are made in other (e.g., Carlisle, 1988; Goodwin et al., 2014; Levesque et al., 2018; White, Power, et al., 1989). As the current study used explicit instruction in morphological analysis for learning vocabulary, it builds directly upon these suggestions that morphological analysis is a key skill in improving reading comprehension.

In a study of 155 Spanish L1 speaking fourth and fifth graders, Goodwin and colleagues (Goodwin, Huggins, et al., 2013) found that both morphological analysis and morphological awareness significantly predicted reading comprehension and when morphological awareness increases, so does reading comprehension ability. They also hypothesized that meaning was more important for reading comprehension than decoding in the older children. These results are supported by those of another study,

with a group of EAL and L1 English speakers of 137 students (55 self-reported speaking only English at home) which found significantly lower levels of morphological awareness with students who were Language Minority (LM) learners (Kieffer & Box, 2013). The authors suggest that morphological awareness might be a source of difficulty with academic vocabulary acquisition and reading comprehension, which aligns with earlier findings that strong morphological awareness makes a significant increase to reading vocabulary which in turn makes a substantial impact on reading comprehension ability (Kieffer & Lesaux, 2012a).

If we look at studies conducted with native speakers of languages other than English, a study in South Korea looked at 188 10th grade students who had Korean as an L1 and had studied English for 7 years 7 months on average and found "that morphology is not only a variable that contributes to reading comprehension independent of phonology but also a powerful contributor to L2 reading comprehension" (Jeon, 2011, p. 230). This is important because if students can transfer the skill of morphology from their L1 to English this might be a way to aid students who are attempting to become fluent in English with accessing texts. Since additional studies or replication studies have not been carried out and it was a small study of only 188 students this idea should be used very cautiously. What this does show however is that students for whom English is an additional language develop morphological awareness in English and that it contributes significantly to reading comprehension. This means that for students who speak a language other than English in their homes but are schooled in English there is still a potential substantial impact of morphology on their reading comprehension. Thus, if an effective method for teaching morphology were determined, it could positively impact EAL students' reading comprehension. Additionally, if a method could be devised so that it could be implemented earlier it would mean that the differences between EAL and L1E students could potentially be narrowed earlier and critically, before gaps between the groups grow even wider. To do so, finding the proper method and the right age at which to intervene is also important. With a number of studies and theories indicating that around year 3 is a critical time, this is a practical reason behind why the present work was conducted with year three students.

2.3.3. A Review of Interventions using Morphology

The previously described studies clearly indicate that Morphological Awareness and Morphological Analysis have a significant influence on reading comprehension. This should entail that improvements in morphological knowledge will lead to improvements in reading comprehension ability. With this in mind we shall now turn to studies that have assessed the efficacy of teaching morphology to primary level students, although they have approached it from multiple perspectives.

The majority of studies on the topic of the teaching of morphology have been concerned with suffixes and students' awareness of the purposes of suffixes used in texts that would be encountered in American primary schools. One such study was undertaken by Wysocki and Jenkins (1987) with 131 students in fourth sixth and eighth grades (9-12 years old) in California. It was a very short intervention focused on twelve morphologically related word pairs (e.g., sapient: sapience, abatement: abate), lasting only three sessions of 15-20 minutes each. The intervention involved students' reading in unison both the target vocabulary and definitions aloud. Then the researcher erased the definitions and left only target vocabulary and asked the students to supply the definitions. They then reversed this process with the researcher giving the definition and students producing the target word. Different sentences were used each day. Students were tested two weeks after the intervention was completed. The older students performed better and "were more skilled in the use of context clues and in the use of morphological clues," than the younger students (Wysocki & Jenkins, 1987 p.78), although the authors also stated that the levels of morphological generalization were impressive for all students. This might be seen as being predictable because the older students would have had more exposure to morphologically complex words as they would have had more opportunities to encounter them and thus more experience with handling morphologically related words. This would also connect with reading development theory: as the learners progress, they have more practice with the skills which are also involved in later phases. An interesting characteristic of this study is that they used two types of scoring including one that was strict and in which participants needed to correctly identify the meaning and syntax. The other type of scoring was a more lenient system, which gave credit if the definition given indicated that the student knew the general meaning of the transfer word. The scores were much higher when the lenient scoring system was used as almost 70% of the answers were the meaning for the transfer word. Another interesting characteristic of this study was there was no control group. Instead of this, the authors constructed a list of control words

that were not taught to some of the groups but were taught to others. This makes it difficult to interpret how much of the observed improvement was caused by learning purely through the intervention and how much may have been affected by the instruction that was given. What is being taken from this study is that generalizable meanings are easier for students to access than specific ones, and that even very short interventions can be effective. These two factors will come into play with the design of the morphology measures used and in the design of the intervention.

A large study on morphology teaching with 230 ninth grade (14 years old) students in an urban setting the United States with 46% of students at the participating school received free or reduced lunches, looked at two different methods of teaching vocabulary to students with disabilities and students without disabilities (Harris et al., 2011). Students with disabilities were classified as such based on active IEPs which included accommodations for learning disabilities, emotional disability, mental retardation, other health impairment and autism spectrum disorders. The methods employed were a word mapping strategy (MAPS)[1]; using a graphic organizer to break down morphologically complex words and map the morphemes' meanings, and the Vocabulary strategy (LINCS)[2]; a set of steps to aid students in learning and recalling vocabulary words. Both groups received 10 lessons of 45 minutes in length delivered in their regular classrooms by the researcher. There was also a control group that was only pre- and post-tested. Both the word mapping and LINCS groups made significant and comparable gains on the strategy use tests, though the group that received the word mapping lessons made significantly larger gains on the morphological analysis test than either the control group or the group that received the LINCS training. There were also no significant differences between the students with disabilities and the students without disabilities. In an intervention study by Apel & Diehm (2014), the authors found that even students as young as kindergarten to second grade (5-7 years old) could make statistically significant and practical gains on morphology measures after an 8-week intervention. The students were seen four times a week in groups of four or five students away from their regular classroom. The study had 151 participants from 6 kindergarten, 5 first grade and 5 second grade classrooms including a randomly assigned "business as

[1] The MAPS strategy involved four steps: 1) breaking words into their morphemic parts, 2) attaching meaning to each part, 3) make a prediction about the meaning of the unknown word based on each part, 4) check a dictionary for the definition.(Harris et al., 2011)

[2] The LINCS Strategy involves 5 steps, ***L*ist** the parts, ***I*dentify** a reminding word, ***N*ote** a Linking story, ***C*reate** a Linking picture, ***S*elf-test** and practice. (Ellis, 1992)

usual control", while they were from a low SES (Social Economic Status) area, they all spoke English as their first language. The intervention covered both common prefixes and suffixes (e.g., plural -s, past tense -ed, prefix un-, prefix re-, suffix -er(person), suffixes -ness, prefix dis-) and had review lessons built in at regular intervals (after a pair of lessons there would be a review session). The sessions contained a variety of activities such as word sorts, circled words containing the target affix from a list of words, raised a magnifying glass when a word with the affix was read. Outcomes were measured by looking at the effect sizes for Rehit and Relatives measures, which are based on common measures in morphology. The Rehit measure was used to determine ability to combine two morphemes into a novel word and define that word, as noted by the authors: "[f]irst, the student was asked to repeat one bound and one free morpheme (e.g., "Say re." [student responds] "Now, say hit." [student responds]) and then combine the two morphemes to create a "silly" word (i.e., "rehit")." (Apel & Diehm, 2014, p. 67). The Relatives measure is based on a measure repeatedly used by Carlisle (1995); a Test of Morphological Structure (TMS), where participants complete sentences using a morphologically different version of a suggested word (e.g. Driver. Children are too young to____ (drive).) (Carlisle, 1995; Carlisle & Fleming, 2003; Carlisle, 2000). Effect sizes of the difference between control and intervention groups on these measures ranged from d= 0.41 to 1.26. One possible issue with these measures is that they only ask participants to manipulate suffixes, there are no questions that ask students to change or remove prefixes. While such tasks have been widely used and seem to be reliable to measure participant ability to recognize and manipulate suffixes, they should not be generalized to all morphological ability or the ability to manipulate prefixes as well.

While suffixes have relevance to the current research, prefixes are less well researched and will be a focus for the current study. White, Sowell and Yanagihara (1989) is one of the oldest and widely cited studies of prefixes. These researchers advocated for "actively *using* word-part clues to derive all or part of the meaning of unfamiliar words" (White, Sowell, & Yanagihara, 1989 p.302). The study they referred to, as part of their support for this method was a 7–8-week intervention they conducted with forty- seven third grade students (8-9 years old). The intervention included lessons on what is a prefix, and then a series of lessons focused on major prefixes (e.g. un-; dis-; in-; im-; ir-; non-; re-; en-; em-; over-; mis). During the lessons students were guided to "peel-off" the prefix and identify the meanings of both the

remaining word and the "peeled off" prefix. Lessons also included examples of prefixes that do not always mean what one would assume (e.g., Disarm: does not mean something that is not an arm). Four measures were used to measure the outcome of the intervention: identification of root words, prefix meanings, transfer test 1- applying prefix knowledge to unfamiliar words and transfer test 2 - defining ten unfamiliar prefixed words in context. At the end of the intervention, the intervention group had scores that were significantly higher than the control group on all four measures. The scores of the instructed group were 71%, 84%, 82%, and 23% respectively while the control group scores were 53%, 43% 54% and 6% on the same measures. This demonstrates that there can be a significant impact on students' use of morphology for understanding new words if they have been instructed on how to do so. This same concept has been reinforced by the results of some research undertaken by Baumann et al (2002) who found that students who were given instruction in morphemic analysis were able to outperform students who were not given instruction on recognition transfer word measures (Baumann, Edwards, Font, Tereshinski, Edward, & Olejnik, 2002). This study was also concerned with context and focused on ninety-two students in the fifth grade (10-11 years old) in the United States. The students were split into four groups: morphemic -only instruction, context only instruction, combined morphemic- context instruction and, an instructed control which had no explicit instruction in vocabulary strategies but who discussed words as they came up from reading a grade level appropriate book. Students in the two morphology instruction groups were also able to outperform students who were given only context instruction or no instruction in almost every measure taken. The lesson plan contained in the article as an example of how the lessons were conducted is very teacher focused. The work with prefixes illustrated, has the teacher defining and then leading students through a guided practice and finally the students defining in written form additional words and the original prefixes. One of the main differences between this intervention and the one designed for the research completed for this dissertation is that the one for this dissertation is much more student focused and led. Also, earlier research has mainly been completed with either very young early primary aged participants or much older secondary school participants. The middle to late primary years seems to be avoided by researchers. This is one of the gaps that the intervention in the current study attempts to fill.

Baumann et al., (2003) looked at teaching social studies vocabulary by comparing the teaching of content specific words and the teaching of morphemic analysis strategies. This study comprised of 157 students

from a mid-sized US city, who were drawn from 8 classrooms at 4 different schools with a range of free and reduced lunch of 42-65%. They were also diverse classrooms with a mix of African American, Asian/Pacific Islander, Hispanic, and European American students. Students who received the morphemic analysis treatment scored on average 9.27 points higher on a word part test than the students who did not undergo this treatment. They also scored an average of 1.84 points higher on a delayed vocabulary in context test which was statistically significant. In their summary, the authors state "instruction in morphology did aid in students being able to unlock meanings of morphologically decipherable transfer vocabulary better than textbook vocabulary instruction," (Baumann et al., 2003 p. 481). A unique aspect of this study is they integrated the research materials with the curriculum that was already in practice. That is, the authors scripted lessons that employed the textbook already in use in the classroom and simply extended content lessons from the text to include morphological instruction that employed vocabulary already present. The study included 25 days of instruction and an additional eight days of pre-tests, post-tests, and chapter tests. Sessions took 45 minutes that included an extension that was the focus of the intervention. The extension was a 15-minute vocabulary lesson that focused on an affix family while the control group received a 15-minute lesson on vocabulary from the textbook where they predicted meanings and then looked them up in the glossary. The focus on whole words or morphemes and their meaning is appropriate for students who are likely in the later stages of reading development and have progressed to the morphographic phase (Seymour & Duncan, 2001) and are using or developing their use of morphological analysis to understand words that would fall into Tier 2 where they are high utility and found in a variety of domains (Beck et al., 2013, p. 9). Because this study was embedded into the existing curriculum and utilized the classroom teachers for delivery, it would appear to be very ecologically valid and adoptable as it would not greatly affect the already established classroom practices.

In a more recent, smaller-scale study (Kirk & Gillon, 2009), the researchers attempted to combine morphological awareness with other types of linguistic skills to improve reading and spelling. In the intervention, they used word sort tasks for both morphologically simple and complex words, picture identification of vowel length, prompted spelling, and they increased the complexity of the word sort task over time. The participants were sixteen children aged between 8 and 11 years old from New Zealand. Students showed improvement in both spelling and reading for both taught and untaught words, showing that the taught skills were generalizable to new words. Additionally, the improvement was maintained

when it was measured six months after the intervention. Some caution should be taken when looking at the results of this study for a couple of reasons. First all, the participants were recruited from a high SES school and were Caucasian. Additionally, they were recruited on the basis of their teachers' determination that they had poor spelling skills but were in the normal range of intelligence. Thus, it may be the case that such an intervention is effective only for this specific type of population. Also, the intervention sessions and testing all occurred individually, so this intervention method is arguably not as ecologically valid as interventions that occur in groups, or with whole classes of students.

Turning now to research involving multilingual groups, Goodwin has carried out a number of studies on the effects of morphological instruction on reading comprehension. In 2016, she studied the impact of a short intervention comprising 4 lessons of a half hour each to small groups of 2-7 students (Goodwin, 2016). The students were in fifth and sixth grades (10-11 years old) and were a mix of 129 native speakers of English (NS), 28 English Language Learners (ELLs) and 46 Language Minority Youth (LMY), and the groups were relatively equally divided between the intervention and comparison groups. The intervention consisted of six steps for each session. The steps were scripted and started with tutors teaching students about comprehension strategies, with the students then applying the strategy. Following this, the main points were displayed on a white board. Then a morphological problem-solving activity for the intervention or defining words for the comparison condition was undertaken. Next, the tutors closed the lesson and finally a reinforcement game was played, which varied between the groups. The morphological problem-solving groups played games that reinforced morphological problem solving whereas the comparison group played a game that reinforced the comprehension strategies that had been taught. The participants were pre- and post-tested using eight different measures: two vocabulary measures, two reading comprehension measures, two measures of word reading fluency and two measures of morphological awareness that were a mix of researcher made and assessments that had previously been used in other research studies. One of the morpheme measures was designed to look at general morpheme awareness and had students answer if they thought a new morphologically complex word came from a base morpheme (e.g., Does magician come from magic?). The second morpheme measure looked at morphologically related words where students were asked to produce in writing as many morphologically related words as possible for a target word (i.e., big: bigger, biggest). This second measure seems like it would be more difficult as it required participants to produce new words not just determine if they are

plausible relatives. This may be the case, given the work discussed in previous sections that has reported that productive vocabulary tends to develop after receptive knowledge. If so, there may have been some gain in knowledge that was not able to be measured, as it had not yet developed to the point of participants' being able to produce it. The LMY and NS showed significant increases in their ability to generate morphologically related words in comparison to the control groups. The findings add to the evidence that instruction in morphological analysis leads to moderate to large effect sizes in comparison to control conditions. The authors also noted that the below average performers on the pre-test, behaved on the post-test akin to the comparison group on the post-test (i.e., average like above average, below average like average). This potentially illustrates that even a short intervention has the ability to narrow or close the gap between groups. If this is indeed possible it could mean that short and effective interventions could be used to aid students who are struggling due to their having smaller vocabulary size or less exposure to English outside the classroom.

Kieffer and Lesaux have also carried out multiple studies of morphology in second language learners. For example, in a study following 90 Spanish speaking LM learners, they found parallel growth in both, English vocabulary and morphological awareness and a strong positive correlation between the two (r=.67) (Lesaux & Kieffer, 2012). The authors later undertook and evaluated an 18-week intervention with 349 LM learners and 133 native speakers in grade six (Age 11-12). The intervention, named Academic Language Instruction for All Students (ALIAS), taught eight or nine target words each unit which lasted two weeks. There were specific targeted lessons that focused on explicit instruction in morphology. The intervention spent 7.5 hours out of the total 54 hours of instructional time on just derivational suffixes. (i.e., –tion; –sion; –ify; –er; –or; –ist; –al; –ical; –ous; –ity; –ness; –ful; –ly) and on relational and syntactic aspects of morphology. They found positive effects when compared with standard classroom practice. Moreover, when compared with data from the previous study, students in the intervention group showed an additional four months of vocabulary growth (Kieffer & Lesaux, 2012a). While this is significant, the amount of time necessary to carry out this intervention was substantial: 45minute sessions, 4 days a week comprising eight two-week units and two review weeks. Indeed, any of the studies with significant effects were carried out over large periods of time or were very intense in the number of sessions or the duration of the sessions with relation to a typical school (e.g., Baumann et al., 2002; Nunes & Bryant, 2006). The amount of time dedicated to some of these interventions is enormous.

A large amount of evidence is going to be necessary to convince teachers to commit this much classroom time to instruction in morphology. While these interventions used materials that could also be covered as part of the curriculum, an intervention that is short and which could fit easily into a small period of time would likely be more adoptable by teachers.

Several meta-analyses of morphology instruction have been completed. One of the earliest syntheses of morphological instruction studies was completed by Reed (2008), who found that there were medium sized effects (.60 to .79) on reading accuracy and efficiency after interventions. She focused on studies published between 1986 and 2006 with students from kindergarten to 12th grade. She found only seven studies that met this criterion and of them, three focused specifically on vocabulary acquisition, three focused on word identification and one was focused on spelling. She also noted that many of the measures were researcher made and not measuring the same facet of morphology. On the basis of the meta-analysis, Reed makes the overall suggestion that that there is a need for more research to be done before conclusive interpretations can be drawn. Later, Carlisle et. al. (Carlisle et al., 2010) carried out a meta-analysis of sixteen studies with instruction in morphological awareness or other related terms with both pre- and post- intervention measures and at least one measure of literacy. The studies included students from kindergarten to high school and at least one control group. These studies were published between 1989 and 2009, five were focused on morphological awareness and phonological awareness, seven were focused on Morphological awareness and orthographic development, only four looked at the effects of instruction in morphological analysis to learn meanings of unfamiliar words. These final four are of the most interest as they are the closest to the goals of this study. Two of them have been discussed in depth already in this review (e.g., Baumann et al., 2002, 2003). In all but one of these four studies that compared morphological awareness instruction with a control group, there were significant effects on multiple measures of morphological awareness. The one exception was close to significant (Carlisle et al., 2010). Bowers, Kirby, and Deacon (2010) had similar findings. They examined 22 studies from preschool through to grade 8 and found that morphological instruction has its best effects at a sub lexical morphological level, indicating that morphological instruction is successful at improving morphological abilities when compared to alternative and control instruction.

Goodwin & Ahn (2013) also carried out a meta-analysis of morphological intervention studies. After synthesizing 30 different studies, a moderate overall effect of morphological instruction was found which suggests that students who received instruction in morphological analysis perform better than comparison groups on measures of literacy. When looking specifically at measures of vocabulary they found the effects for vocabulary establish the link between vocabulary knowledge and morphological knowledge and advocate for instruction in morphological analysis as a means to aid students in using the meanings of morphemes to find the meaning of a morphologically related word. The meta-analysis also found that knowledge of affixes and roots aids students in their ability to access word meanings (Goodwin & Ahn, 2013, p. 278).

In a random controlled trial with 717 second graders in Norway von Koss Torkildsen et al, (2021) with 366 in the morphological training group, found that with an 8 week app based morphology program resulted in "immediate and long-term improvements in children's meaning-based and code-based knowledge of multimorphemic words that were exposed during training"(von Koss Torkildsen et al., 2021, p. 850)(von Koss Torkildsen et al., 2021, p. 850)They also suggested that the results from their work could support instruction in derivational affixes and compounding patterns might not need to wait until later primary grades as has previously been suggested.

This section has attempted to provide a brief summary of the literature relating to morphological intervention studies. So far it has been shown that there is a significant positive effect on students' understanding of vocabulary when they have received instruction in morphological analysis. The age of the student seems to not have a large impact, although older students do seem to be able to apply and utilize the information more successfully than younger students. Looking back to the discussion of the literature on morphology and its connections to reading in earlier sections, it was found that this is being explored as a source of access to vocabulary and thus in turn, to better comprehension. There does not seem to be a consensus on the ideal length or intensity of the intervention needed to make an impact, with intervention lengths spanning from 18 weeks (Kieffer & Lesaux, 2012a) eight weeks (Baumann et al., 2003) or even four sessions totalling two hours(Goodwin, 2016). Interventions can be stand-alone short activities, or they take up significant periods of class time or be embedded in content lessons with only a short period being dedicated to specific morphology instruction. Many of the interventions reported on

above have focused on teacher led interventions with limited student input and focus. The intervention designed for this dissertation is based some of these same ideas while being student focused and led. We will now turn to studies of Greek and Latin root interventions and their connections to morphological analysis and reading comprehension as they have shown another possible path for increasing vocabulary size and improving reading comprehension.

2.3.4. Why Teach Greek and Latin Roots explicitly

There is a distinctive issue when it comes to examining roots. While it has been established that a large number of words in English come from Greek or Latin roots, and even more when Academic vocabulary is examined, there has been very little research into the instruction of these roots. One of the unique characteristics of the study that was conducted for this dissertation was the specific and planned inclusion of Greek and Latin roots. To examine these issues, this section will discuss the research into the numbers of Greek and Latin roots present in English and Academic Vocabulary and the few studies that have been published investigating the instruction of Greek and Latin roots.

It has been estimated that around 60% of all words in the English language have Greek or Latin roots, this rises to 90% when looking at scientific or technology lexis (Green, 2014). Additionally, a very significant percentage of words that are used in academic settings come from Latin roots. When Lubliner and Hiebert (2011) examined the Academic Word List constructed by Coxhead (2000), they found that 75% of head words were bound roots of Latin origin. This means that an overwhelming majority of words used in higher education have Latin roots and that if students do not understand them, they are effectively barred from accessing study at a higher level.

To date, few studies have looked at the effects of instruction in Latin or Greek roots, but those that have, have found positive effects. A preliminary study by Crosson and McKeown (2016) found that after relatively minimal instruction (11 lessons over two school years), students were more likely to apply a direct route to analysing an untaught but morphologically related word than students in the control group. A ‘direct route’ in this study entailed that “students applied bound root to infer meaning of unfamiliar words” (Crosson & McKeown, 2016 p.159). The direct route was applied 71.7% and 65.38% of the time in two instructed groups verses 55.77% of the time for the control group. This study was done with 59 NS

in the United States. The students were tested for reading comprehension before the intervention and there was a sample of students from low- middle and middle to high reading comprehension. The second year of the study also comprised students from an honours class, and they were measured separately. In a study of 84 EL students from grades 6-12 (11-18 years old), a ten-week program of Latin roots instruction, when compared with the Robust Academic Vocabulary Encounters (RAVE) program, the Latin roots condition showed large positive treatment effects for morphological analysis to problem solve the meanings of unfamiliar words. Latin roots showed an advantage for lexical access, as lexical decisions were more accurate for the target words after Latin roots instruction (Crosson et al., 2018). With these two studies, Crosson and colleagues have started to establish that if students are given guided instruction in Latin roots and how to decompose morphologically complex words to find them, students are able to apply known meanings and thus are able to understand unknown or unfamiliar words in context. As these studies are relatively recent, there still needs to be more research in this area to build support for this idea. Additional research or application of similar interventions with primary grade students might also open up new areas in order to bridge the gap earlier between students with smaller vocabularies who then struggle with reading comprehension if this research correlates with the work that has been done with morphology. As roots are a part of morphology, this should be a logical connection. This could also help students who are not L1 speakers of English, giving them an additional tool that could help them to comprehend higher level texts by giving access to what is typically viewed as a more prestigious word form.

2.3.5. Summary of Studies of Morphology

This section has illustrated that there are multiple key areas of morphology, morphology intervention that have been researched and found to have had an impact on vocabulary and reading comprehension. However, a method for instruction has yet to be determined. The section began with a discussion of the large number of words in school texts that are analysable once affixes are known (White, Power, et al., 1989). We then discussed different models of morphological processing and how they fit into the larger frameworks of reading comprehension and development with a focus on how they connect as skills develop and shifts occur as vocabulary develops and more morphemes are recognized as carrying meaning to a reader. This area of detail is still evolving but the recent progress aligns with the research that has been conducted.

Next, we progressed to the development of multiple aspects of morphology such as knowledge, awareness and analysis and how they are split and are yet overlapping in some respects. The development of these different aspects illustrates that there are multiple dimensions that should be considered when investigating morphological skills and knowledge. This would indicate that there should be both multiple manners of assessment and instruction if these multiple facets are eventually found to impact equally upon reading or vocabulary knowledge.

Additionally, there has been consideration paid to the strength of correlations between morphological analysis and awareness with reading as students progress for both EAL and L1E students. This is important, as the need for instructional methods that can be applied for both populations is necessary with the increase in EAL students in many classrooms. The lack of any impact of language background is unique and might need more work as it has not been found consistently. Part of why this is important is connected to what was seen when work with vocabulary was discussed because many EAL students arrive at school with smaller vocabularies and potentially less exposure to English. This lack of exposure means that they start with smaller vocabularies and are required to catch up at the same time L1E students have had more exposure. This seems to have a compounding effect on vocabulary growth (via the Matthew effects) where it is easier for students who have larger vocabularies to read more, so they do, thus increase their vocabulary and reading skills at a more rapid pace. This would logically entail that the earlier gains can be made for students who start at a disadvantage, the better. This leads to the need to develop interventions that could be used in a variety of settings and for a variety of skills levels.

Thus far, interventions have found that generalizable meanings are stronger for students in post-test measures (Wysocki & Jenkins, 1987). Word mapping has a greater impact on morphological analysis than general vocabulary strategies for determining meanings (Harris et al., 2011). The length of interventions does not seem to be a factor if there is an impact on morphological analysis (e.g., 18 weeks (Lesaux & Kieffer, 2012)8 weeks: Baumann, Edwards, Font, Tereshinski, Edward, & Olejnik, 2002; or 4 sessions over two hours: Goodwin, 2016). A large majority of the work completed has developed interventions that are very teacher focused, whereas current trends in primary education is for either balanced instruction or student focused instruction, where students drive the instruction via investigating and discovering knowledge for themselves. Finally, interventions can be successful when embedded in

existing curricula and other lessons, meaning that they do not need to be stand-alone (Baumann et al., 2003).

2.4. Summary of Findings from Literature

Collectively, the work examined in this chapter has shown that there is evidence of influence on both reading and vocabulary by morphology. The exact nature and timeline of the impact is still under investigation, but overall, the findings suggest that the understanding of words increases for students once their ability to recognize morphemes as carrying meaning has developed. This appears to occur in the mid to later primary school years, with normally developing readers. This is based the findings of research reported above which support a combination of theories of reading development, reading comprehension frameworks and the morphological processing model. These together illustrate the process at different levels of detail and from different perspectives. The optimum age of this development has not been exactly defined but it should be noted that it has been observed to develop around year 3 of formal schooling. This would most likely be the earliest point in which an intervention or instruction in morphology could make an impact as students in earlier phases would likely struggle and not have developed the skills or knowledge necessary to utilize this knowledge appropriately. This would mean that the positioning of the current study to align with this development is important and slightly different from the majority of studies that have been aimed at either younger groups (i.e., kindergarten to second grade) or significantly older groups (i.e., grade 5 through 8). The choice to position this study in this understudied area is intentional as it could add to the information available about when these skills develop. Furthermore, if this is a potential earlier period to implement work with morphological analysis into classrooms it can aid students who might be struggling with reading and vocabulary issues at an earlier point in time.

Secondly, taken together, the work presented and discussed on vocabulary growth demonstrates that there remain differences in vocabulary size between both EAL and L1E students and Higher and lower SES groups and no conclusive method for equalizing them. The current literature suggests that multiple exposures and use of vocabulary in various types of contexts is one of the strongest methods for increasing vocabulary size, but it has not yet been used successfully to bridge the distance between students starting with low versus high levels of vocabulary. The current trend in instruction is to focus on

words that have opportunities for both high utility and repeated exposure for students. Applying this trend for selection of both affixes and roots will add to the current body of research that demonstrates that this is an area that has a potential for growth with students.

Thirdly, the work that has been completed with interventions has indicated that there are certain activities that seem to promote the development of morphological analysis by students. These conclusions are drawn from work both into morphology and vocabulary development and will be combined in development of the intervention implemented in the present study. They include the use of word mapping to deconstruct and apply meaning to individual morphemes which makes the meaning of separate morphemes explicitly clear as they are both broken down and recombined to indicate the meaning of the morphologically complex word. This potentially could develop the skills with students to decompose morphologically complex words and then even to partially identify unknown words so that they could develop a general sense or meaning of an unknown word. Even this partial knowledge could be beneficial to students who are struggling to understand texts.

Fourthly, the selection of measures for this work is a combination of tried and tested measures that have been utilized in similar work over time. These measures (e.g., YARC and BPVS3) have been reliably used to measure reading comprehension and receptive vocabulary respectively. The use of the YARC was as a productive measure, as the comprehension questions do not give a selection of answers but the ideas must come from the participants' understanding of the piece of writing they have read aloud. The use of a receptive measure in the BPVS is as discussed in the section on receptive and productive knowledge based on the idea that receptive vocabulary develops sooner and as this is a short study, it is more likely to change than productive vocabulary.

Finally, with what is known about reading and morphology, the ability to connect these ideas in a way that is accessible and potentially easily applicable to current curricula would be useful for teachers especially if it can help the students who may not be on the same level as other students in the class due to their starting with a smaller vocabulary or just less exposure to English. Or it may help students who are on level to progress further so that the instruction can be incorporated with the current work already being undertaken in the classroom. While many researchers have deployed intervention studies, the interventions themselves seem to take up large portions of class time, and with the increasing demands for

teachers to use instructional time wisely to increase scores on government mandated exams programs, having interventions that are short and easily implemented would be easier and more attractive for teachers to adopt and utilize.

3. Methodology

3.1. Research Questions:

The main focus of this research was to explore the impact of morphological instruction on reading comprehension, morphological awareness, and vocabulary in primary school children. An additional focus explored whether the impact was affected by the child's language background i.e., if the child is a monolingual English speaker or is learning English as an additional language. This thesis addresses two research questions:

1. Is there a difference between EAL and L1 children in the acquisition of the meanings of derivational morphemes that are taught as part of a small group intervention. Is there an effect of individual differences in gender, attendance and age in the effectiveness of teaching the meaning of derivational morphemes? Is there a relationship between reading ability and knowledge of derivational morphology and does this relationship differ between EAL and L1 English children?
2.

What is the effect of measures of reading ability on derivational morphology knowledge? In order to look at these questions several steps were taken. First an intervention that could be used by a classroom teacher was designed. Using work previously conducted by Beck, McKeown and Kucan(Beck et al., 2013), and Bimiller and Boote(Biemiller & Boote, 2006), combined with Baumann(Baumann, Edwards, Font, Tereshinski, Edward, Olejnik, et al., 2002). Second, tests were designed to measure the participants' knowledge of meanings for both affixes and roots, focusing on derivational morphology. Thirdly, standardized measures of vocabulary and reading comprehension were selected. A pilot study using both the newly created measures and a shortened version of the intervention was undertaken. This was due to a request by the pilot school to fit more easily with the school calendar and cause fewer interruptions. As the focus of the pilot was to test the measures and to ensure that the intervention was feasible with the target age group and the timing of the sessions, the shortened version was acceptable. After modifications, the main study started in the autumn of 2019, running over eight months of the 2019-2020 school year. The data were then analysed looking at whole group effects, followed by an investigation of subgroup (L1 and EAL) effects. To begin this chapter, we will look at the design of the intervention and measures, followed by the pilot and main study.

3.2. Organization of the Intervention and How It Was Implemented

The goal of the intervention was to instruct students in morphological concepts and allow them to build their semantic knowledge of morphology. The intervention was loosely based on a themed spiral curriculum (Figure 3: Intervention Plan). Activities were chosen that were deemed to be familiar and accessible to the average primary level teacher. By using activities that teachers may have seen before or used in a slightly different way, it was anticipated that adoption by teachers in the future would be easier. Some of the activities were adapted from interventions that had demonstrated some impact on students' growth in morphological analysis abilities (i.e., word mapping from the MAPS intervention, c.f., (Harris et al., 2011). Additionally, some of the activities were drawn from current practices or were based on current theories of reading comprehension and vocabulary (i.e., Word Tiers from Beck et al., 2013). The decision to focus on morphological analysis with regards to learning the sematic meanings of morphemes is multi-layered. The theoretical frameworks and some of research described in the previous chapter have indicated that as reading skills develop, there is a shift from focus on individual letters or phonemes to morphemes and whole word meanings being accessed. This is at the centre of this intervention and supports the idea that if improvement can be made in either or both knowledge of morphemes and the skill of morphological decomposition, a greater number of word meanings would be accessible to students. This would in turn assist students with being able to read and access the meaning of a larger number of words. To this end, the intervention was designed with the intention to focus on student led discovery of morpheme meanings and then utilizing this new knowledge in various activities.

The intervention lasts ten weeks and each week has a theme such as numbers, science, position, or communication. The lesson cycle lasts a fortnight, with each cycle containing two themes and a total of four lessons. This was modified from the pilot which had six lessons each cycle (see Figure 9). Having established that the majority of research has been on inflectional morphology, the focus of this intervention is on the less explored territory of derivational morphology. Morphemes were selected from lists of the most common prefixes, suffixes and roots in English, using resources that would be easily accessible to the average primary level teacher (e.g., Marzanno and Sims, 2011; Ganske, 2008). The exact detail of which affixes were drawn from which sources is given in appendix 12.9. The majority were taken from an article by Manyak, Baumann, and Manyak (2018). In this work they synthesize a significant amount of the work that had been completed with regards to instruction of

morphemes. They also drew on research from other scholars compiling high frequency affixes and empirical lists of most frequently occurring affixes. The work included two lists: one of affixes for instruction in Grades 3,4 and 5, and a list of Latin and Greek word roots for instruction at Grades 4 and 5. As the ages of participants in this study are approximately the same as students in Grades 3 and 4, the lists for these grades were used when selecting which affixes and roots would be taught. Other work that was utilized to a lesser degree when selecting which morphemes would be instructed were Marzano and Sims (2011) on the instruction of vocabulary in the common core curriculum, and White, Sowell, Yangihara (1989). All three of these sources used research conducted into word families and occurrences of words in primary texts to determine which affixes and roots were most frequently occurring. It is acknowledged that it might have been more rigorous to only choose affixes and morphemes that occurred most frequently and solely focus on lists that had been compiled as the most frequent. The use of a blend of sources that have all approached this work from slightly different angles resulted in a list that did not give rise to issues such as the intervention covering a morpheme that was also being taught at the same time by a classroom teacher and thus it not being able to determine which was effective the classroom teacher's instruction or the intervention. They were also checked to ensure some overlap with both the English national curriculum and the Common Core State Standards used in the United States, in order to be more widely applicable. Both curricula have included instruction in morphology. The Common Core State Standards has instruction of morphemes for meaning starting with grade 3 and going through to the end of compulsory schooling. The English National Curriculum includes morphology instruction as part of the English curriculum under spelling, starting with year one and inflectional suffixes and the derivational prefix un. The affixes that were included in the intervention that are also in the curriculum were selected to be instructed in higher years than the intervention took place to ensure there was not an overlap in the instruction. The themes progress from common or familiar (e.g.: tri-, bi-) to more difficult or less common (e.g., -rupt, fract), while also keeping a balance of prefixes, roots and suffixes. There are 14 prefixes, nine suffixes and nine roots for a total of 31 morphemes taught.

All participants in the main study used the same intervention plan. The pilot study tested a plan with more activities and sessions each week that was later changed to better fit with the needs of the research schools, while still aligning with current research findings about instruction.

Table 2:Intervention Plan

Intervention Plan

Week # Theme	1 Numbers	2 Not	3 Science	4 Person Who	5 Look/ Light
Session #1 A) Reading 1 B) Word Map	bi- tri-	dis- un-	geo- bio-	-er -or	Scope Photo
Session #2 A) Reading 2 B) Word Sort	uni- mono-	Review: A) Word Sort B)Sentence	geo- astr-	Review: A) Word Sort B)Sentence	vis/ vid
Week # Theme	6 Position	7 Communication	8 States	9 Build or break or movement	10 Useful
Session #1 C) Reading 1 D) Word Map	Pre- post- mid-	phone/ phon	-able -ible	-rupt -fract	de- co-
Session #2 C) Reading 2 D) Word Sort	Review: A) Word Sort B)Sentence	scribe/ graph	-less -ness	-struct port	Review: A) Word Sort B)Sentence

3.2.1. Session 1 Week 1

The first activity of each session was a short reading passage, the idea being that the target morpheme would be introduced in context to allow students a chance to build their own meaning and mental connections with the meaning of the morpheme. This would allow them to take ownership and build interest in the morpheme. It then progressed to students completing two-word maps shown and described below to build the skills of morphological decomposition and analysis, based on the success of that reported in Harris et al. (2011) but slightly modified to be more accessible and visual for younger learners. Harris et. al., worked with 203 ninth grade students to instruct them in vocabulary using two completing methods the LINCS method(Ellis, 1992) covered earlier and word mapping. Word mapping was found to have significant gains over both the control (test only) group and the LINCS strategy group when identifying both the meanings of word parts and new words. Based off of these findings, a simplified version of word mapping was employed in the intervention. Each activity will now be described in more detail with examples below.

The first session of each week contains a short reading passage (e.g., Figure 4) that includes the focus morphemes or roots for that session in context multiple times. This is read aloud by the students as a group and the focus morpheme is identified, discussed and a meaning is agreed on by participants and

the instructor. The morphemes were not emphasised in the printed text in any manner except through repetition.

The Brown **tri**plets were <u>three</u> of the wildest children you ever saw. They liked to ride their **tri**cycles and imagine they were taking part in a **tri**athlon. They used a mud puddle to swim in and their mum yelled, "You are not a **tri**o of ducks! Come in and get a bath!" Afterward, they begged for a **tri**logy of stories to be read to them before bedtime.

Figure 4: Example passage for the prefix tri-

The use of bold for the prefix tri- in the example is for emphasis here. The passages were short: between 50-100 words. A clue as to the meaning of the target morpheme was provided in the passage (in the example, it is underlined word three, the emphasis on the clues was not present in the texts given to the students). This was to encourage students to use skills they likely had already acquired to analyse the morpheme and to develop a more personally meaningful meaning and mental schema. Participants first read the passage aloud as a group, then they were asked to identify the target morpheme and underline all words containing it. Then they would circle the morpheme and think about what it could mean. As a group they would come to a consensus as to what it meant in the contexts shown. If as a group they were correct, they were praised for their hard work. If incorrect, they were led to clues that might give them a better idea in the passage. This way students took ownership of the instruction and potentially gained confidence in their own skills as they progressed through the intervention sessions.

The second activity is a word map (Figure 5) where students take two words containing the target morphemes from the reading passage, break them down by their morphemes and the meanings of those morphemes and then reconstruct the meaning of the original words. The first word is broken down and rebuilt as a group and the second word is worked on independently. This gives students focused practice and understanding that words are made of morphemes or pieces of the words that carry meaning and can be combined in different ways. It also allows a chance for them to notice that some morphemes are bound and must be combined with others to form a complete word (e.g. *dis*-: *dis*turb) whereas others are unbound and might stand alone (e.g., *scribe*).

The example given below (Figure 6) is for the prefix tri- after choosing one of the words from the passage in this case tricycles. The students would break the word into its morphemes (tri-, cycle, -s) and place them in the appropriate box for prefix, root and suffix. Then the meaning of each morpheme was filled in. It could be a simple synonym or a multi word phrase that would convey the meaning. In the case of tricycles, tri: three, cycle: a circle or wheel, -s: plural or more than one. Finally, students would combine the meanings. The meaning of tricycle being more than one three wheeled thing. The mix of group and independent work allowed the students and researcher to assess on a continuing basis if individual students were understanding both the instructions for the activities and the morphemes themselves. This was to ensure that students understood each morpheme before progressing to the next one. This feedback was intended to encourage students if they had difficulties and give positive reinforcement to those that were successful.

Word Map

Word:		
Prefix:	**Root:**	**Suffix:**
Meaning of Prefix:	**Meaning of Root:**	**Meaning of Suffix:**
Combined meanings:		

Figure 5:Blank word map

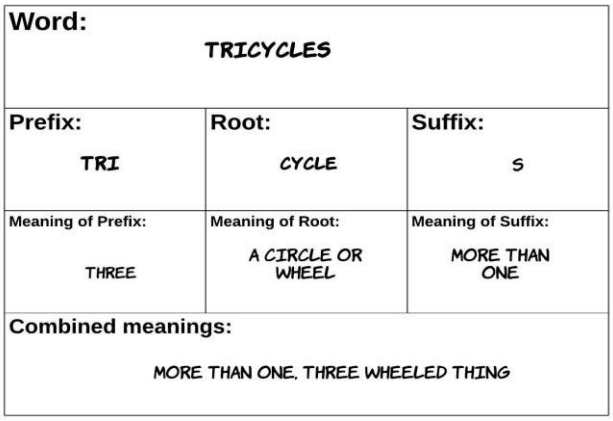

Figure 6:Example of a completed word map for the prefix tri-

The morphemes that were introduced during session 1 each week received more exposures as participants were exposed to them in more ways than those which were introduced in session two. The morphemes introduced in session one were also used in the word sorts and sentences that were part of session two, which will be covered in more detail next.

3.2.2. Session 2 Week 1

The second session each week began similarly with a reading passage with new target morphemes that are thematically associated with the ones focused on in the first session. For example, in the week on science the target morphemes are *astr*, *bio* and *geo*. The difference from the first session was that the second activity is a word sort. Participants are given a set of cards with a word on each one (see Appendix 12.3). The cards were made by printing the words containing the target morphemes on cardstock and printing them out and cutting them into individual words. Words are sorted by meanings, affixes, roots, and student generated groupings (e.g., words that could be names, subjects to study in school and, words that have number affixes). Word sorts are common in phonics instruction so it would not be unusual for a student to have done something similar before. The intention behind this activity was to reinforce students' recognition of different affixes and roots. It allowed them a chance to encounter new words containing each and as the sets of cards to sort had a minimum of four affixes or roots that had been studied that would need to be sorted. This activity tended to feel like a game for the students as they would try to race other students or come up with their own categories for sorting.

3.2.3. Session 3 Week2

The third session is the same as the first introducing the new theme for the week and target morphemes. Participants would read a passage and complete word maps for the target morpheme. This structure enabled students to gain more knowledge of morphemes.

3.2.4. Session 4 Week 2

Every fourth session is a review of the three previous sessions' morphemes with a word sort and a writing activity. This session did not contain a reading passage. The word sort would contain all the cards for the morphemes that had been used in the previous three sessions this would include words that included the morphemes instructed in session 3 which were not used in the sort during session 2. They would be sorted similarly either as individual participants or in pairs. This gave some support for students who were struggling to remember the previously introduced morphemes. It was also an opportunity for students to play games as they sorted and thus keep interest in participation and the activities from becoming stale after multiple cycles.

For the writing activity the students choose one of the words they had learned containing a target morpheme and write a sentence using the morpheme correctly. This ensured that the participants saw each morpheme multiple times and had multiple chances to work with and gain familiarity with each. This aligns with current ideas that students need opportunities to not only be exposed to new words but to use them in different ways (Crosson et al., 2018; Nagy et al., 2012; Beck, McKeown and Kucan, 2013). Examples of sentences from students written during the intervention are in Appendix12.4.3.2.5 Exposures to Morphemes

Due to the spiral nature of this intervention, students were exposed to the morphemes from session 1 potentially four times. These exposures were in the reading, word map, word sort and sentence writing. While morphemes introduced in Session two were likely to only have been exposed to participants three times, in two-word sorts, a reading and potentially the sentence writing. Morphemes introduced in session 3 were potentially only given 4 exposures to the participants in the reading, word map, and word sort and sentence writing. As there were no morphemes introduced in session 4, this is where more consolidation of learning was predicted to take place.

This section has attempted to provide a brief summary of the principles and rationale that were used to design this intervention. This design while seeming simple has the benefits of being familiar to many teachers. It also conforms with current advice for vocabulary instruction and allows for repeated exposure, exploration of words and practice using the words. In the next section, I will present the design and rationale behind the measures of derivational morphology that were designed for this study.

3.3. Selection and Design of Measures

This section will cover the rationale for the selection and design of measures to asses if there were changes in knowledge of morpheme meanings, particularly over the period of the intervention and establish both baseline and any changes in participants reading ability based off of comprehension, accuracy and reading rate.

3.3.1. Design of Morphology Measures

There were two measures selected for use: the Word Part Levels test and the Test of Root Knowledge. The Word Part Levels test was selected as it was already designed and tested, though it had been used with an older population. It would cover some of the aspects of affix meanings. This left a need for a measure of root words. As there was not a standardized measure available to this researcher, one was designed and piloted by the researcher. These two measures are described in the following sections covering their usage and development.

Deacon et al (2008) outline a taxonomy of morphological tasks considerations are made as to the presentation of task and how responses are gathered whether orally or in writing. To reduce the cognitive load and potential issues both tasks were presented orally and in writing. However, responses were gathered only in written format, with participants circling what they thought was the correct answer. As productive tasks are considered more difficult, the tasks designed and chosen for this study were judgement tasks, although these also have issues when used with sentence completion activities as participants could simply select what sounds best (e.g., Diamanti et al., 2018). Ceiling effects have also been found in younger participants when judgement tasks are used to ascertain if works are related (bake- Bakery). The ceiling effects are found much later for sentence completion or analogy tasks (e.g., Fejzo et al., 2018; Tyler & Nagy, 1989). This was considered and attempted to be avoided by not giving context for words to be inserted instead using the same format of sentence for each item.

3.3.2. Modified Word Part Levels Test

The first measure is a measure of comprehension of written derivational morphemes. This was a method for analysing a dimension of morphology mainly focusing on the extent to which the participants could identify the meaning being carried by the target morpheme. It was selected because one of its sections covers meaning of common prefixes and suffixes. The measure is based on Sasao and Webb's Word Part Levels test (Sasao and Webb, 2017), but has been adapted as the original measure was designed for university students (aged 18-23). It was designed with multiple levels of difficultly. As it was going to be used with upper primary aged students, the easy form of the test was selected for use. The easy form of the test comprises 87 questions and covers three sections: form, meaning and use. These 87 questions came from 118 affixes that appear in more than one of the most frequent 10,000-word families of the British National Corpus. This aligns with current theory that knowledge of the 9,000 most frequent words along with proper names is desired for comprehension of authentic text (Nation, 2006; Laufer and Ravenhorst-Kalovski, 2010). Each question contained an affix, two words containing the affix and a choice of four synonyms for participants to select as the meaning. (See Figure 7 for an example.) This format asked them to see a morpheme and then identify a meaning, which would make it more of a receptive task.

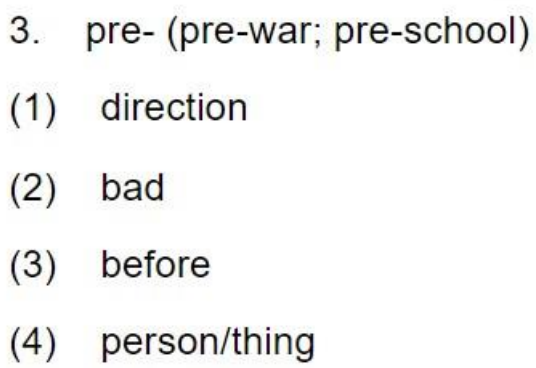

Figure 7:Example WPLT question

The original measure covered more topics than were necessary for the current study.
In order to make it accessible to a younger group and to align it closely with the focus of the study, the decision was made to use only the meaning section of the test with a reduced question set from 34 questions to 24.The reduced measure was shared with several colleagues who have experience teaching the target age group. Feedback was requested on accessibility and whether they thought it

would be reasonable for students at this level to be able to complete. Based on the feedback, the number of questions was then reduced further to ensure that the difficulty level was appropriate. The 19 affixes that were selected include 13 prefixes and six suffixes and thirteen of the selected affixes are explicitly taught in the intervention. The items are being defined as taught in this work as: the item was the primary focus of instruction, and the meaning was discussed during the intervention. As many suffixes tend to function to change grammatical class, only the suffixes that are primarily used to carry meaning were included, this made the available pool of questions much smaller for suffixes than prefixes. The measure was then piloted with a group of year 3 students and the findings are in the pilot summary below. This measure was administered in small groups with the questions and answers being read aloud with the participants being able to read along and select the correct answer. This was done to reduce testing time and ensure that if there were issues where a participant did not know how to read a word it would not impede their participation. This should also be a familiar format for the participants as many standardized exams are given this way in the lower primary years. A copy of this measure is contained in Appendix 12.1.

3.3.3. Test of Root Knowledge

The test of root knowledge was developed to assess the participants' knowledge of the meaning carried by common roots. This would indicate if they acquired knowledge as a result of the intervention. As discussed earlier while there has been research and measures made for inflectional morphemes there are very few if any for derivational morphemes and even fewer for knowledge of roots. This led to a need to design a measure specifically for this purpose. This was not easy or straightforward as there were no measures to base this on. First, a standard phrasing that would work for a multitude of words was developed. The format for the questions was standardized to read, "Which word below do you think has something to do with **looking** at something?" followed by three choices matched for length, and position of morpheme in the word (see Figure 8 for an example). By providing the meaning and then asking for the target morpheme, this measure would work in the reverse of the other measure by being more of a productive measure. Using the lists of Greek and Latin roots prepared by Manyak et al. (2018), words were selected that matched the roots being taught in the intervention, and an additional set that were not being taught. This resulted in a list of 22 words with 11 containing roots that would be taught in the intervention. Using an online dictionary (https://www.onelook.com/) that could generate words that matched in length and position of root in the word, alternate choices were selected. Before being confirmed they were additionally checked to

ensure that they appeared in the Collins online primary school or secondary school dictionaries. Collins dictionaries are commonly used in UK schools as references, so they are a good way to check to see if the words would be of the level and usage that could be expected in a UK classroom.

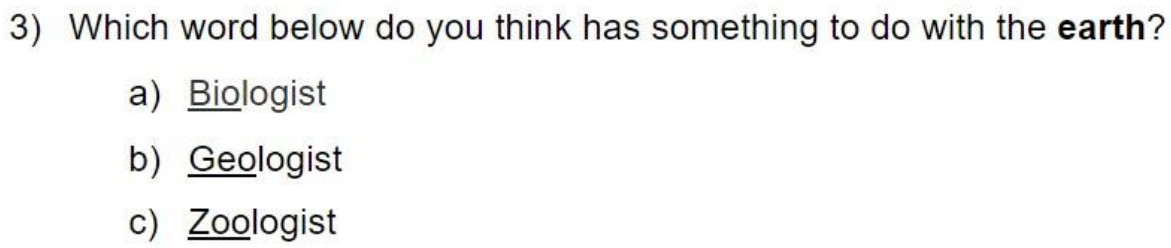

Figure 8:Example question from Root Measure

As a first step, an initial version of the measure that was twenty questions long was piloted with a group of 10 graduate students that was comprised of 5 EAL and 5 L1 speakers. Scores were then calculated, and discussions were had about why certain questions were hard and if the wording was clear. Some questions were modified or dropped due to difficulties with knowing what the correct answer would be with students being split between choices, or if they had issues with understanding the question. The measure was then re-piloted with a smaller group of 5 graduate students who had not participated the first-time round. At this point all questions were clear and accessible to graduate students and the measure had 16 questions and four alternate questions that could be used if needed. The alternates could be used of there were issues such as all participants got the question correct or incorrect when piloting with year 3 students. Additionally, they could be used to replace questions that appeared to have obvious miscues where students consistently answered a specific incorrect answer.

The measure was then used in the pilot study with year 3 primary students and the scores were correlated with the scores on the standardized measures. Analysis of the results revealed that some of the questions were answered correctly by every student and others were answered incorrectly by every student. As such, four new questions were added at the end of the pilot to replace these problematic items. The final version of the test contained 15 questions of which eight are taught in the intervention and seven are not taught. A copy of the final full measure is contained in Appendix 12.2.

This measure was administered to small groups, with the questions and answers being read aloud with the participants being able to read along and select the correct answer. This was done to reduce testing time and ensure that if there were issues where a participant did not know how to read a word it would not impede their participation. This should also be a familiar format for the participants as many standardized exams are given this way in the lower primary years. This measure and the Word Part Levels test were administered at the same time with the Word Part Levels test being given first each time.

3.3.4. Selection of Standardized Measures

Two standardized measures were selected to complete the battery. The British Picture Vocabulary Scale (BPVS) was selected as a measure of receptive vocabulary. The second standardized measure selected was the York Assessment of Reading Comprehension (YARC) to measure reading comprehension. Standardised measures were included in the test battery in order to provide a benchmark of the children's performance and to explore the relationship between these measures and the direct measures of knowledge of affix meaning and knowledge of root meanings. The rationale for the selection of these tests and descriptions of both follow. These measures were all designed to be administered individually. They were administered with each participant in a quiet area or small room in the participating school. Total time for administering each set of measures was approximately 35-40 minutes. The pilot participants were tested pre-intervention and post-intervention. The main study participants were tested three times as explained in more detail in Section 5.3 Procedure.

3.3.4.1. British Picture Vocabulary Scale

The British Picture Vocabulary Scale III (BPVS) measure is widely used in the United Kingdom to measure the size of receptive vocabulary among people aged 3-16. This was selected as knowledge of how the size of vocabulary could have changed due to the intervention was needed to evaluate if there was an impact on the vocabulary size due to the intervention. It is also easily administered as it consists of sets of pictures that increase in difficulty and students merely have to identify the picture associated with the word spoken. Administration typically takes between 11-15 minutes and can only be administered to one participant at a time. The reliability for this measure has been reported as 0.91. Since it is a widely accepted and utilized measure it was selected to measure the receptive vocabulary

for the participants. This third version of the BPVS was standardized in 2009 with a sample of 3,278 students ranging in age from 3 years to 16 years old. It was drawn from 161 schools in the UK which means in many ways it is standardized for the population of this study. The one short coming is that only 45 of the participants in the standardization were reported as EAL students. This means that the standard scores, percentile and age ranks might not be applicable for EAL students. This is part of the rational as to why the raw scores and not standard scores were used in this research. Receptive vocabulary was used as it is typically viewed as developing before productive vocabulary, as noted in the literature review chapter, therefore changes might be observed more rapidly than with a productive vocabulary task. As discussed in the literature review, vocabulary is a major component of reading comprehension and very closely associated with morphology it makes sense to measure this aspect specifically. Reading development is additionally based on acquisition of vocabulary. By selecting a measure of receptive vocabulary that removes reading from the equation as participants do not need to read anything to answer the questions. This allows the focus to solely be on the vocabulary size of the participant. This allows the researcher to utilize the measure as an identifier of change in size while not being dependent on any change in reading ability.

3.3.4.2. York Assessment of Reading Comprehension

The second standardized measure that was selected was the York Assessment of Reading Comprehension (YARC), which measures comprehension, accuracy, and fluency. Additionally, the Single Word Reading Test that is imbedded in the YARC gives a measure of reading individual words. These were necessary to assess the baseline and change for the participants reading ability. As discussed in the literature review reading compromises multiple skills and thus to be effectively measured needs to be examined in multiple dimensions.

The YARC is administered individually and consists of the SWRT and two reading passages that are usually set based on the participants' performance on the SWRT. However, it was noticed that students were performing well on the SWRT and then unable to read the passages at the level indicated by the level of words they could read. This is likely due to the emphasis on phonics instruction in lower primary grades, resulting in students being able to pronounce the words but having little to no comprehension of the meaning of the words. Therefore, it was decided that all students would start with the passage levelled for year three students (Form A-3) and the second

passage would then be either the higher passage or lower passage depending on questions answered correctly (greater than four questions correct moves up) or number of reading errors (more than 20 moves down) scored in during the reading. The comprehension score was calculated based on eight open ended questions that were asked after the students read each passage. Their answers were scored as either correct (1) or incorrect (0). Open ended questions allow for less guessing by students as there are no choices to choose from. The participant would need to answer a question about the passage from the information in the passage or by connecting it to prior knowledge.

The form that all participants read was Form A-3, which was a fiction passage about a burglar being caught in a house. The comprehension questions for this passage which participants answered were a mix of Literal information (1), Knowledge based inference (4), Cohesive device (1), Elaborative inference (1), and Vocabulary dependant (1). If participants answered 4 or fewer questions correctly for this passage or had more than 20 reading errors the second passage, they read Form A-2. Form A-2 was a non-fiction passage about robins and their nests. This passage also had a mix of reading comprehension questions. They were Litteral information (6), Knowledge based inference (1) and, Elaborative inference (1). If the participant scored more than 5 or more questions correctly for form A-3 and less than 20 reading errors, the second passage was Form A-4. Form A-4 was a nonfiction passage about a type of lizard. This passage also had a mix of reading comprehension questions, that included Litteral information (3), Knowledge based inference (2), Cohesive device (2), and Vocabulary dependant (1).

These three forms A-2, A-3, A-4 were the same at all three times the participants were assessed. The participants only read the two passages and from the one reading, the accuracy and reading rate were assessed. This same passage was the passage on which the reading comprehension questions were based. There were discussions about utilizing the alternative form of the YARC form B for time 2 and then reverting to Form A for time 3. This was decided against as it would introduce an additional variable as the question types and topics for Form B were different from Form A. As participants were given no feedback and were not told if they had answered any of the questions correctly and there was ten weeks between each testing, it was decided that it was not likely that there would be significant practice effects to utilizing the same passage three times with the participants.

The other scores reported on the YARC are accuracy and reading rate. These connect to the concepts of decoding and fluency. As previously discussed, decoding and fluency are key aspects of reading

comprehension as identified in the Simple View of Reading. This assessment utilizes the construct of reading accuracy to assess decoding ability, and the construct of reading rate for fluency. The concept behind this being that if students can accurately read words aloud, they are decoding them properly. Fluency being measured by how long it takes a student to read the passage. The accuracy score is calculated as an ability score based on the average number of errors from the two passages. Errors included in this calculation are as follows: Mispronunciations, Substitutions, Reversals, Refusals, Additions and, Omissions. Self- corrections and omissions or repetitions of a line or more of text are not included in the errors(Snowling et al., 2009). Reading Rate (Fluency) was calculated by taking the number of words in the passage divided by the time taken by the participant to read the passage in seconds then multiplying this by five. This would produce the raw score for fluency.

Reliability is reported for different levels and forms of the YARC. Since the majority of participants were administered the forms for level three, I will report those measures of reliability here. The reliabilities for Accuracy, Reading Rate and Reading comprehension are Cronbach's alphas of 0.93, 0.92, 0.48 respectively. If we look at the pairs of passages that were most commonly administered the Cronbach's alphas for reading comprehension are 0.71 and 0.73 which are much more acceptable. The YARC was standardized with a sample of 1,376 participants from ten regional test centres that were distributed across the UK. The sample was also split between girls and boys 694 and 682 respectively and across year groups from Reception to Year 6 with approximately 200 students from each year. However, only 14.02% of the sample was EAL though, this is very close to the national statistics for pupil characteristics that was released in the year the standardization occurred. This would mean that it is adequately standardised for use with EAL students in the UK.

As one of the major research questions of the study is the impact of morphological instruction on reading comprehension a measure of reading comprehension needed to be included. This one was selected as it was easily administered and contained sub measures that could give more depth to the analysis. It generally took approximately 25-35 minutes to administer all the parts of this assessment.

6.3 3.3.3.5 Mapping of Measures to the Research Questions

The variables are firstly ones that are innate to the participants, such as their language background, gender, age and attendance at the intervention sessions. The dependant set of variables are the knowledge of derivational morpheme meanings and reading abilities.

The reading ability of the participants is measured by the YARC which is broken down into several components such as accuracy of reading a passage allowed, the rate at which the participant reads the passage and the number of questions about the passage answered correctly. This measure was designed to reflect the components of the simple view of reading of comprehension and fluency, with fluency being demonstrated by the ability to accurately read a passage in a time period.

The participants' knowledge of derivational morpheme meanings is measured by the Test of Root Knowledge and the Modified Word Parts Levels test. These assess if a participant can identify the correct either meaning or morpheme for a meaning from given choices.

The independent variable in this study is the instruction that was given in the intervention.

The British Picture Vocabulary Scale was administered to detect if there was any significant shift in the vocabulary size of the participants over the time periods of both the control and intervention period. As it has been demonstrated that morphemes are one of the building blocks of vocabulary knowledge it there is the potential that if there were significant changes to the participants' knowledge of morphemes there is potential for a change in vocabulary size. 3.4 Recruitment of schools and participants

Recruitment for schools to participate began in early 2019.Over 100 Email and print invitations to participate in the study were sent to all primary schools who were reported to have at least a 25% EAL population on the 2018 List of schools in England and located in Yorkshire. This area was selected as it had the population of both L1E and EAL students that was being looked for and it was within reasonable transportation area to fit the researchers visa requirements to reside near the university. From this group there was a 10% response rate and after further discussion, one school agreed to pilot the study with potential to agree to participate in the main study, and two other schools agreed to participate in the main study. Other schools declined to participate as they stated they did

not have the space in the school for the intervention to be conducted or the teachers did not feel there was time in the students' timetables for interruptions.

3.5 Additional work that was planned but not carried out.

There was additional post testing that was planned for 10 weeks after the end of the intervention to see if there were long term benefits or residual effects from the intervention. This was to coincide with interviews with the teachers and students who participated. Additionally, there was a survey about speaking and reading habits of the EAL students which was planned to be completed at the follow up testing. It was planned for the last time so that students would feel comfortable with the researcher and be willing to answer honestly about personal information such as what language they speak with their family and if they receive any instruction in their additional language. This was brought forward to the Time 3 testing and data is included at the end of the data analysis and included in an amendment to the discussion.

4. Pilot Study Summary

A pilot study was conducted in the spring of 2019. The purpose was to check the interventions accessibility with year three students, ensure feasibility of activities within a typical primary school setting, and check the researcher designed measures for accuracy and utility in a field setting. This also allowed time to modify the intervention schedule and measures before the main study started the following autumn.

4.1. Participants

Students were recruited from a local primary school. All children were recruited from Year 3. Invitation letters and consent forms were sent to all students and 12 students took part in the pilot study after returning signed consent forms. There were 10 L1 students and 2 EAL students (7 male students).

4.2. Materials

Intervention materials were written and designed by the researcher as detailed in the intervention design section. A shortened version of the intervention only lasting 5 weeks was used for piloting as it was requested by the pilot school to fit with the school calendar and minimise interruption of the curriculum. The shortened version is shown in Figure 9. The measures for the pilot were both the researcher designed variation on the Word Parts Level test and the Test of Root Knowledge and he the standardized York Assessment of Reading Comprehension (Snowling et al., 2009) and the British Picture Vocabulary Scale. III (Dunn & Dunn, 2009)

Week # Theme	1 Numbers	2 Not	3 Science	4 Person Who	5 Look/ Light
Session #1 **A) Reading 1** **B) Word Map** **C) Web**	bi-	dis-	astr-	-er -or	Scope
Session #2 **A) Reading 2** **B) Word Sort** **C) Word Map**	tri-	un-	geo-	-ist	vis/ vid
Session #3 **A) Reading 3** **B) Word Sort** **C) Web** **D) Sentence Activity**	uni- mono-	Review: A) Word Sort B) Web C) Sentences	bio-	Review: A) Word Sort B) Web C) Sentences	Photo

Figure 9:Pilot Intervention Schedule

4.3. Procedure

Pre-intervention testing was carried out in April. Testing took approximately 45 minutes per student and was administered in a quiet area of the student's school. Tests were administered in two sessions by the researcher. In the first session each participant was tested individually using the York Assessment of Reading for Comprehension (YARC) Second Edition (Snowling et al., 2009) and the British Picture Vocabulary Scales 3rd Edition (BPVS IIIi; Dunn & Dunn, 2009), following the guidelines set out in the manuals. In a second small group setting students completed the modified Word Part Levels Test (WPLT; Sasao & Webb, 2017) and the researcher made Test of Root Knowledge (TRK). Both measures were read aloud to students and students followed on printed copies, marking their answers in pencil.

Post-pilot intervention testing occurred after a one-week half term break. The tests were the same as the pre-test with the addition of four questions to the root measure to pilot possible questions to replace those that showed ceiling effects on the pre-test. Testing occurred in a quiet room off the Assistant Headteacher's office.

4.4. Results

As this was a pilot, the focus was on the functionality and feasibility of the measures and intervention. The expectation was that the morphology measures would correlate with each other and potentially with the reading comprehension measure. It was also expected that each session of the intervention

would be able to be carried out in a reasonable time period of between twenty to thirty minutes to reduce the disruption to the classrooms where participants were drawn from. Finally, it was predicted that the participants reading, and vocabulary scores would improve though possibly not significantly as they were only receiving approximately 8 hours of instruction over the five weeks.

The data was analysed using IBM SPSS 25 and R. Initial correlations were carried out to look at the relationship between the morphology measures and the measures of reading and vocabulary. This was to ensure the accuracy of the measure. It was expected that the two morphology measures should have strong correlations as they are measuring very similar concepts. The WPLT was correlated with the TRK r= .948, p=.000 on the pre-test and r= .754, p= .012 on the post-test. This very strong correlation demonstrates that there is a correlation between the participants initial understanding of affixes and root word meanings.

Additionally, the WPLT post test score was strongly correlated with the score for BPVS r=.653 p=.021. The TRK correlated with YARC reading comprehension r=.602 p=.602, YARC reading rate r=.594 p=.042, SWRT r=.682 p=.015, and the BPVS r=.769 p=.003. Using the BPVS raw score is an emerging trend with LMY and EAL students in research (e.g., Geva, Farnia, Geva, & Farnia, 2012; O'Connor, Geva, & Koh, 2019) as the scaled scores for the BPVS were not standardized for EAL participants.

Table 3:Pilot Pre-Test Measures Correlations

Variable	*M*	*SD*	WPLT	TRK
WPLT	10.00	3.86		.948**
TRK	11.00	2.98	.948**	
YARC Comprehension	109.58	11.69	.515	.602*
YARC Reading Rate	108.75	15.28	.494	.594*
YARC Accuracy	105.92	12.96	.263	.363
BPVS III	107.67	22.32	.653*	.769**
SWRT	41.67	9.95	.558	.682*

Note. *M* and *SD* are used to represent mean and standard deviation, respectively.

* indicates p<.05. ** indicates p<.01.

Table 4:Pilot Pre and Post Test Data

	Pre-Pilot			Post Pilot		
	Mean	Standard Deviation	Range	Mean	Standard Deviation	Range
Test of Root Knowledge	12	2.5	8	13.2	1.5	5
	10.8	3.7	10	13.6	3.6	9
	44.7	6.9	20	47	6.3	16
	108.9	12.1	39	115.1	10.7	31
York Assessment of Reading Comprehension (YARC): Reading Rate	113.1	12.6	39	114.6	10.5	36
York Assessment of Reading Comprehension (YARC): Comprehension	112.1	11.2	42	119.4	9	27
British Picture Vocabulary Scale: Raw Score	113.9	18.6	59	119.9	19.5	54
British Picture Vocabulary Scale: Scaled Score	105.4	17.6	55	109.1	18.5	55

The post pilot measures showed increases across all measures with reductions in ranges of scores (see Table 2). Notably the range on the YARC Comprehension scores saw a significant reduction from 42 to 27, this possibly shows a narrowing of the performance by the students. The Accuracy section also showed gains and a reduction in range, while the Reading Rate showed the smallest change. Part of this might be due to the students taking slightly more time and being more cautious in the post test. Maximum standard score on the YARC for all sections is 130 and there were students on all sections both pre and post that were near or at that score.

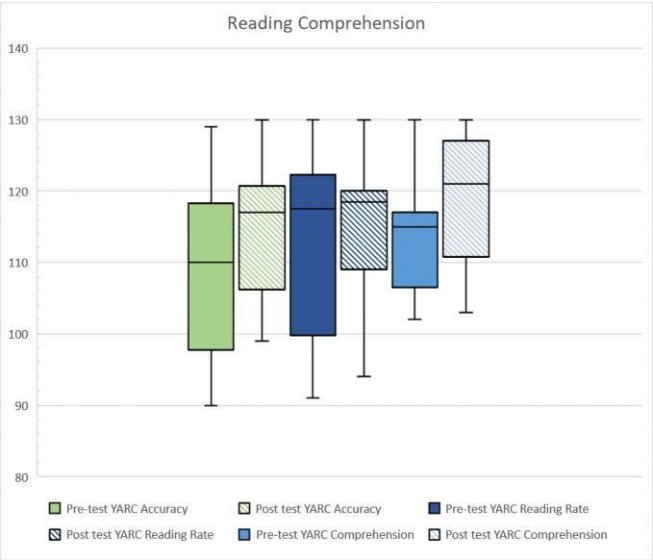

Figure 10:Pre and Post Test Pilot YARC scores

For the measures of receptive vocabulary (BPVS), both the raw and scaled scores showed improvement. The interquartile range stayed approximately the same for the raw scores while the scaled scores showed a noticeable difference with the size of the distance between the median and the Q3 growing. As most of the students in the pilot were L1 speakers of English, the focus should be on the scaled score. The raw score is given so it can be compared with the main study where there will be a significant sample of EAL students.

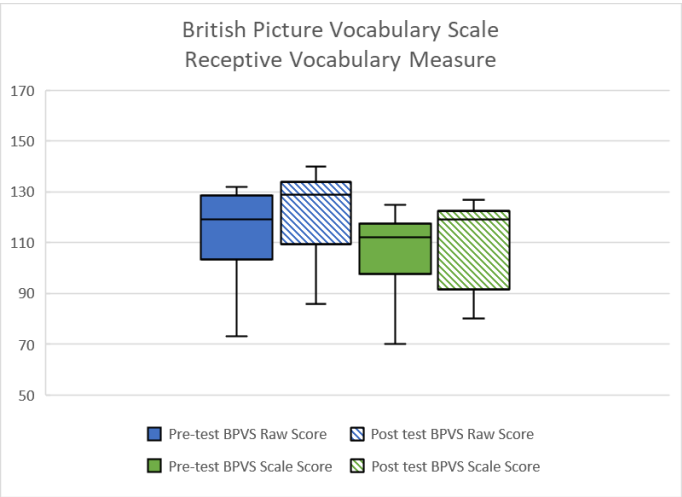

Figure 11:BPVS Pre and Post Pilot

The morpheme measures showed increases as well. Both the Word Parts and the Roots measure showed increases in the means and decreases in the range. The Roots measure showed some ceiling effects on the post-test so it will be modified before the main study using the additional four questions that were added for the post test on the pilot. The questions that were answered correctly by all participants on the pre-test and questions where the nine out of ten participants answered correctly were exchanged. The balance of the questions for the main study will still be eight taught roots and eight untaught roots.

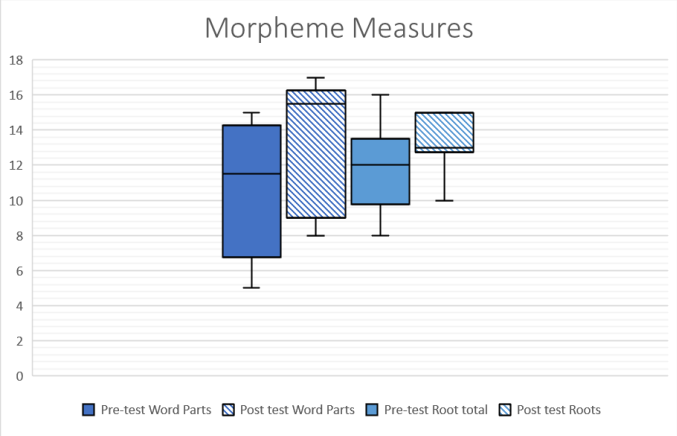

Figure 12:Pre and Post Pilot Morpheme Measures

4.5. Reflections on the intervention:

The first week of the intervention started on April 30th after students returned from a school break. One L1 student dropped out after the first week due to scheduling conflicts with additional services he was receiving from the school. Three groups of 4 students each were trialled then dropped in favour of two groups of five after the loss of a second student.

Sessions were conducted in a relatively quiet area of the school. There were occasional interruptions but overall, it was a good and conducive place to work. Students seemed to enjoy the first week. They did seem to struggle a bit at the beginning with trying to figure out some of the activities potentially because students were not familiar with Morphological Analysis. As they practiced the skills over time, they became very good at finding the different morphemes. From the second week on, students enjoyed the challenge of trying to figure out what morphemes we were working with each session. Several students commented that they enjoyed learning this way. Teachers who have passed by have

seemed interested and at least one has commented that they are impressed by the level of work and the discussion the students are having about the words.

Three sessions were conducted each week with the fortnightly rotation containing six sessions, with each focusing on a theme for example science, or people who. Over the five weeks each student participated in a total of 15 sessions. Each session typically ran for 25 to 35 minutes and were held in the afternoons to minimise interruptions to the regular classroom activities, at the request of the school. One issue that occurred was that year 3 students go to swimming class one afternoon a week so that only left four potential afternoons for the intervention to run. This is a common issue with schools in North Yorkshire and will be an issue with the schools for the main study. This has led to a redesign of the intervention to only have two sessions per week to ensure that all sessions can be conducted in a week with minimal interruptions to the regular class schedule. The session that was dropped contained a reading passage and a word web of words participants knew that contained the target morpheme. Though time may be tight as there are only really two hours in the afternoon (after accounting for attendance being taken, time to pack up at the end of the day, and afternoon recess), which means that four sessions could be conducted each afternoon, or 24 students could participate. With a projected sample size of 50 participants this will be tough but possible to get all sessions in each week.

5. Main Study

5.1. Participants

The participants for the main portion of the study were recruited from three schools in Yorkshire. The three schools all were in similar areas with a similar proportion of students receiving free school meals (between 33-41.6% of the student body). All three schools had similar attendance rates with approximately a 4-5% absence rate over the 2018/2019 school year (the year in which this study took place). There was a variance in the percentage of students in the school population that were EAL, the two schools that remained at the end of the study had student populations that were between 21-17% while the school who was not available for time 3 tests was 58%. These schools were consistent with average school characteristics of 21.2% EAL and 23% on free lunch. This data came from the gov.uk website and provided by national statistics. Students were identified as being EAL by the schools they attended. When participating schools were selected, they were asked for students who were coded as EAL on the annual school census. The definition of EAL used by the Department for Education for the 2018 Census was "– where the pupil has been exposed to a language other than English during early development and continues to be exposed to this language in the home or in the community" (School Census autumn (2017), p 64) Using this definition meant that there was a common identifier for schools to use to identify students who met the criteria for participation.

Invitation letters and consent forms were sent to all qualifying students at each school in the first month of the school year. Students who returned consent forms were then tested the following week.

All students were in year three with no diagnosed learning disabilities. EAL participants had all attended school in the United Kingdom from reception to year 3. The participant count at the beginning of the study was 18 EAL students and 23 L1 English speaking students. The EAL participants' home languages were Bengali, Latvian, Italian, Polish, Portuguese, Tigrinya and Urdu. The average age of the participants at the time of the first measure was 7.5 years old. There were 20 male participants and 21 females. There were 6 male EAL students and 12 females.

Table 5:Participants by gender and language

	Male	Female
EAL	6	12

L1	14	9
Total	20	21

Table 6:Reported home languages of EAL participants by the schools

Home Language	Number of Participants
Bengali	1
Latvian	1
Italian	2
Polish	3
Portuguese	1
Tigrinya	1
Urdu	8

A statistical power analysis was conducted to determine recruitment goals using Apel and Diehm's (2014) study that included an 8 week intervention designed to increase awareness of affixes with kindergarteners(n=27), first (n=22) and second grade (n=26) groups of students. As the second-grade group was the closest in age to the participants that were recruited for the present study we used the effect sizes for that group in the calculations. They used two measures: one for relatives and one for affix identification, so power analyses were conducted for both. The effect sizes they present were calculated using the differences between pre-test and post-test measures using both adjusted and unadjusted means. The effect size was large (0.82) for the relatives and very large (1.52) on the affix identification. With an alpha = .05 and power = 0.80 with a statistical test of t-test, means: Difference between two independent means (two groups) selected, the projected sample size needed with this effect size (GPower 3.1) is approximately N =20 for relatives and N=7 for affix identification. This led to the original recruitment goals of 20 participants for both EAL and L1 groups, for a combined total of 42 participants. If using these same settings, the effect sizes from the pilot study are utilized, the sample sizes are significantly different. As the effect size is 0.57 for the test of root knowledge, the sample size for each group would need to be n= 39. If we use the effect size of 0.76 which was the

effect size for the Prefix and Suffix measure, the sample size for each group would need to be n= 23. If we use the sample sizes calculated from Apel and Diehm (2014) The final L1 group is fully powered while the EAL group is slightly underpowered. However, if we use the sample sizes from the pilot study both groups are underpowered. This will be taken into account in the analysis of the data.

5.2. Materials

The materials used were the same as designed and used in the pilot study with the exception of the redesigned schedule with the reduction in sessions from three to two each week (see Figure 3).

5.3. Procedure

The main study has a double base line design with initial measures given to all participants in September or October of 2019 and the second measures given ten weeks later in November or December of 2019 (see Figure 10). This theoretically gives an estimate of growth or change for each student over ten weeks. This meant that all participants were able to receive the intervention and be their own control. This additionally reduced some of the possible uncontrolled variables such as differences in class instruction, possible influences from differences in home language, or variations in starting points. All of these would be stable as the participants remained in the same school environment, with the same class and same teachers. The plan was for the intervention was delivered from January to March of 2020 over a period of 10 weeks. Followed by post-testing over two weeks before the April holidays. Additional delayed post testing was to be carried out a further ten weeks later in June 2020. However due to the rapidly evolving situation with the COVID-19 pandemic. It was decided to cut the intervention short by one week and start post testing as soon as possible. This led to the loss of one school as they closed their doors to visitors before the testing was able to be carried out. The other schools were tested in what would have been week 10 of the intervention. The planned delayed post-test was not able to be carried out due to a nationwide lockdown in England.

Figure 13:Timeline for main study

Attendance was tracked for all sessions with 8 EAL and 7 L1E participants attending all sessions, 9 EAL and 12 L1E participants attending greater than 75% of the sessions and 3 EAL and 0 L1E participants attending between 50% and 75% of the sessions. If participants attended less than 50% of the sessions their scores were dropped from the study. This led to 0 students' scores being excluded from the final totals.

Participants were grouped by class to minimise disruption to their classes, per the request of the classroom teachers. If a class had more than five participants, they were split into smaller groups, if they had less, they were combined with other classes, if possible, without going over five participants per group. There were eight groups ranging in size from three to five participants.

Sessions were held in the afternoons for two of the three schools at the request of the schools to minimize disruption to the reading and math block conducted by the classroom teachers every day. Sessions for the third school were held in the last hour before lunch. Each session took approximately 30 minutes. If there was a disruption to the schedule due to school activities, make up sessions were held that week or the following week at the convenience of the school and classroom teachers.

6. Data Analysis

This chapter examines the data gathered in response to the research questions constructed to examine the impact of the intervention on both EAL and L1 English (L1E) participants' understanding of the meaning of morphemes and reading comprehension. This chapter progresses from examining the whole sample (n=22) to looking at the subgroups of EAL and L1E (n=11) participants. The small sample size is due to a loss of data as the schools closed due to COVID-19. As the research questions included comparisons between both the EAL and L1E subgroups, this data has been reported as well. The means and SD are reported for each measure at T1, T2, T3. One-way ANOVAs were used to establish if there was a change over time, and repeated planned contrasts (T1 vs T2, T2 vs T3) were conducted to see if the change was driven from the change during the baseline period or during the interventions.

In the first two of the main sections of this chapter, the quantitative data and measure results were reported, first with both EAL and L1E groups combined as this is how most educational settings currently exist, then again separately to see if there are different effects for each group. The results are reported in the following order: receptive vocabulary, single word reading, reading rate, reading accuracy, reading comprehension, morphology: prefix and suffix, and morphology: roots. The measure of receptive vocabulary is the British Picture Vocabulary Scale version 3. The measures of single word reading, reading rate, reading accuracy and reading comprehension were all part of the Yorkshire Assessment of Reading Comprehension. The morphology: prefix and suffix is a reduced version of the Word Part Levels Test by Sasao and Webb (2017). The morphology: roots was developed specifically for this intervention. The results from the measures of receptive vocabulary, single word reading, reading rate, reading accuracy and, reading comprehension all relate to the second question, is there a relationship between reading ability and knowledge of derivational morphology and does this relationship differ between EAL and L1 English children? The two morphology measures are being used to answer the other research question: Is there a difference between EAL and L1 children in the acquisition of the meanings of derivational morphemes that are taught as part of a small group intervention. Is there an effect of individual differences in gender, attendance and age in the effectiveness of teaching the meaning of derivational morphemes?

The third main section includes the qualitative data gathered from teacher interviews about any impact of the intervention on student performance in their classrooms. This was gathered to obtain additional insights, as students may develop skills or knowledge that were not found in the quantitative measures. Classroom teachers are however in a place where they can observe changes in student behaviour or performance in a less formal manner. They are in a position to see whether a student is applying skills that were taught in the intervention on a regular basis as part of their classwork without prompting by outside influences.

The fourth section includes data from a questionnaire completed by the EAL participants at T3 about their language usage and instruction in home language. This was done post hoc as it was part of an original planned extension of the study but the majority of that was dropped due to issues surrounding school closures due to Covid-19. While it is not directly related to the original study it is being included as it raises interesting questions about designation of students as EAL in UK schools and differences between EAL and L1E students after multiple years of instruction in the same context.

Details of the 22 participants who completed the intervention and testing can be found in Table 2. Only their data was used for the analysis. The 19 participants who did not complete the measures at T3 due to the closure of schools in March 2020 as a result of the COVID-19 pandemic, are not included in this analysis. The baseline data from them is reported in appendices 12.6 and 12.7. Tests and analysis were carried out using SPSS. Repeated Measure ANOVAs were conducted using the participants' scores from the receptive vocabulary measure, single word reading test, reading rate, reading accuracy, reading comprehension, and two morphology measures. Where the data was found to violate sphericity, the Greenhouse- Geiser correction was applied. A paired samples t-test was conducted to compare the mean gains obtained after the intervention with those obtained during and equivalent period before the intervention for each participant. This was done to highlight the change over each period, baseline and intervention.

Table 7:Breakdown of Participants

	Total who started the study	Total Tested at T3	Male at T3	Female at T3	EAL at T3	L1 English at T3
School A	16	0	0	0	0	0
School B	7	7	4	3	7	0
School C*	18	15	5	10	4	11
Total	38	22	9	13	11	11

*School C had all participants complete the intervention however three were absent on the day of testing for T3.

6.1. All Participant Data

The whole group data is examined first, as most primary education classes in both the US and UK comprise a mix of both EAL and L1E learners. This would mean that if the intervention was successful, it would likely be implemented for both groups at the same time so looking at all of the data combined would give a better picture as to how this would work overall for a regular class. Sections 6.1.1 to 6.1.5 analyse the data for the research question about reading ability and knowledge of derivational morphemes. Section 6.1.6 and section 6.1.7 are being analysed to answer the research question: Is there a difference between EAL and L1 children in the acquisition of the meanings of derivational morphemes that are taught as part of a small group intervention. Is there an effect of individual differences in gender, attendance and age in the effectiveness of teaching the meaning of

derivational morphemes? The sub questions for both research questions will be addressed in Section 6.2.

6.1.1. Receptive Vocabulary (BPVS)

The participants' average scores on the BPVS test increased with time, from 87.73 (SD 11.91) at T1, to 97.73 (SD=11.75) at T2 and 104.68 (SD=14.23) at T3. The distribution of scores at each time point, and the change over time are summarised in Figure 11.

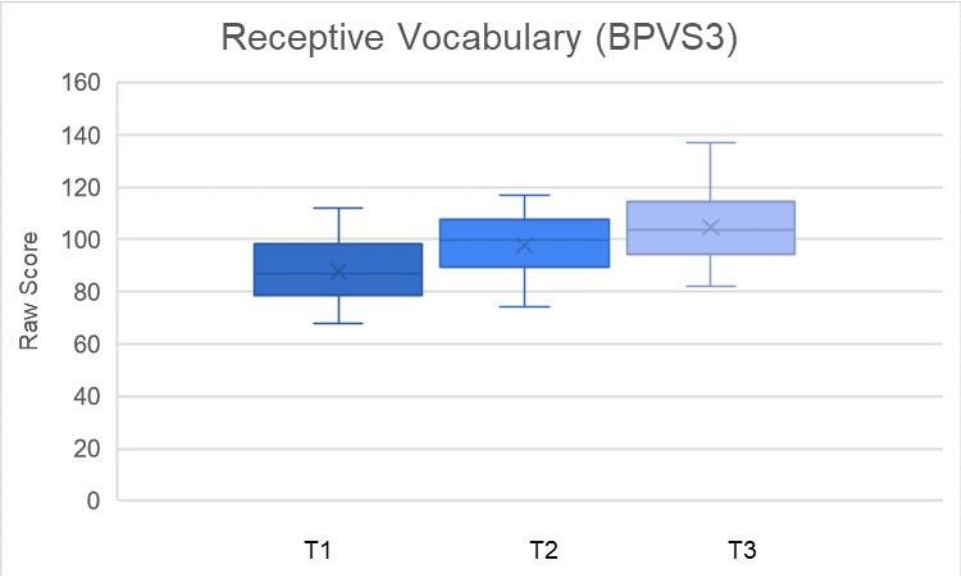

Figure 14: Receptive Vocabulary (BPVS Scores) by Time

A one-way repeated measure ANOVA confirmed that this change was statistically significant: $F(2,42)=25.18$, $p=.000$. The size of receptive vocabulary increased as time passed. Repeated contrasts revealed that the average score at T2 was significantly higher than at T1 ($F(1,21)=14.86$, $p=.001$, $r=.64$), and the average score at T3 was significantly higher than at T2 ($F(1,21)=9.47$, $p=.006$. $r=.56$). Both periods showed large effect sizes.

The average gain on the BPVS test between T1 and T2 (baseline measure) was 10.00 (SD=12.17), and between T2 and T3 (after the intervention), it was 6.96 (SD=10.60) (see Figure 12). A dependent (paired samples) t-test indicated that this difference was not statistically significant ($t(21) =0.71$, $p=.483$, $r=.15$).

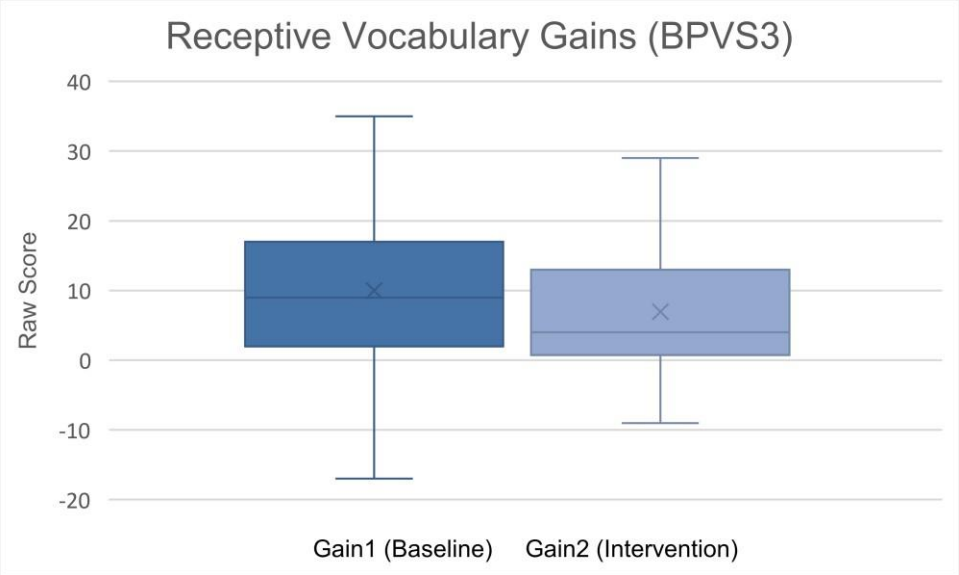

Figure 15:BPVS Gains

This illustrates that while there were gains in receptive vocabulary, they were not statistically significant. Thus, while the participants vocabulary did increase it was likely not due to the intervention as the average baseline gain is larger than the average gain from the intervention period. This means that there is likely no effect from the intervention on the participants general vocabulary knowledge. It does, however, show that there was student vocabulary growth over the entire period studied.

6.1.2. Single Word Reading Test (SWRT)

The participants average scores increased over time from 33.27 (*SD*=7.54) at T1, to 35.32 (*SD*=7.78) at T2, and 42.00 (*SD*=8.56) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 13.

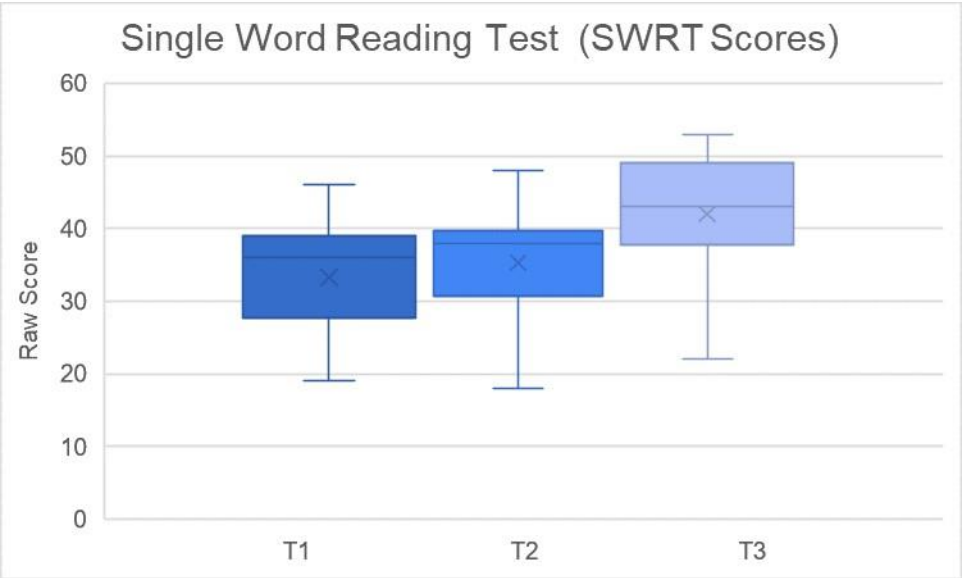

Figure 16:Single Word Reading Test Scores by Time

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 45.33$, $p=.000$. The ability to read single words increased over time. Repeated contrasts revealed that the average score at T2 was significantly higher than at T1 with a medium effect size ($F(1,21)=4.34$, $p=.050$, $r=.41$), and the average score at T3 was significantly higher than at T2 with a large effect size ($F(1,21)=51.47$, $p=.000$, $r=.84$).

The average gain on the SWRT test between T1 and T2 (baseline measure) was 2.05 (SD=4.60), and between T2 and T3 (after intervention), it was 6.68 (SD=4.37) (see Figure 14). A dependent (paired samples) t-test indicated that there was a statistically significant difference between the baseline and intervention periods with a large effect size (t (21) =2.80, p=.011, r=0.52). This is potential evidence that the intervention affected the participants' ability to read words without additional context.

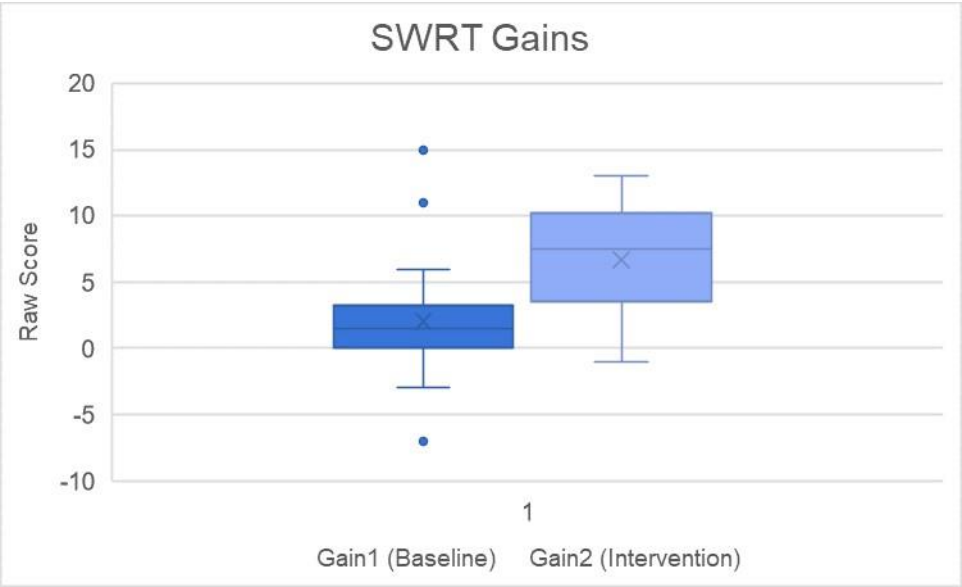

Figure 17: SWRT Gains

This data indicates that there was a potential effect on the participants ability to read words individually. It shows that there were larger gains over the period that included the intervention than the baseline period. This could reflect some improvement on participants' ability to read words over the period that the intervention covered and possibly an improvement in the participants' reading subskills.

6.1.3. Reading Rate (YARC RR)

The participants average scores increased over time from 58.95 (*SD*=13.74) at T1, to 63.45 (*SD*=12.17) at T2, and 66.68 (*SD*=10.57) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 15.

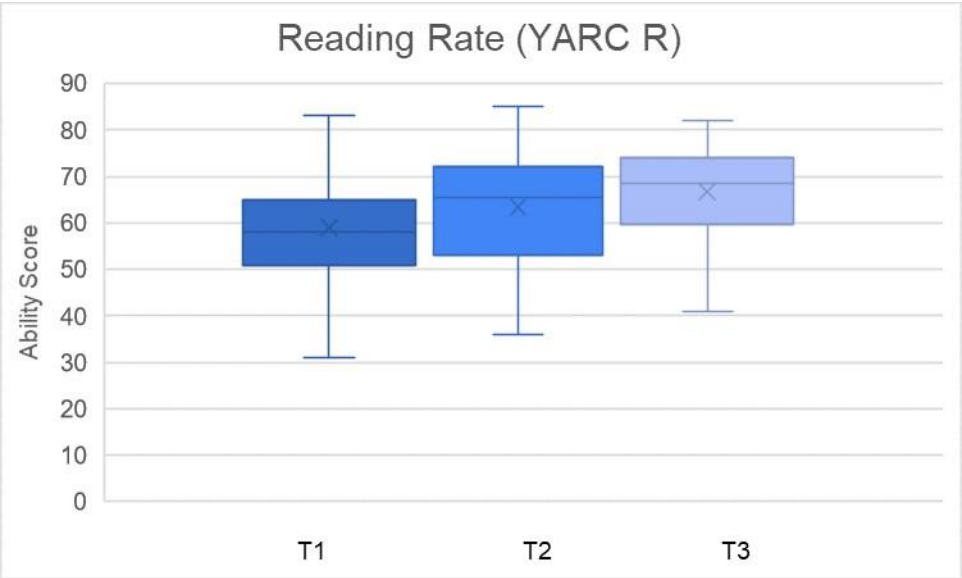

Figure 18:Reading Rate Scores by time

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 11.18$, $p=.000$. The speed at which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was significantly higher than at T1 ($F(1,21)=6.173$, $p=.021$, $r=.48$), and the average score at T3 was significantly higher than at T2 ($F(1,21)=9.150$, $p=.006$, $r=.55$). Both measures also demonstrated large effect sizes.

The average gain on the reading rate measure (YARC RR) between T1 and T2 (baseline measure) was 4.50 (SD=8.495), and between T2 and T3 (after intervention), it was 3.23 (SD=5.004) (see Figure 16). A dependent (paired samples) t-test indicated that this difference was not statistically significant ($t(21)=0.560$, $p=.58$, $r=0.12$). If this is taken in consideration with the repeated contrasts, it could be evidence that the repeated exposures to the measure did result in some effects but not statistically significant ones.

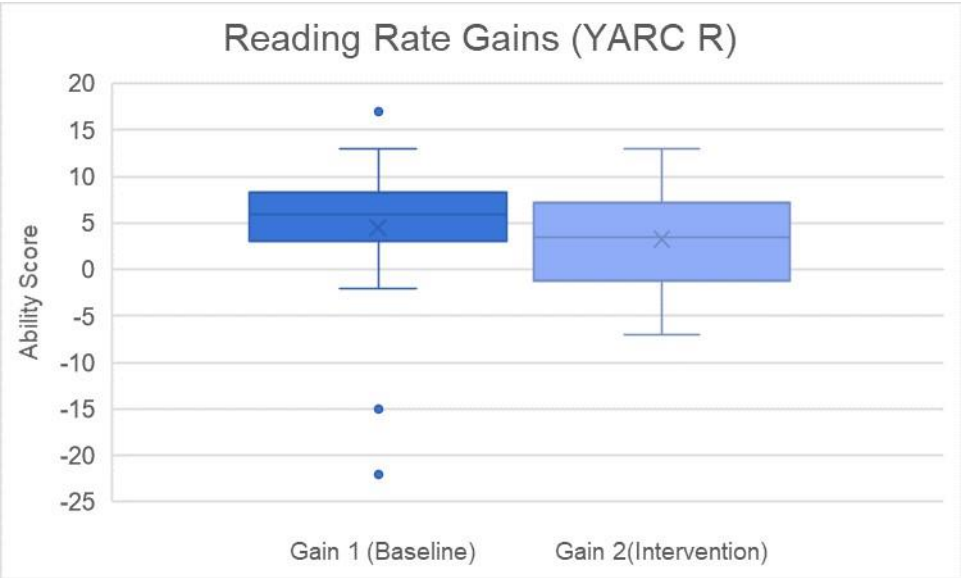

Figure 19:Reading Rate Gains

This analyses above demonstrate that while there may be some small effects to the repeated use of the same measures, they were not statistically significant. This illustrates that there is improvement it is at an approximately consistent pace between each point and potentially consistent with practice effects.

6.1.4. Reading Accuracy Measure (YARC A)

The measure of accuracy simply measures how correctly the participant could read the passage aloud. This could indicate if there were effects to additional reading skills from the intervention. It does not directly connect to the questions about reading comprehension or morphology but rather whether the measures were affected by their repeated usage. The participants mean scores decreased between times one (M=47.68, SD=7.026) and two (M=44.36, SD=6.374) and increased between two and three (M=52.64, SD=6.230). The distribution of scores at each time point and the change over time are summarized in Figure 17.

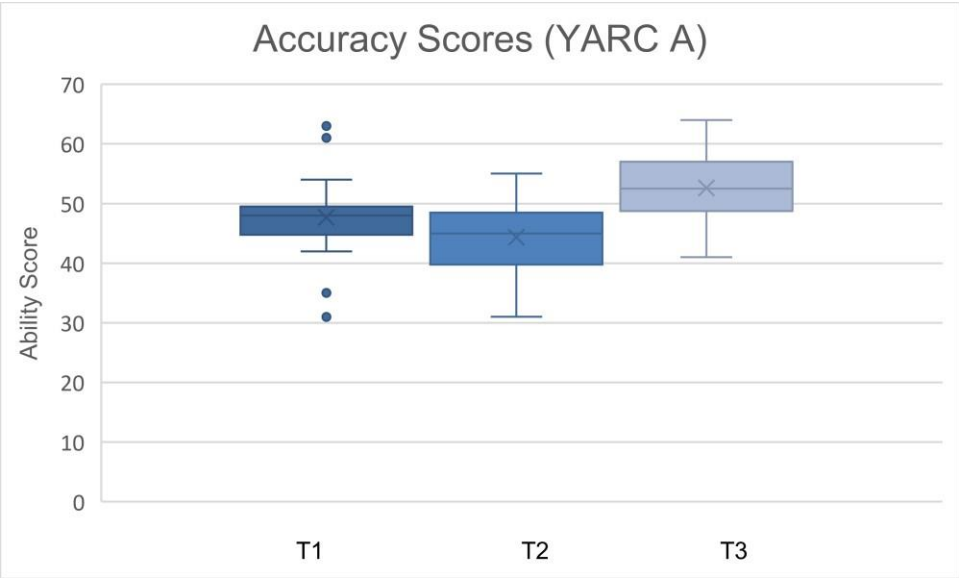

Figure 20:Accuracy scores by Time

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 18.507$, $p=.000$). The accuracy with which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was significantly lower than at T1 ($F(1,21)=5.996$, $p=.023$, $r=.47$), and the average score at T3 was significantly higher than at T2 ($F(1,21)=39.904$, $p=.000$, $r=.81$).

The average gain on the accuracy measure (YARC A) between T1 and T2 (baseline measure) was 3.32 (SD=6.357), and between T2 and T3(after intervention), it was 8.27 (SD=6.143) (see Figure 18). A dependent (paired samples) t-test indicated that this difference was statistically significant with a large effect size (t (21) =-5.164, p=.000, r=.75). A potential explanation for this could be the repeated exposure to the passages. However, this should be viewed with caution as the results at time two do not support this as those scores were significantly lower than at time one. This could have been a result of the relationship between the participants and the test administrator as all the measures and interventions were carried out by the researcher.

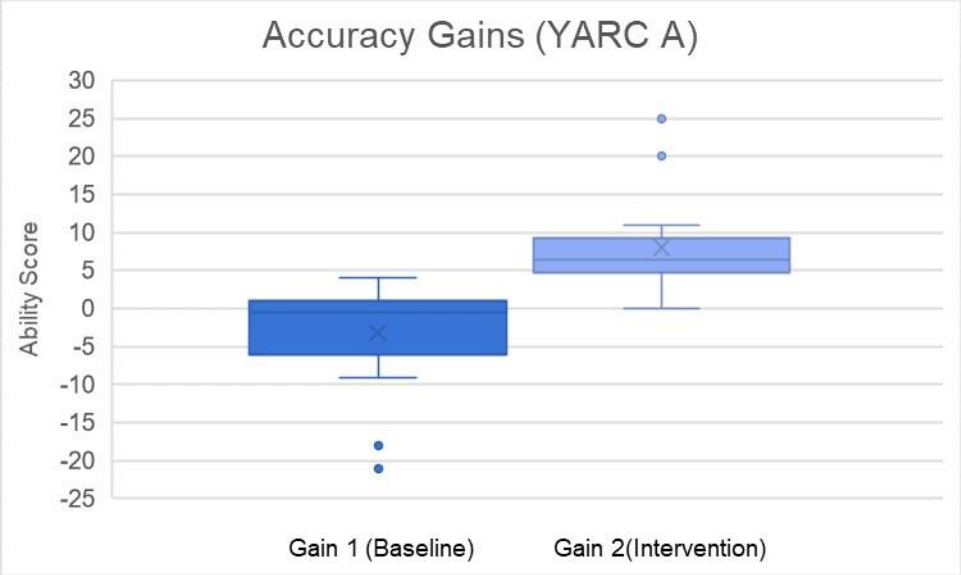

Figure 21:Accuracy Gains

6.1.5. Reading Comprehension (YARC RC)

The participants average scores increased over time from 50.77 (*SD*=5.66) at T1, to 52.09 (*SD*=6.80) at T2, and 64.23 (*SD*=6.44) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 19.

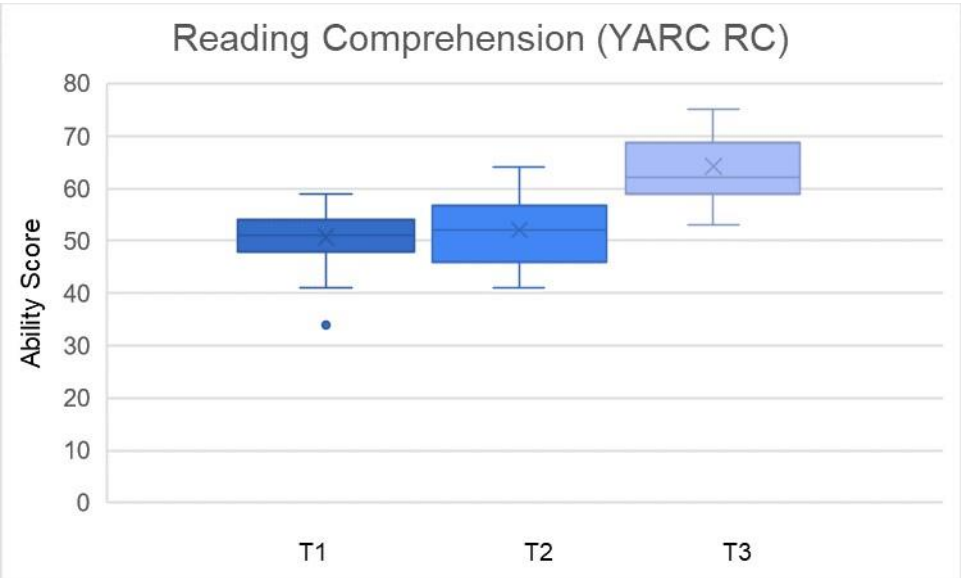

Figure 22:Reading Comprehension Scores

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 59.22$, $p=.000$. The comprehension of the participants on the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,21)=0.953$, $p=.340$, $r=.20$), and the average score at T3 was significantly higher than at T2 with a large effect size ($F(1,21)=70.255$, $p=.000$, $r=.88$).

The average gain on the reading comprehension measure (YARC RC) between T1 and T2 (baseline measure) was 1.32 (SD=6.34), and between T2 and T3 (after intervention), it was 12.14 (SD=6.79) (see Figure 20). A dependent (paired samples) t-test indicated that this difference was statistically significant with a large effect size (t (21)=-4.349, p=.000, r=0.69).

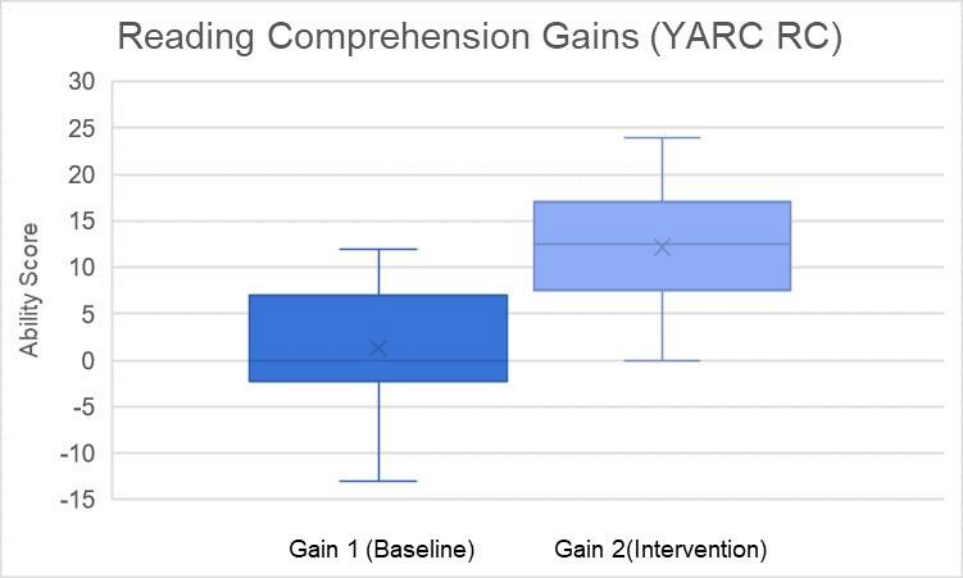

Figure 23:Reading Comprehension Gains

These results demonstrate a significant positive effect of the intervention on reading comprehension, though some caution should be used as there is potential for some of the effects to be caused by cumulative repeated exposures to the measures. The presence of statistically significant improvement shown by the ANOVA and by the t-test on the gains illustrates that there was a demonstratable effect of the intervention.

While there are improvements in reading comprehension, and this was shown by the measure of reading comprehension, the above measures illustrate that there was not a significant measured effect due to the repeated exposures to the measures. This supports the idea that this intervention has the potential to improve reading comprehension for primary students without taking into account the first language of the participant.

6.1.6. Morphology Measure (Modified Word Part Levels (WPLT)

The morphology measures are being utilized to answer the research question Is there a difference between EAL and L1 children in the acquisition of the meanings of derivational morphemes that are taught as part of a small group intervention. These measures looked at slightly different types of morphemes. The first measure looked solely at prefixes and suffixes and their meanings without context. The second measure looked at morphemes that also have roots in Greek and Latin and are often used as root words where participants were given the meaning and had to select the correct word containing the target morpheme.

Prefixes and suffixes are more predictable as they have fixed positions and thus students are more likely to be able to isolate them from the rest of a morphologically complex word and recall the meaning that is carried. Root words can be found in multiple positions of a word and thus are not as easily identified.

The participants average scores increased over time from 6.50 (*SD*=3.05) at T1, to 7.50 (*SD*=3.28) at T2, and 9.64 (*SD*=3.92) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 21.

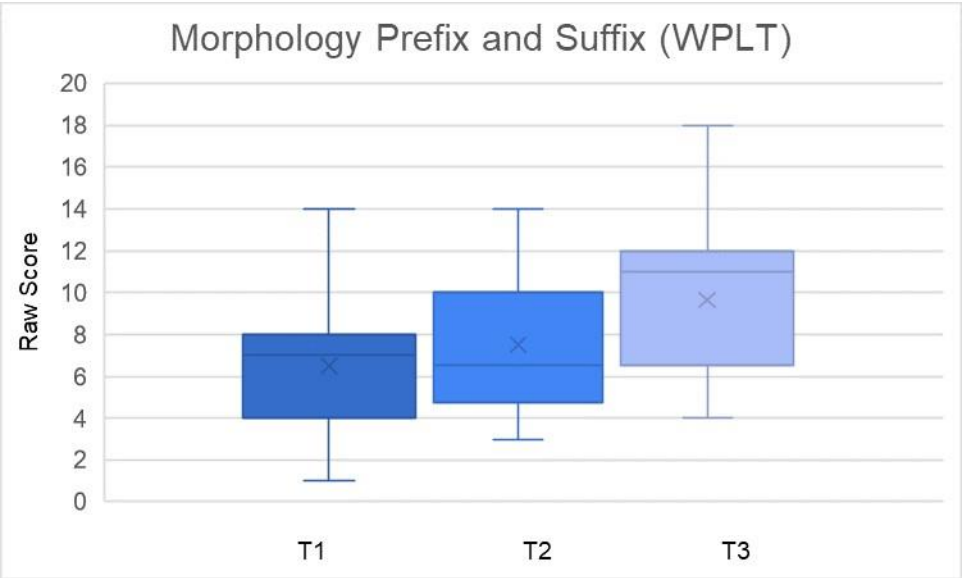

Figure 24:Morphology Measure (Prefixes and Suffixes)

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 9.839$, $p=.000$. The ability of students to identify meanings of prefixes and suffixes increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,21)=2.750$, $p=.112$, $r=.20$), but the average score at T3 was significantly higher than at T2 with a large effect size ($F(1,21)=9.066$, $p=.007$, $r=.88$).

The average gain on the Morphology, prefixes and suffixes between T1 and T2 (baseline measure) was 1.00 (SD=2.828), and between T2 and T3 (after intervention), it was 2.14 (SD=3.33) (see Figure 22). A dependent (paired samples) t-test indicated that this difference was not statistically significant but had a large effect size (t (21) =-1.116, p=.277, r=.69). These results taken together show there was some impact of the intervention on participants understanding of prefixes and suffixes, even if it did not reach the threshold of statistical significance.

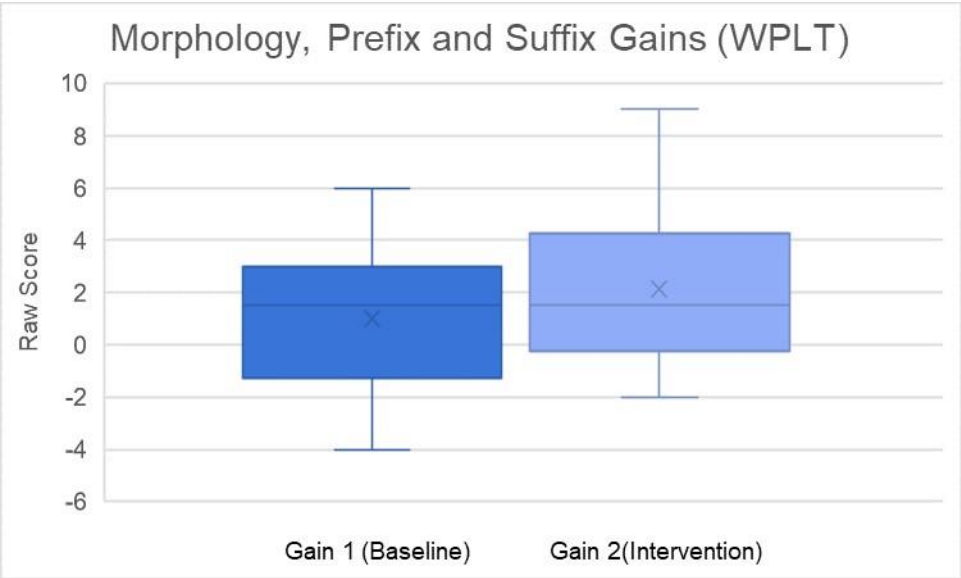

Figure 25:Morphology Measure Gains (Prefixes and Suffixes)

6.1.6.1. Taught Morphemes WPLT

Taking a closer look at both the taught and untaught morphemes demonstrates the participants average scores for taught morphemes increased over time from 4.45 (*SD*=2.32) at T1, to 5.41 (*SD*=2.81) at T2, and 6.91 (*SD*=2.63) at T3 The distribution of scores at each time point and the change over time are summarized in Figure 23

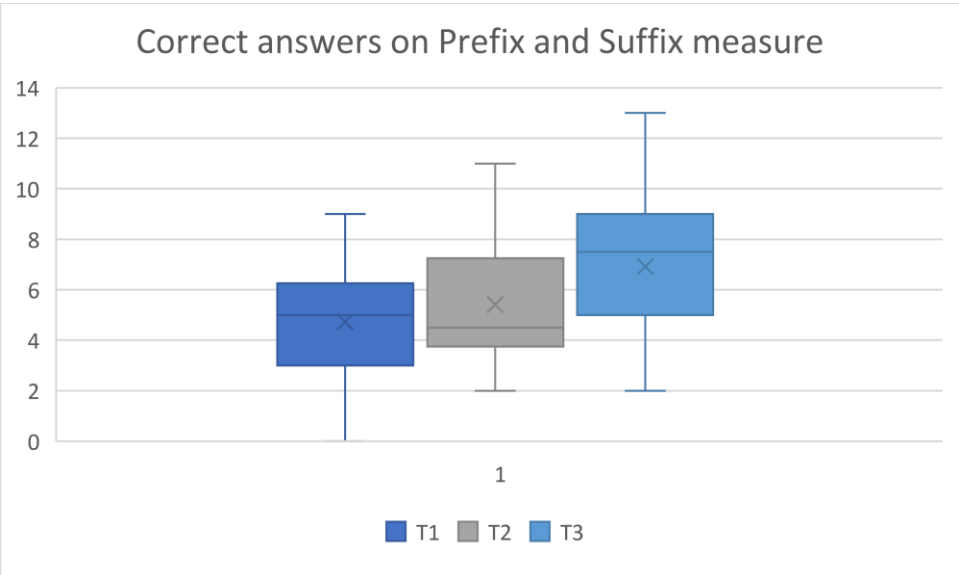

Figure 26:Taught Morphology; Prefix and Suffix correct answers out of 13

A one way repeated measure ANOVA confirmed this change was statistically significant for the taught items: $F(2,42)= 13.626$, $p=.000$. The ability of students to identify meanings of the taught

prefixes and suffixes increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 with a small effect size ($F(1,21)=0.888$, $p=.357$, $r=.20$), but the average score at T3 was significantly higher than at T2 with a medium effect size ($F(1,21)=16.399$, $p=.001$, $r=.66$

Together these results would suggest that there were gains in knowledge of prefix and suffix meanings on the taught morphemes. They also suggest that repeated exposures to the measures did not have a significant impact on participant knowledge or performance as there were no significant increases on untaught measures or effect sizes greater than r=.15 for any of these tests. The results for the taught measures indicate that there were significant gains during the intervention period in relation to the baseline period with large effect sizes of r= .66 for both tests.

6.1.7. Morphology Measure Roots

The participants average scores increased over time from 7.27(*SD*=2.66) at T1, to 6.64 (*SD*=2.84) at T2, and 8.82 (*SD*=2.99) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 24.

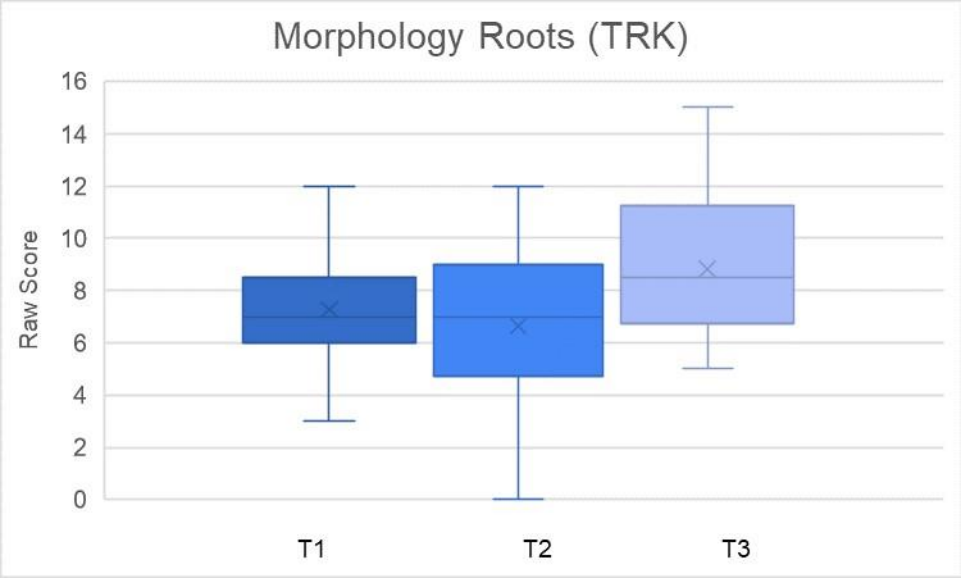

Figure 27:Morphology Roots correct answers out of 15 questions

A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 5.404$, $p=.008$. The ability of students to identify meanings of roots increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,21)=0.695$, $p=.414$, $r=.18$), and the average score at T3 was significantly higher than at T2 and had a large effect size ($F(1,21)=9.191$, $p=.006$, $r=.55$).

The average gain on the Morphology, roots between T1 and T2 (baseline measure) was -0.64(SD=3.58), and between T2 and T3(after intervention), it was 2.18 (SD=3.38) (see Figure 25). A dependent (paired samples) t-test indicated that this difference was not statistically significant but had a medium effect size (t (21) =-2.043, p=.054, r=0.40). Together these statistics demonstrate potential positive effects of the intervention on participants' understanding of roots and on their understanding of words with related roots.

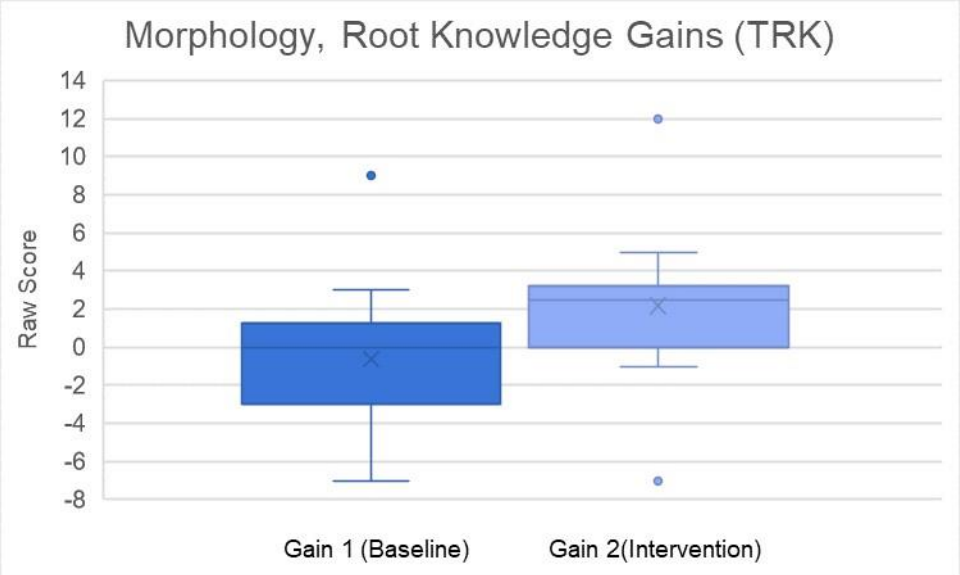

Figure 28:Morphology, Root Knowledge Gains

6.1.7.1. Taught Morphemes (Roots)

Taking a closer look at the taught morphemes demonstrates that the participants average scores for morphemes included in the intervention decreased from 3.59 (*SD*=1.68) at T1, to 3.27(*SD*=1.86) at T2, but increased to 4.64 (*SD*=1.76) at T3 (i.e., after the intervention). The distribution of scores at

each time point and the change over time are summarized in Figure 26. This demonstrates that there were gains in knowledge of roots, for those taught in the intervention.

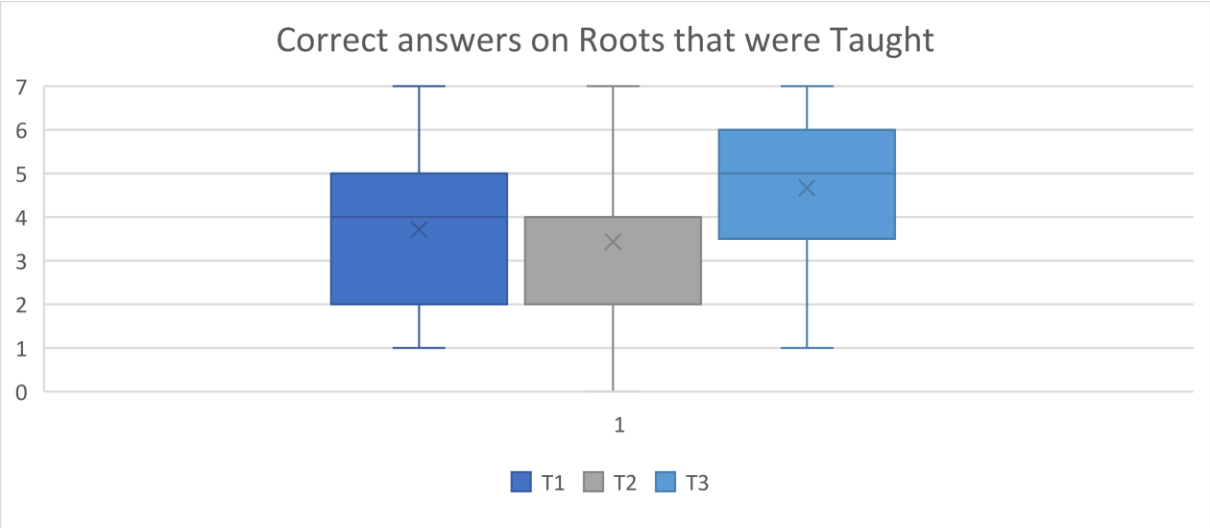

Figure 29: Taught Morphology Measure; Roots out of a possible 7 questions

A one way repeated measure ANOVA confirmed this change was statistically significant for the taught items: F(2,42)=12.946, p=.000. The ability of students to identify meanings of roots increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 (F(1,21)=0.488, p=.492, r=.20), but the average score at T3 was significantly higher than at T2 with a medium effect size (F(1,21)=37.580, p=.000, r=.66). A dependent (paired samples) t-test indicated that the difference between T1 and T2 was not statistically significant but had a small effect size (t (21) =0.699, p=.492, r=0.20). A dependent (paired samples) t-test indicated that the difference between T2 and T3 was statistically significant and had a medium effect size (t (21) =6.130, p=.000, r=0.66).

If we consider these results, it shows that the participants were successful at learning the meaning of roots from the intervention over the time covered in the study. This additionally suggests that repeated exposures to the measures did not have a significant impact on participant knowledge or performance. These results indicate that there are potential positive effects for participants' knowledge of morphological roots from the intervention.

6.2. EAL and L1E Participant Data

Details about the 11 EAL participants and 11 L1E participants can be found in Table 2. All tests and analysis were carried out following the same method as used for the whole group measures. If the data was found to violate Sphericity the Greenhouse- Geiser correction was applied. If this is not mentioned, it can be assumed that data is normally distributed.

6.2.1. Receptive Vocabulary (BPVS)

The EAL participants' average scores on the BPVS test increased with time, from 88.64 (SD=12.53) at T1, to 97.27 (SD=9.25) at T2 and 103.82 (SD=14.06) at T3. The L1E participants' average scores on the BPVS test also increased with time, from 86.82 (SD=11.78) at T1, to 98.18 (SD=14.26) at T2 and 105.55 (SD=15.04) at T3. The distribution of scores at each time point, and the change over time are summarised in Figure 27.

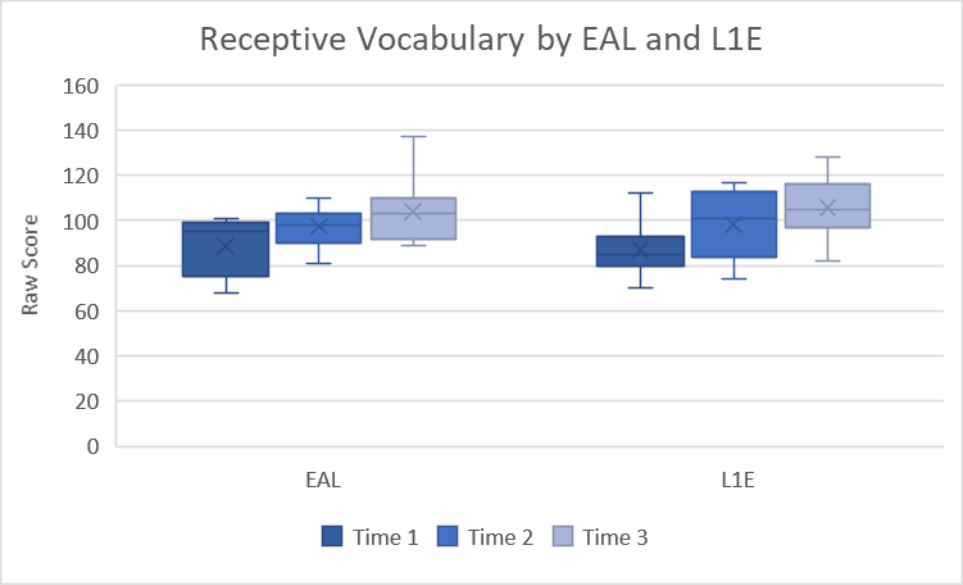

Figure 30:EAL and L1E Receptive Vocabulary Score

For EAL participants, repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=4.492$, $p=.060$, $r=.56$), and the average score at T3 was also not significantly higher than at T2, although there was a medium effect size ($F(1,10)=2.498$, $p=.145$ $r=.44$). However, a one-way repeated measure ANOVA confirmed that the change from T1 to T3 was

statistically significant: F(2,20)=8.124, p=.003. The size of receptive vocabulary increased as time passed.

The average gain on the BPVS test between T1 and T2 (baseline measure) was 8.64 (SD=13.52), and between T2 and T3 (after the intervention), it was 6.55 (SD=13.74) (see Figure 28). A dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10) =0.274, p=.790, r=.09), with larger gains at baseline than after the intervention.

For L1E participants, repeated contrasts revealed that the average score at T2 was significantly higher than at T1 (F(1,10)=11.450, p=.007, r=.73), and the average score at T3 was also significantly higher than at T2 (F(1,10)=12.730, p=.005 r=.75). Although, a one-way repeated measure ANOVA also confirmed that the change from T1 to T3 was statistically significant: F(2,20)=18.513, p=.000. The size of receptive vocabulary increased as time passed.

The average gain on the BPVS test between T1 and T2 (baseline measure) was 11.36 (SD=11.14), and between T2 and T3 (after the intervention), it was 7.36(SD=6.85) (see Figure 28). A dependent

(paired samples) t-test indicated that this difference was not statistically significant (t (10) =0.949, p=.365, r=.29). This indicates that there was not an effect on receptive vocabulary for L1E participants due to the intervention.

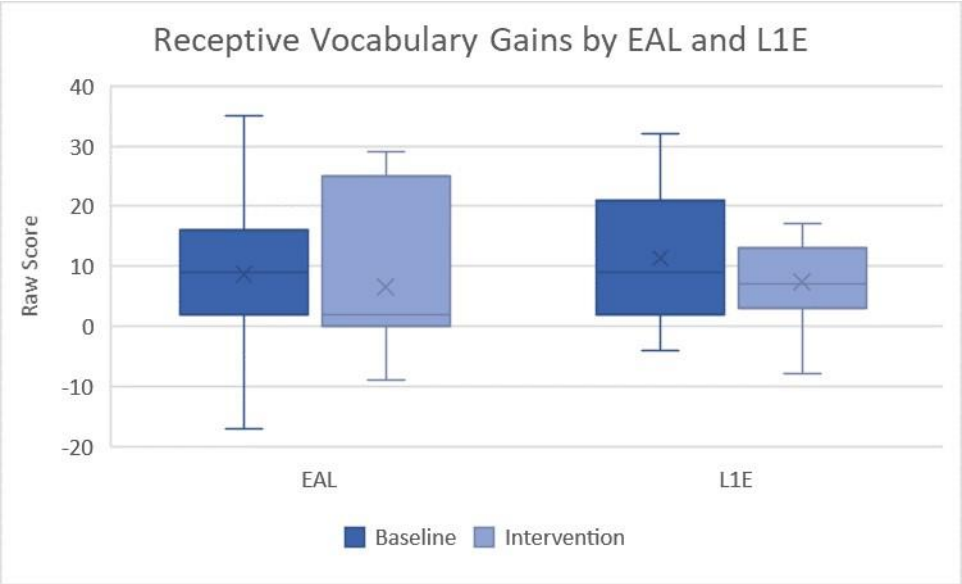

Figure 31: Receptive Vocabulary Gains by EAL and L1E

6.2.1.1. Summary

Both EAL and L1E participants showed statistically significant gains on the measure of receptive vocabulary over time based on the one-way repeated ANOVA. The L1E participants displayed changes with statistical significance on the repeated contrasts with large effect sizes while the EAL did not, although they did show medium sized effects. Neither group showed significant differences on the paired samples t-test conducted on the gains scores. This illustrates that there may be potential for larger impacts of the intervention on L1E participants' receptive vocabulary, which could then impact their reading comprehension in a positive manner.

6.2.2. Single Word Reading Test (SWRT)

The EAL participants average scores increased over time from 32.91 (*SD*=6.19) at T1, to 34.64 (*SD*=7.66) at T2, and 42.18 (*SD*=8.32) at T3. The L1E participants average scores increased over time from 33.64 (*SD*=8.98) at T1, to 36.00 (*SD*=8.21) at T2, and 41.82 (*SD*=9.21) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 29.

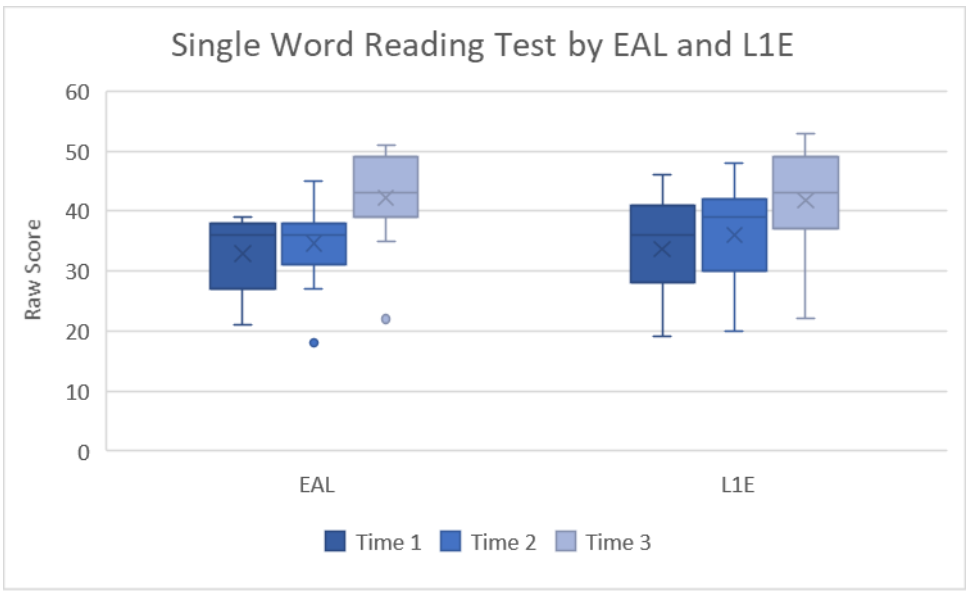

Figure 32: Single Word Reading by EAL and L1E

For EAL participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 20.185$, $p=.000$. The ability to read single words increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=0.982$, $p=.345$, $r=.299$), and the average score at T3 was significantly higher than at T2 with a large effect size ($F(1,10)=32.495$, $p=.000$, $r=.87$).

The average gain on the SWRT test between T1 and T2 (baseline measure) was 1.73 (SD=5.78), and between T2 and T3(after intervention), it was 7.55 (SD=4.39) (see Figure 30). A dependent (paired samples) t-test indicated that this difference was not statistically significant yet there was a large effect size (t (10) =2.17, p=.055, r=0.57). The large effect size with a p-value close to significance and a very small sample size could indicate a trend of improvement on word reading with EAL participants in a positive direction.

With the L1E participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 25.785$, $p=.000$. The ability to read single words increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=5.662$, $p=.039$, $r=.60$), and the average score at T3 was significantly higher than at T2 ($F(1,10)=19.431$, $p=.001$, $r=.81$).

The average gain on the SWRT test between T1 and T2 (baseline measure) was 2.36 (SD=3.30), and between T2 and T3(after intervention), it was 5.83(SD=4.38) (see Figure 30). A dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10) =1.714, p=.117, r=0.48). These results show evidence of potential positive results of the intervention on word reading with L1E participants.

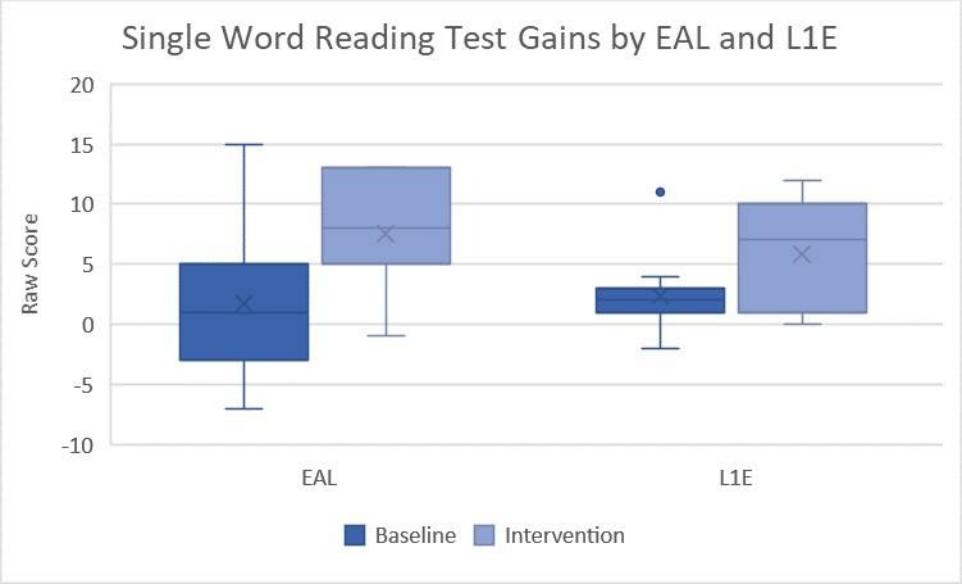

Figure 33: EAL and L1E Single Word Reading Gains

6.2.2.1. Summary

Both groups showed statistically significant gains on the measure of single word reading based on the one-way repeated measures ANOVA, additionally they both were significant on the repeated contrasts for the intervention period with large effect sizes. When examining the gains, there were differences between the groups: the EAL group had a large effect size and significance whereas the L1E group only had a medium effect size and did not reach significance. Potentially there was more impact on the single word reading from the intervention. There is also the risk that the repeated use of the same measure had a larger practice effect on the EAL group than then L1E group. This would need to be taken into account when looking at the overall potential for improvement in reading comprehension, as there should be some caution taken with the repeated use of measures.

6.2.3. Reading Rate (YARC RR)

The EAL participants average scores increased over time from 58.73 (*SD*=15.03) at T1, to 62.82 (*SD*=12.13) at T2, and 66.00 (*SD*=11.37) at T3. The L1E participants average scores increased over time from 59.18 (*SD*=13.05) at T1, to 64.09 (*SD*=12.76) at T2, and 67.36 (*SD*=10.22) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 31.

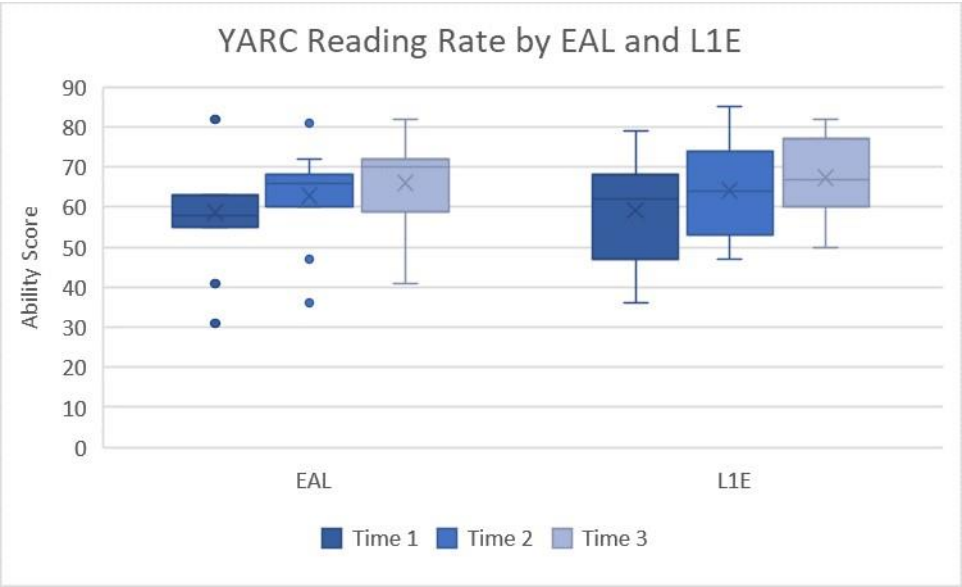

Figure 34:YARC Reading Rate by EAL and L1E

With EAL participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 3.525$, $p=.049$. The speed at which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 with a medium effect size ($F(1,10)=2.071$, $p=.181$, $r=.41$), and the average score at T3 was not significantly higher than at T2 with a large effect size ($F(1,10)=4.726$, $p=.055$, $r=.57$). This potentially could be due to a practice effect of participants seeing the measures more than once even with more than two months between usage.

The average gain on the reading rate measure (YARC RR) between T1 and T2 (baseline measure) was 4.09(SD=9.43), and between T2 and T3(after intervention), it was 3.18 (SD=4.85) (see Figure 32). A dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10)=0.321, p=.75, r=0.10).

For L1E participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 9.050$, $p=.002$. The speed at which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=4.256$, $p=.066$, $r=.55$), and the average score at T3 was not significantly higher than at T2 ($F(1,10)=4.060$, $p=.072$, $r=.54$).

The average gain on the reading rate measure (YARC RR) between T1 and T2 (baseline measure) was 4.91(SD=7.89), and between T2 and T3(after intervention), it was 3.27 (SD=5.39) (see Figure 32). A

dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10) =0.443, p=.667, r=0.14). This seems to illustrate that there is not a practice effect for L1E participants with regards to reading rate over the three administrations of the measure.

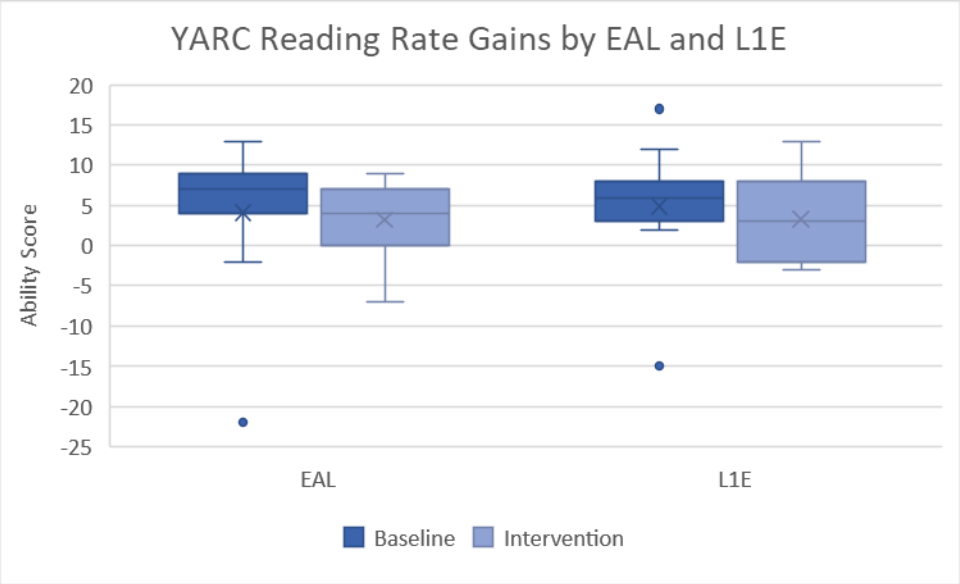

Figure 35:YARC Reading Rate Gains by EAL and L1E

6.2.3.1. Summary

Both EAL and L1E participants showed positive gains and significant results on the one-way repeated measure ANOVA on reading rate, however repeated contrasts showed no significance for either group or period. The dependent paired samples t-tests for both groups did not indicate statistically significant results or effect sizes. This illustrates that while there may be some effects from the repeated use of measures it may not be significant.

6.2.5. Reading Comprehension (YARC RC)

The EAL participants average scores increased over time from 49.91 (*SD*=7.41) at T1, to 51.82 (*SD*=6.88) at T2, and 63.27 (*SD*=7.30) at T3. The L1E participants average scores increased over time from 51.64 (*SD*=3.30) at T1, to 52.36 (*SD*=7.05) at T2, and 65.18 (*SD*=5.65) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 35.

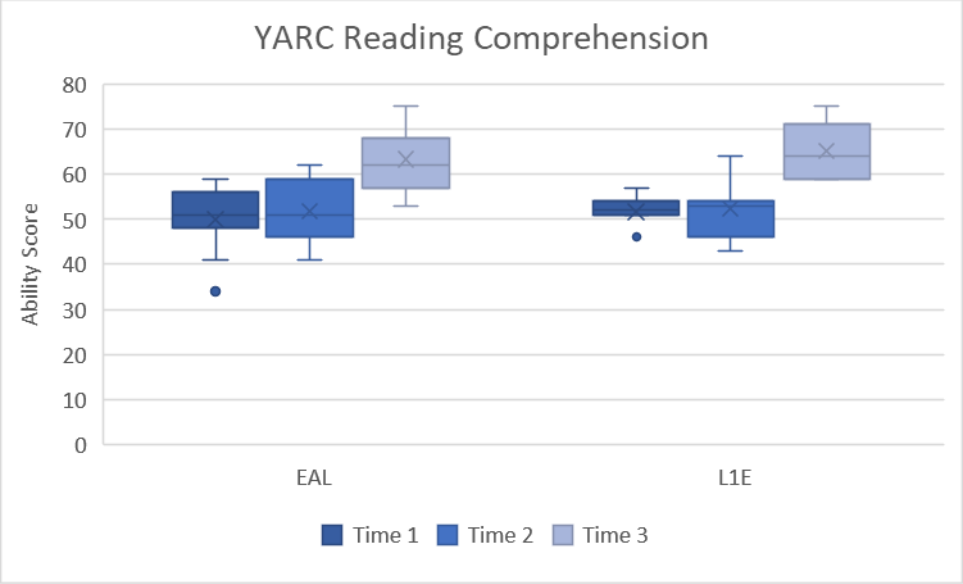

Figure 36:YARC Reading Comprehension by EAL and L1E

For EAL participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 19.390$, $p=.000$. The comprehension of the participants on the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=0.697$, $p=.423$, $r=.255$), and the average score at T3 was significantly higher than at T2 with a large effect size ($F(1,10)=20.656$, $p=.001$, $r=.82$).

The average gain on the reading rate measure (YARC RR) between T1 and T2 (baseline measure) was 1.91(SD=7.58), and between T2 and T3(after intervention), it was 11.45 (SD=8.36) (see Figure 36). A dependent (paired samples) t-test indicated that this difference was not statistically significant though it has a large effect size (t (10) 2.215, p=.051, r=0.57). This shows that there is a potential positive effect from the intervention on the reading comprehension of EAL participants beyond the effects of regular instruction.

With L1E participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 49.213$, $p=.000$. The reading comprehension of the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=0.224$, $p=.646$, $r=.15$), and the average score at T3 was significantly higher than at T2 with a very large effect size ($F(1,10)=69.611$, $p=.000$, $r=.94$).

The average gain on the reading comprehension measure (YARC RC) between T1 and T2 (baseline measure) was 0.73(SD=5.10), and between T2 and T3 (after intervention), it was 12.82 (SD=5.10) (see Figure 36). A dependent (paired samples) t-test indicated that this difference was statistically significant (t (10) 4.538, p=.001, r=.82). These results signify a possible positive effect of the intervention on L1E participants reading comprehension beyond practice effects.

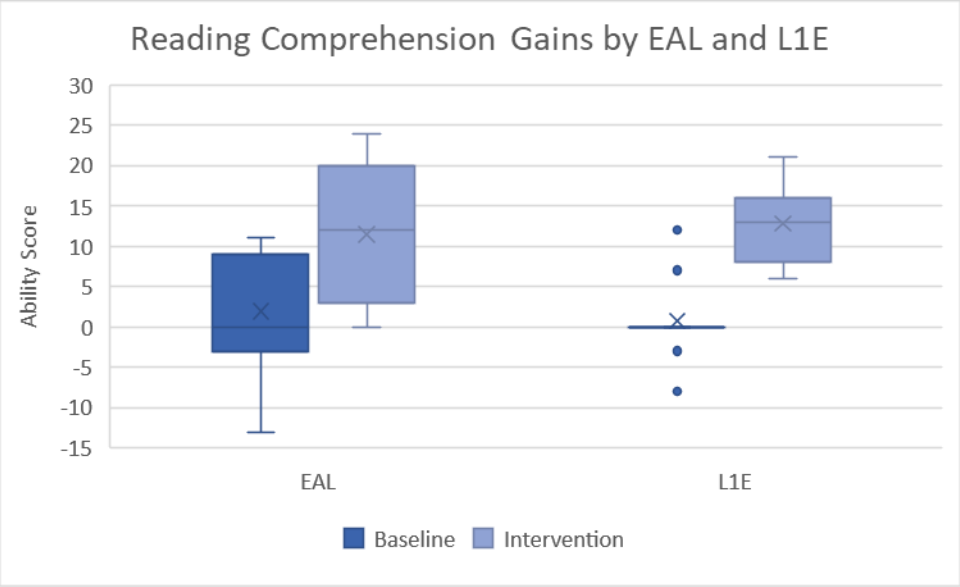

Figure 37:Reading Comprehension Gains by EAL and L1E

6.2.5.1. Summary

There were positive effects on both EAL and L1E participants on the reading comprehension measure. The one-way ANOVAs both showed significant gains and repeated contrasts showed significant gains at time three with large effect sizes. The dependent t-tests indicated statistically significant increases on the gains for L1E participants, though the EALs had a large effect size they did not meet the threshold for significance. This suggests strong evidence for positive effects on reading comprehension for both groups from the intervention though there is potentially more of an impact on L1E participants than EAL. This ties directly to the research question about reading comprehension. It illustrates a potential positive effect for both EAL and L1E groups with a stronger and significant effect on the L1E groups.

6.2.4. Accuracy Measure (YARC A)

The participants mean scores decreased between times one (M=46.91, SD=7.49) and two (M=44.45, SD=4.89) and increased between two and three (M=52.18, SD=6.21). The participants mean scores decreased between times one (M=48.45, SD=6.80) and two (M=44.27, SD=7.84) and increased between two and three (M=53.09, SD=6.52). The distribution of scores at each time point and the change over time are summarized in Figure 33.

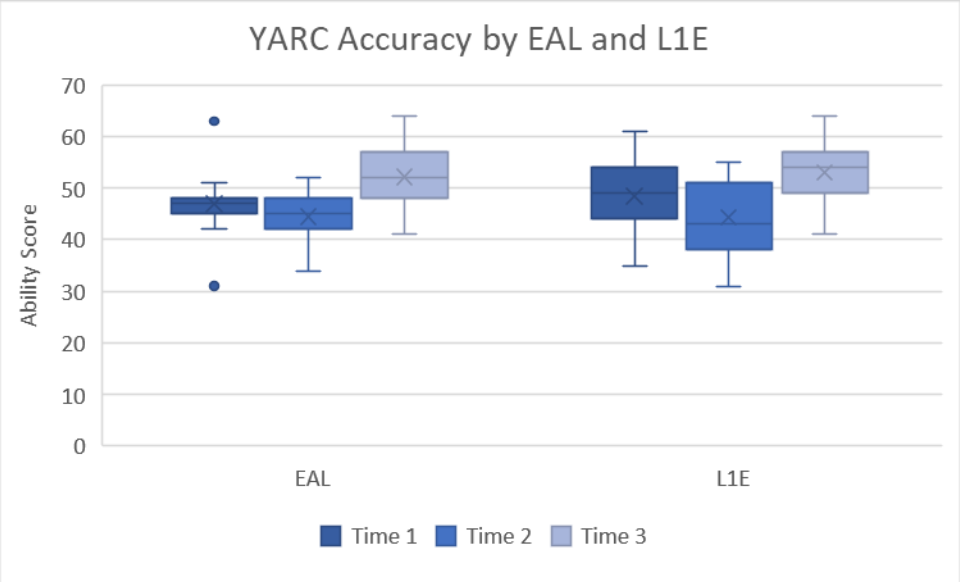

Figure 38:YARC Accuracy by EAL and L1E

For EAL participants, a one-way repeated measure ANOVA confirmed this change was statistically significant: F (2,20) = 6.685, p=.006. The accuracy with which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly lower than at T1 (F(1,10)=1.334, p=.275, r=.34), and the average score at T3 was significantly higher than at T2 with a large effect size (F(1,10)=17.840, p=.002, r=.80).

The average gain on the accuracy measure (YARC A) between T1 and T2 (baseline measure) was 2.45 (SD=7.05), and between T2 and T3(after intervention), it was 7.73 (SD=6.07) (see Figure 34). A dependent (paired samples) t-test indicated that this difference was statistically significant with a large effect size (t (10) =3.287, p=.008, r=0.72). Considering both the repeated contrast results and the results of the t-tests on gains part of the large effect could be due to repeated use of the same measure but, it does not explain all of it. Therefore, part could be due to the intervention.

For L1E participants, a one-way repeated measure ANOVA confirmed this change was statistically significant: F (2,20) = 12.446, p=.000. The accuracy with which participants read the passages increased over time. Repeated contrasts revealed that the average score at T2 was not significantly lower than at T1 (F (1,10) =5.731, p=.038, r=.60), and the average score at T3 was significantly higher than at T2 (F (1,10) =20.481, p=.001, r=.82).

The average gain on the accuracy measure (YARC A) between T1 and T2 (baseline measure) was 4.18 (SD=5.80), and between T2 and T3(after intervention), it was 8.82 (SD=6.48) (see Figure 34). A dependent (paired samples) t-test indicated that this difference was statistically significant (t (10) =3.891, p=.003, r=0.78). These results seem to indicate an effect beyond a practice effect for the accuracy of L1E participants on account of the intervention.

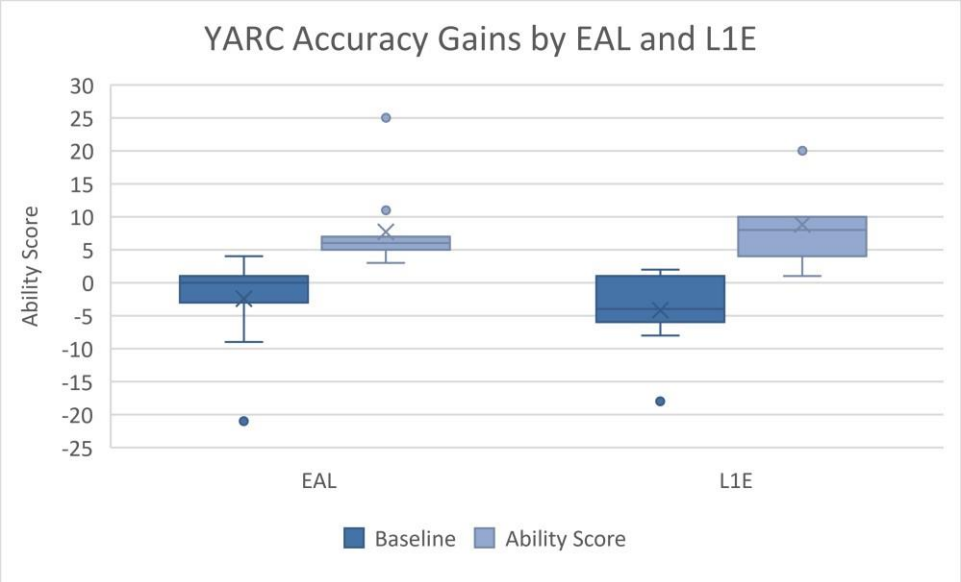

Figure 39: YARC Accuracy Gains by EAL and L1E

6.2.4.1. Summary

The results for this measure were very similar for both groups. Both had statistically significant one-way ANOVA results and repeated contrasts that were significant with large effect sizes at the intervention period. They also had very similar results on the dependent (paired samples) t-tests on the gains. The improvement in accuracy in reading could improve the comprehension of the participants

as they would not be mistaking similarly spelled words for each other and this may in turn make the cognitive load lower and improve their comprehension.

6.2.6. Morphology Measure (Prefixes and Suffixes)

The EAL participants average scores increased over time from 7.00 (*SD*=2.41) at T1, to 7.27 (*SD*=3.07) at T2, and 9.36 (*SD*=3.01) at T3. The L1E participants average scores increased over time from 6.00 (*SD*=3.63) at T1, to 7.73 (*SD*=3.61) at T2, and 9.91 (*SD*=4.81) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 37.

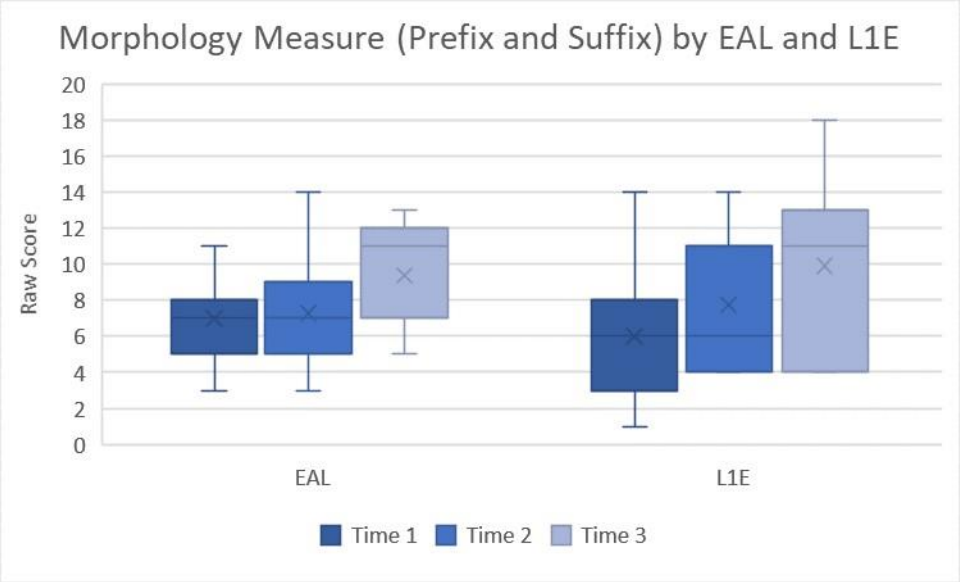

Figure 40:Morphology Measure: Prefix and Suffix by EAL and L1E

With the EAL participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 3.635$, $p=.045$. The ability of students to identify meanings of prefixes and suffixes increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=0.095$, $p=.764$, $r=.097$), and the average score at T3 was not significantly higher than at T2 though it had a large effect size ($F(1, 10)=3.977$, $p=.074$, $r=.53$).

The average gain on the Morphology, prefixes and suffixes between T1 and T2 (baseline measure) was 0.27(SD=2.94), and between T2 and T3(after intervention), it was 2.09 (SD=3.48) (see Figure 38). A dependent (paired samples) t-test indicated that this difference was not statistically significant though it has a medium effect size (t (10) =1.070, p=.310, r=0.32). This suggests a potential increase in participant knowledge of affixes.

For the L1E participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,20)= 6.416$, $p=.007$. The ability of students to identify meanings of prefixes and suffixes increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=4.676$, $p=.056$, $r=.56$), and the average score at T3 was not significantly higher than at T2 ($F(1, 10)=4.691$, $p=.056$, $r=.57$).

The average gain on the Morphology, prefixes and suffixes between T1 and T2 (baseline measure) was 1.73 (SD=2.65), and between T2 and T3(after intervention), it was 2.18 (SD=3.34) (see Figure 38). A dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10) =0.388, p=.706, r=0.12). These results indicate that there was not a significant effect on L1E participants knowledge of affixes, though there may be a slight increase in that knowledge.

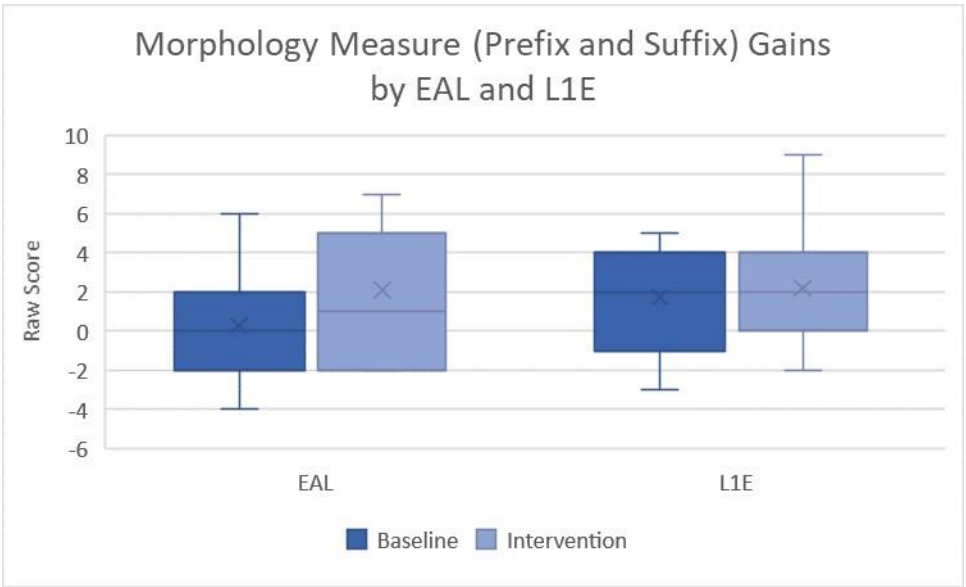

Figure 41:Morphology Measure: Prefix and Suffix Gains by EAL and L1E

6.2.6.1 Summary

Both groups showed significant changes on the one-way ANOVA, the repeated measures did not show significance for either period or group, both showed large effect sizes at time three.

Additionally, on the dependent (paired samples) t-test on the gains were not statistically significant. There were however small (L1E) and medium (EAL) effect sizes. This shows potential effects of the intervention in a positive direction. This would suggest that there are positive effects on participants understanding of prefixes and suffixes from the intervention which ties to the research question about understanding of morphology.

6.2.7 . Morphology Measure Roots

The EAL participants average scores increased over time from 7.55 (*SD*=2.58) at T1, to 6.27 (*SD*=2.41) at T2, and 8.18 (*SD*=2.32) at T3. The L1E participants average scores increased over time from 7.00 (*SD*=2.83) at T1, to 7.00 (*SD*=3.29) at T2, and 9.45 (*SD*=3.42) at T3. The distribution of scores at each time point and the change over time are summarized in Figure 39.

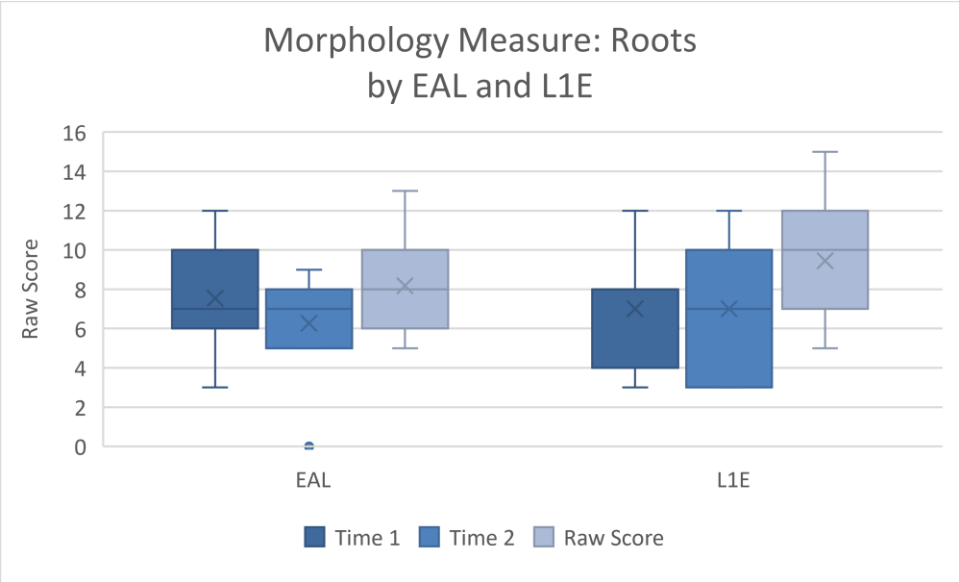

Figure 42:Morpheme Measure: Roots by EAL and L1E

For EAL participants, a one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 5.404$, $p=.008$. The ability of students to identify meanings of roots increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,21)=0.695$, $p=.414$, $r=.18$), and the average score at T3 was significantly higher than at T2 and had a large effect size ($F(1,21)=9.191$, $p=.006$, $r=.55$).

The average gain on the Morphology, prefixes and suffixes between T1 and T2 (baseline measure) was -0.64 (SD=3.58), and between T2 and T3(after intervention), it was 2.18 (SD=3.38) (see Figure 40). A dependent (paired samples) t-test indicated that this difference was not statistically significant but had a medium effect size (t (21) =-2.043, p=.054, r=0.40). Together these statistics demonstrate potential positive effects of the intervention on EAL participants understanding of roots. Though this could potentially be as a result of the negative average gains for participants at time 2 in comparison to time 1.

With L1E participants, a one way repeated measure ANOVA confirmed this change was not statistically significant: $F(2,20)= 3.030$, $p=.071$. The ability of students to identify meanings of roots increased over time. Repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 ($F(1,10)=0.000$, $p=1.00$, $r=.00$), and the average score at T3 was not significantly higher than at T2 ($F(1,10)=3.335$, $p=.098$, $r=.50$).

The average gain on the Morphology, prefixes and suffixes between T1 and T2 (baseline measure) was 0.00(SD=4.38), and between T2 and T3(after intervention), it was 2.45 (SD=4.46) (see Figure 40). A dependent (paired samples) t-test indicated that this difference was not statistically significant (t (10) =-0.950, p=.365, r=0.29). This result suggests that there was a medium sized effect on the L1E participants' knowledge of root meanings as a result of the intervention, even though the growth did not reach the level of statistical significance.

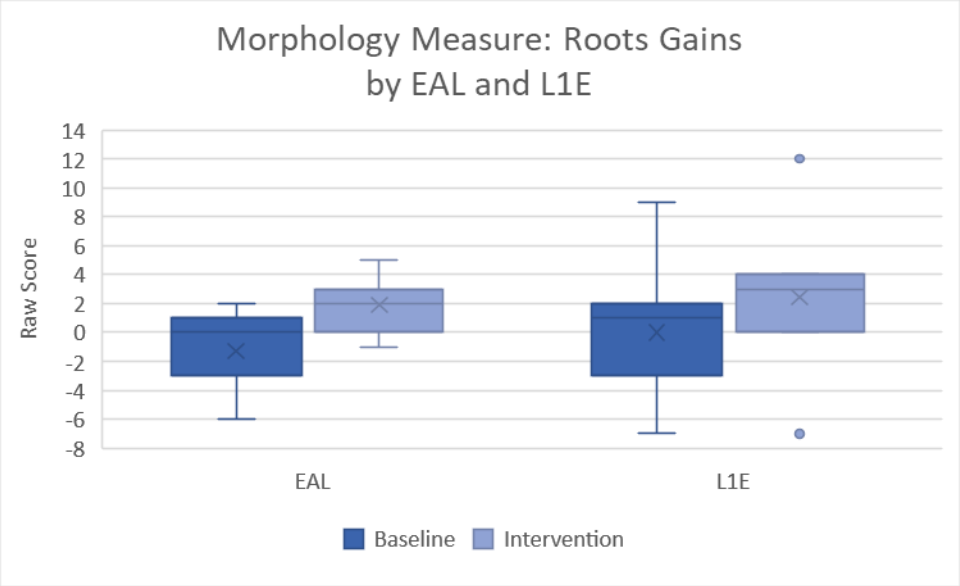

Figure 43: Morphology Measure: Roots Gains by EAL and L1E

6.2.7.1 Summary

According to the repeated measure ANOVA there were not statistically significant gains on the measure of roots. However, for the EAL participants the repeated contrasts revealed significant increases with a large effect size at time three (T3). This was additionally confirmed by the dependent (paired samples) t-test results that were both statistically significant and had large effect size. The L1E participants showed positive trends due to the intervention but, even though they did not meet the threshold of statistical significance, there was a medium effect size. This could indicate a positive relation between the intervention and the ability of the participants to understand the meanings of roots and other morphemes. The significant results for the EAL population would indicate that the intervention had a larger positive effect for them than for the L1E participants, although there were positive effects for both. This will be discussed more in depth in the discussion chapter.

6.3 Models to answer Research questions.

6.3.6 Differences in EAL and L1 children in effectiveness of teaching semantic meanings of derivational morphemes.

Is there a difference between EAL and L1 children in the effectiveness of teaching the semantic meaning of derivational morphemes? Is there an effect of individual differences in gender, attendance, and age in the effectiveness of teaching the semantic meaning of derivational morphemes?

In order to address RQ1, scores on the two tests of semantic morphology knowledge were compared between the two learner groups in two separate mixed regression models, the first testing the knowledge of roots (scores on the TRK) and the second testing the participants' knowledge of affixes (scores on the WPLT).The predictor variable TimeT1 was the intercept, as the models used R's default treatment coding. It was assumed that increased scores over the three testing periods throughout the intervention. Table 8. shows the mean scores of the participants on the two tests of knowledge of semantic morphology.

Affixes

Table 8: Affix Means and SD

			Affix	(sd)
EAL	Time	T1	0.41	0.49
		T2	0.44	0.50
		T3	0.53	0.50
L1	Time	T1	0.33	0.47
		T2	0.40	0.49
		T3	0.53	0.50

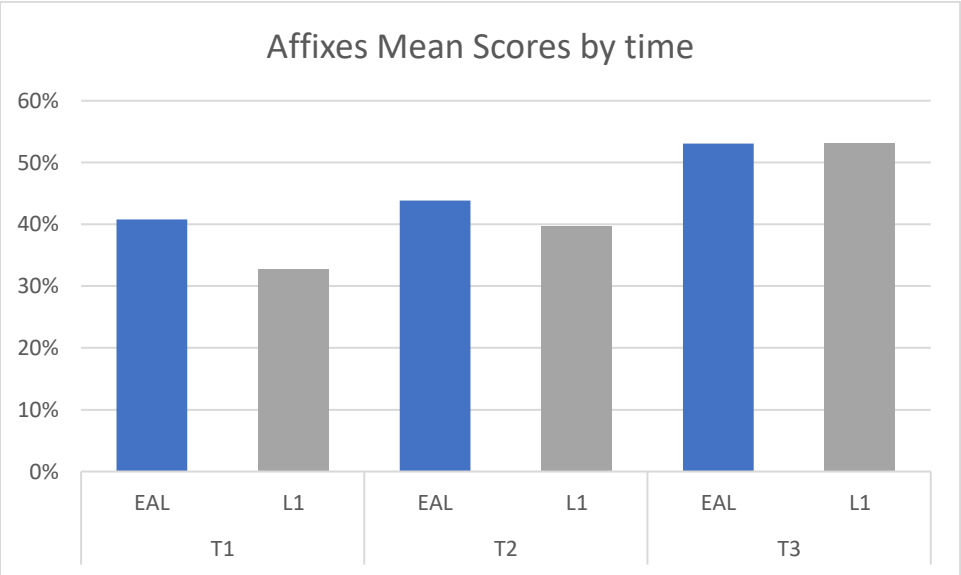

Figure 44: Affixes Mean scores by Time

Roots

Table 9: Roots Means and SD

				Root	(sd)
EAL	EAL	Time	T1	0.56	0.50
			T2	0.50	0.50
			T3	0.67	0.47
	L1	Time	T1	0.48	0.50
			T2	0.44	0.50
			T3	0.65	0.48

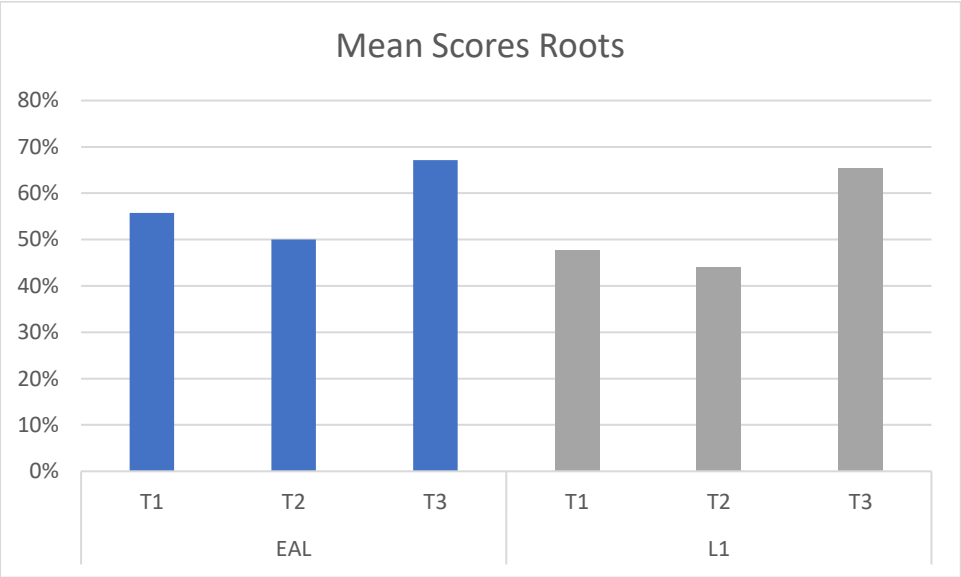

Figure 45: Roots Mean scores by time.

As can be seen from the mean scores on the two measures of semantic morphology knowledge, both the L1 and EAL groups scored more highly by T3. The EAL children scored more highly than the L1s at T1, but both showed an improvement between T1 and T3 for both measures. These observations were supported by the findings of the mixed models. Although there was an overall improvement across the three time slots, the significant effect on scores was seen in the data for the affix knowledge (WPLT) was observed in the large improvement between T1 and T3.

Affixes

glm(formula = WPLTA ~ Time * EAL + (1 | Participant) + (1 | WPLTitem), , family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.2324	-1.0234	-0.8933	1.2841	1.4910

Coefficients: (2 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.3735	0.1785	-2.093	0.0364 *
TimeT2	0.1261	0.2512	0.502	0.6156
TimeT3	0.4967	0.2505	1.983	0.0473 *
EALL1	-0.3391	0.2471	-1.372	0.1701
TimeT2:EALL1	0.1703	0.3451	0.493	0.6217
TimeT3:EALL1	0.3442	0.3431	1.003	0.3157

Roots

lm(formula = RootT1 ~ Time * EAL + (1 | Sub), family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.492	-1.1774	0.8926	1.0816	1.2806

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.22957	0.24062	0.954	0.34
TimeT2	-0.22957	0.33918	-0.677	0.498
TimeT3	0.48508	0.35022	1.385	0.166
EALL1	-0.32488	0.32500	-1	0.317
1 | SubTRUE	NA	NA	NA	NA
TimeT2:EALL1	0.08565	0.45943	0.186	0.852
TimeT3:EALL1	0.25027	0.47228	0.53	0.596

There was no difference in performance across the two groups, with both showing overall an improvement in semantic morphology knowledge.

In order to examine possible effects of individual differences across the participants in age, gender and their attendance records on their semantic morphology knowledge, two separate mixed regression models were undertaken, one for affix knowledge and one for root knowledge. The table below shows the mean scores on these ID variables.

Table 10: Age, Attendance, and means by participant.

		Age	Attendance	WPLTA	Roots
EAL	20126	92	1.0000	0.51	0.43
	20133	93	1.0000	0.69	0.81
	20137	90	0.9444	0.36	0.67
	20140	99	1.0000	0.49	0.71
	20141	97	1.0000	0.49	0.48
	20142	96	0.8333	0.54	0.57
	20143	87	1.0000	0.46	0.62
	20144	93	0.9444	0.44	0.71
	20145	98	0.8333	0.31	0.52
	20146	96	1.0000	0.31	0.24
L1	20121	89	0.9444	0.31	0.14
	20122	105	0.6666	0.31	0.52
	20123	92	1.0000	0.33	0.71
	20124	96	1.0000	0.23	0.29
	20125	93	1.0000	0.46	0.90
	20127	94	1.0000	0.28	0.48
	20129	93	1.0000	0.67	0.76
	20130	84	0.9444	0.46	0.57
	20131	92	0.9444	0.72	0.71
	20132	85	0.8333	0.13	0.24
	20134	92	1.0000	0.72	0.48
	20135	83	1.0000	0.41	0.48
EAL		94	0.9555	0.46	0.58

L1	92	0.9444	0.42	0.52

There was a significant effect of attendance on scores for the affix knowledge (WPLT), such that the better the attendance record, the higher the WPLT score, for both groups of children. There was also a marginally significant effect of gender, with girls outperforming boys in the WPLT.

glm(formula = WPLTA ~ EAL + AGE + Gender + School + Attendance +

(1 | Participant), family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.2812	-1.0811	-0.8516	1.2168	1.5551

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.47284	2.350175	-1.052	0.2927
EALL1	-0.333059	0.210821	-1.58	0.1141
AGE	0.002077	0.018411	0.113	0.9102
Gender Male	-0.325126	0.170681	-1.905	0.0568 .
School C	0.159014	0.24442	0.651	0.5153
Attendance	2.361483	0.967624	2.44	0.0147 *

In contrast, for the measure of roots knowledge, there were no significant effects.

glm(formula = RootT1 ~ EAL + AGE + Gender + Attendance + (1 |

Sub), family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4375	-1.2525	0.9551	1.0764	1.3289

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.01044	3.03004	-0.664	0.507
EALL1	-0.2494	0.21843	-1.142	0.254
AGE	0.01124	0.02339	0.481	0.631
Gender Male	-0.30600	0.21851	-1.4	0.161
Attendance	1.51318	1.26975	1.192	0.233

Is there a relationship between reading ability and knowledge of derivational morphemes and does this relationship differ between EAL and L1 English children?

Each of the participants undertook a number of measures of reading. To control for the potential effects of these scores on measures of morphological knowledge, two mixed regression models were undertaken for each of the morphological knowledge test scores (WPLT and roots). Table x shows the mean scores for these reading measures.

For the WPLT (affixes), there were main effects of SWRT and BPVSR.

lm(formula = WPLTA ~ Time * EAL + YARCcomp + BPVSR + SWRT + YARCRR + YARCacc + YARCcomp + (1 | Sub), family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.5293	-1.0398	-0.7501	1.1812	1.9055

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.4403	1.5382	-0.9360	0.3491
TimeT2	0.1297	0.2547	0.509	0.6107
TimeT3	0.510446	0.2539	2.0100	0.0444 *
EALL1	-0.1739	0.2570	-0.677	0.4985
YARCcomp	0.0035	0.0166	0.213	0.8314
BPVSR	0.0196	0.0094	2.0890	0.0367 *
SWRT	0.0772	0.0233	3.3180	0.0009 ***
YARCRR	-0.0284	0.0172	-1.6540	0.0981 .
YARCacc	-0.0100	0.0268	-0.3750	0.7080
1 | SubTRUE	NA	NA	NA	NA
TimeT2:EALL1	0.195371	0.3525	0.5540	0.5794
TimeT3:EALL1	0.3972	0.3509	1.1320	0.2576

For the roots scores, there was only a main effect of BPVSR

glm(formula = RootT1 ~ Time * EAL + YARCcomp + BPVSR + SWRT +YARCRR + YARCacc + YARCcomp + (1 | Sub), family = binomial(logit), data = roots)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.075	-1.13	0.624	1.036	1.771

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-5.3115	2.1201	-2.5050	0.0122 *
TimeT2	-0.2445	0.3500	-0.6980	0.4849
TimeT3	0.5131	0.3602	1.4240	0.1543
EALL1	-0.1724	0.3489	-0.4940	0.6212
YARCcomp	-0.041123	0.0233	-1.7670	0.0773 .
BPVSR	0.0666	0.0141	4.7310	2.23e-06 ***
SWRT	-0.002597	0.0317	-0.0820	0.9346
YARCRR	0.0091	0.0244	0.3740	0.7083
YARCacc	0.0244	0.0368	0.6640	0.5066
1 | SubTRUE	NA	NA	NA	NA
TimeT2:EALL1	0.0782	0.4832	0.1620	0.8715
TimeT3:EALL1	0.3294	0.4951	0.6650	0.5059

6.4 Post Hoc analysis of language use survey

During the testing at time 3 the participants that schools had identified as EAL additionally participated in a short 10 question survey about language use. This was brought forwards as initially it was to be part of the delayed testing and interviews with participants. As there was uncertainty as schools were closing due to COVID this was discussed with the author's supervisors and decided to bring it forward in case the delayed measures and interviews were unable to be conducted, as was the case in the end. The full survey is in Appendix 12.8. The questions were designed to quickly and easily be answered by young students. Questions covered topics such as what language is spoken with friends and siblings, what is the preferred language to speak in, what is the preferred language to read in. There were two questions that asked the participant to estimate what percentage of time they read in a language other than English and the same question but for speaking. The final question was if the participant had instruction of any type in their language other than English.

The ten EAL identified participants who completed testing at time 3 agreed to completing the survey and had all also returned consent forms for the additional delayed testing and interviews. The students self-reported the language they spoke other than English, of the 10 students, 4 speak Pakistani Urdu, 3 speak Polish, and one each for Italian, Tunisian, and Wolof. Please note that this is different than what was initially reported by the schools to the researcher. One participant when speaking with the researcher during the intervention stated, "My mum says not to tell people where we're from and what we speak, it isn't their business." This could mean potentially, that parents are giving the schools incorrect information, students do not know the official name of the language they speak at home, or that there was a clerical error with the school records are all possible explanations for the differences in information reported by the schools and what was reported by the students.

The first question presented some interesting information as two of the ten students stated they spoke English at both home and school There was an even split with 5 students stating they spoke the same language with their teacher as their parents. When asked about what language they spoke with their friends 7 of the students stated they spoke a mix of languages with friends and only 3 stating that they spoke their language other than English with their friends. There was a larger mix when asked about what language they spoke with siblings. Three spoke a mix of English and their home language, four

stated that they spoke only English and two stated that they only spoke their home language, and one participant did not have any siblings.

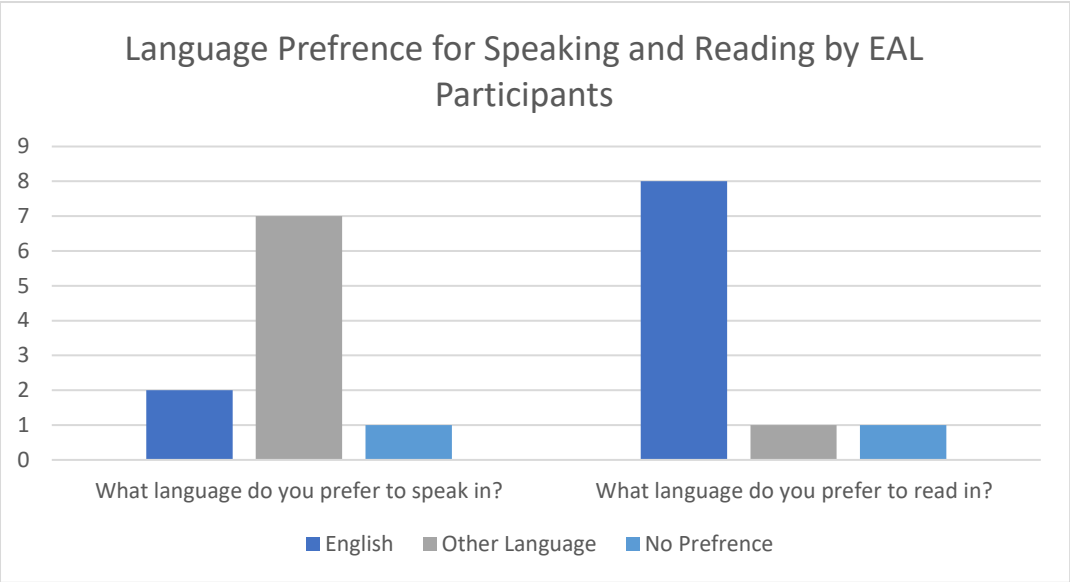

Figure 46: Language preference for EAL participants

This seems interesting as the next question asked about what language they preferred to speak in and only two preferred English while seven preferred their home language and only one participant stated they did not prefer either. These results were almost the reversed when participants were asked what language they preferred to read in, Eight of the ten participants stated that they would rather read in English than their other language. Only 1 said they preferred reading in their other language, and one said they had no preference. This could be the result of lack of instruction in their language other than English as three stated they attended classes in their other language, two stated they had instruction at home, and one said they used practice books to learn. Four of the participants received no instruction at all in their language other than English at all.

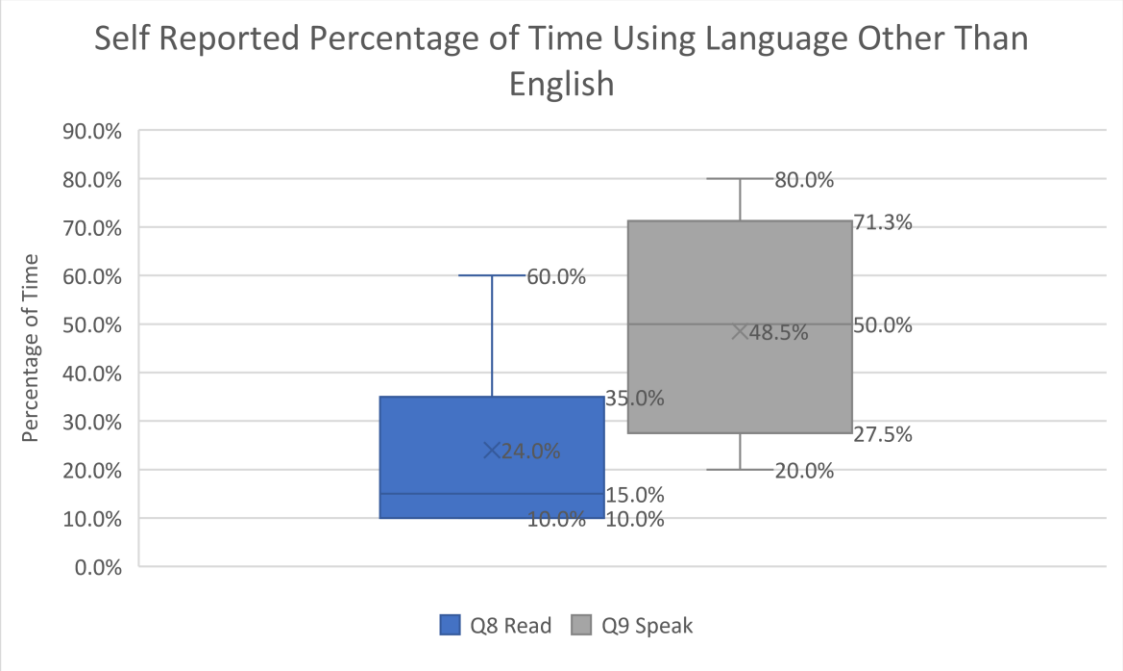

Figure 47:Self-Reported Percentage of Time Using Language Other Than English

The above graph illustrates that there is a significant disconnect for these students between how much they speak and how much they read in their other languages. With a mean of 24% of the time for reading that means that on average 76% of the time when they read, they read in English. This is interesting to consider as there are theories about BISC and CALP. Does the variance which ranges from 60% to 10% of the time have an impact on the participants reading comprehension in English which is being observed in this study? Could the participants instruction in their home language and use of it affect the results of the morphology intervention? Same with speaking in other languages, the mean is 48.5% and the range is 80-20%, 6 of the participants indicated an amount of 50% or greater timewise for speaking in their language other than English. This is a significant amount of time to be functioning in their other languages. This was not a tracked amount and more the participants impression of their usage. Which means there could be a discrepancy between what the participants believe that they are doing both speaking and reading and what is happening in actuality. There could be effects from the amount of language use and the quantity and quality of language instruction in the participants language other than English. The effects of participants working in multiple languages would need to be more closely studied and evaluated to decern if there was an effect and what the effects are. It is also notable that only one student stated that they read more in their home language than they speak. This is also a student who reported that they received no instruction in their home language and had no siblings. This could reduce the amount of

communication they perceive to be conducting by speaking or reading in their home language as they have fewer people at home to speak with. It is fairly logical to see that students speak more in their home language than they read as they would likely be speaking with parents and siblings in the home language.

6.5 Data Analysis: Teacher Interviews

The original research plan included both teacher and participant interviews to get a fuller picture of the impact of the intervention. The goals of the intervention beyond the quantifiably measurable improvements to reading comprehension and vocabulary were to develop an intervention that could be easily implemented and accessible to teachers. It was additionally predicted that there could be some additional effects from the intervention for participants. These would be increased confidence and development of reading skills from the small group work and potential successes during the intervention. Finally, interviews would give feedback as to the design of the program and parts that were beneficial and enjoyed and if there were parts that were confusing or difficult for participants. This information could be used to further refine the intervention for additional research.

While the original intention was to interview all of the teachers who had students participating in the intervention, this was not able to be completed due to both the pandemic and teacher availability. The teachers at one of the schools who had participated in both the pilot study and the main study were gracious enough to participate in follow up interviews approximately 3 months after the intervention ended. This additional time gave them the ability to observe and reflect on any growth that had occurred with participants both during the intervention and post intervention. There were several common themes that arose from the two interviews; they were student interest and engagement, confidence, and application post intervention in regular class activities. The intended outcomes were focused on improvement of participant reading comprehension and morphological knowledge. These unintended outcomes from the work are also important as they illustrate the potential for longer term effects in the absence of delayed testing data. The examples from the interviews below give support to the concept that the intervention has the potential to aid students with developing skills that will support their reading and understanding of vocabulary over a longer period.

6.3.1. Methodology

The interviews were originally planned to take place three to four months post intervention to give teachers time to observe any changes in their students who had participated. They were planned to take place at the schools most likely in the classroom during their planning time. However due to the lockdowns associated with the COVID 19 pandemic this was modified to be a video interview via google meet as both the researcher's institution and the participating school used google meet regularly. This had the additional benefit of the interviews being able to be recorded and transcribed automatically. Consent forms were signed digitally and returned with permission to record the interview and store the data for use. The interviews were arranged via email at a time convenient for both the teacher and researcher.

An interview schedule (See Appendix 14.5) had been developed, and included thirteen open ended questions about teaching practices, changes observed in the participants, and impressions of the intervention. The plan was for each interview to last approximately thirty minutes and be relaxed with discussion going more depth if teachers mentioned anything that could give more information about the effects of the intervention or ideas about how it could be improved.

The recordings were transcribed digitally using the tools included in google meet and then checked for accuracy. Teacher and participant names were anonymized in the transcripts and in the following discussion. All recordings and transcripts have been stored securely in the university provided google drive and on the university mainframe, protected by passwords that are changed regularly and with digital backup. The transcripts were then coded using thematic analysis. Codes that emerged were student interest and engagement, student confidence, application and transfer of skills, teacher practice, practices included in the intervention, and evaluation of the intervention. Each of these themes will be discussed in the following sections.

6.3.2. Student Interest and Engagement with the intervention

The first theme of student interest in and engagement came from multiple comments from both classroom teachers that participants were excited to come out of class for the sessions. Teacher L said in refence to the sessions that she observed in passing, "They were engaging. You always had... little pieces of paper with strips on and they were they were sorting words; they were finding words. So, the

activities are engaging. It wasn't just you talking about them and saying this is a word. What's the meaning of what you know, it was more of those activities that we do in class as well."

Teacher B also confirmed that students were interested and enjoyed the work. "…it was very nice to see them enjoying the intervention. And I think children like the quiet ones, like [student name] did as well." Participant interest and engagement are important parts of education in the primary years as they assist with building skills and ensuring that participants pay attention to the ideas and concepts they are working with. If they are not engaged, they are not learning and thus the time or session is not effective. The feedback from the classroom teachers that students were actively engaged when observed and brought that interest back into the classroom could be a signal that they were developing skills that they could then use later on and demonstrate competency during class.

6.3.3. Student Confidence

Both teacher participants mentioned repeatedly that their students who participated developed more confidence in the classroom. Teacher B mentioned a specific EAL participant and that their confidence improved noticeably, and they contributed to more class discussions about spelling and reading particularly when they pertained to prefixes and suffixes. They attributed it to the extra exposure and training they received from the intervention.

The willingness of participants to try in class when working with vocabulary was noted by Teacher L in that students were volunteering that they had learned parts of words in the intervention so they could guess that a new word might have something to do with a concept post intervention. The attempt to make educated guesses based on information takes confidence on the part of the student. The participant in this story also said that it was something they had learned from the intervention. This direct attribution by the participant to the intervention illustrates that the skills being taught and practiced in the intervention were absorbed and were being applied in new contexts with confidence. This could be viewed as evidence that the intervention's unquantified goals of confidence and application of skills post intervention were being achieved.

6.3.4. Application and transference of skills

Teacher L spoke about an incident several weeks post intervention where a participant came upon the word autograph in a story and took the time to break it apart. Neither morpheme was directly taught in the intervention there were mentions of graph when participants completed word maps, but neither were the focus of instruction. The participant segmented the word into auto and graph and then blended it so she could read it and understand the meaning of graph and give a correct guess as to what autograph meant. This is significant as the participant had some of the lowest vocabulary and reading scores but was able to take the skills from the intervention and apply it at a later time in class to understand and read aloud a word that she previously would not have been able to. This skill is an important part of the intervention and while this is anecdotal evidence it illustrates that there was some transfer of skill with even the weakest of the participants.

Both Teacher L and B mentioned that it was interesting to see what participants brought back into the classroom from the intervention. Teacher L noted that, "the higher ability ones that you had, they were using that vocabulary in their work, but in the correct context as well." Teacher B also mentioned that participants would come back and talk about what they had learned and then later those same words would appear in their regular writing assignments. This is indicative that participants were able to acquire knowledge about vocabulary from the intervention and then apply it in different settings and use it correctly. This is a significant point from the point of view of a teacher as it shows that they absorbed the vocabulary containing the morpheme and then were able to produce it in a correct context outside the intervention. This could also illustrate that the interest in morphemes was sparked. This transference of knowledge and consolidation of information is key in student learning.

6.3.5. Teacher Practices

An additional line of discussion in the interviews was what methods they currently used if they taught about morphemes. It was interesting to note that there were differing views and methods and frequency of teaching. This could be evidence of the differences and compliance with curriculum in relation to instruction in morphology. As one of the two interviewees had students participate in the pilot study, they might have absorbed some of the information and been more willing to utilize and teach morphology as part of the work with vocabulary.

Teacher B said her instructional practices mainly focused on phonics with prefixes and suffixes rarely being discussed in her teaching experience. She mainly followed a very structured program that in later parts addressed affixes as part of spelling but not as carrying meaning. While there has been work about using patterns in spelling to help students learn and it has been part of spelling texts for decades the lack of information about meaning being carried is disappointing.

The interviews did contain some discussion that indicated that while the teacher was aware and cognizant of the patterns and use of morphemes to carry meaning that there was not enough instructional time to really work with the concept with students. That the idea of exploring versus telling students about morphemes would be something they would be open to as it felt beneficial to the intervention participants.

6.3.6. Practices that were also included in the intervention.

Some of the practices that were mentioned in the discussion of current pedagogy and methods that were being utilized were also used in the intervention. Teacher L spoke about how she taught vocabulary and one of the habits she was building with her class was noticing vocabulary when it occurred in texts. Part of this practice was observed and noted informally when the researcher was in the classroom retrieving students for the intervention. It was observed that students would write sentences with the newly introduced vocabulary and then underline the vocabulary word used. This would connect with one of the steps in the intervention where students would draw a line under words containing the target morpheme and then circle the morpheme in the reading passage.

An additional practice that was used in both the intervention and classroom was work with sorting words. The Teacher B said that she used it as a small group activity for phonics work with students sorting words containing different phonemes. This could be evidence as to why participants picked up the directions for the activity quickly during the use of sorting words by morphemes in the intervention.

The final practice that paralleled the intervention, although not as closely as the others, was that one of the teachers described that when introducing new spelling words that contained the same affix, she would have the students note the common affix as a spelling pattern. While this is not the same exact practice as noting affixes and common meanings it is similar in the concept of getting participants or

students to engage closely with the vocabulary and notice similarities and how they could be employed.

That there were similar methods and skills that were already being used in the classroom is important as one of the goals of the intervention design was that it would be familiar to teachers and thus easier for them to adopt. The intervention being familiar and thus more accessible would potentially have additional benefits for teachers of them becoming more interested in morphology and then learning more on their own and bringing it to the classroom with interest and enthusiasm.

6.3.7. Teacher Evaluation of the intervention

Overall, the general feedback from teachers was that they were happy with the participation of their students in the intervention. They felt that it benefited their students by expanding their vocabulary and giving them additional skills with regards to vocabulary development. They also felt that the intervention aided the lower ability students by building their confidence in reading and encouraging them to approach "larger words that they would have avoided," previously.

They were pleased with the timing and content. There was positive feedback about what was observed as teachers were invited to observe sessions and were briefed about the session content before the interventions began. During the interviews both teachers stated that if they were to teach year 3 or higher again, they would be open to using materials like those found in the intervention in their instructional practices.

6.3.8. Summary

Overall, the interview data showed that there were positive effects from the intervention both in intended and unintended manners for the participants. There was good engagement in the intervention and the participants also were able to transfer skills and knowledge from the intervention to regular class work. Students also gained confidence with both vocabulary and reading that was seen in the classroom through increased participation and student work. Finally, there were effects for both the EAL and L1E participants in different areas, but all were positive.

If we look at the information provided by the teachers, it supports some of the quantitative data analysis that was presented previously.. The increases in participant confidence noted by the teachers could be a contributing factor for the increases in accuracy and reading rate. The usage of vocabulary gained through the intervention in classwork could be an early indication of later gains in vocabulary size. This could also support the gains that were shown on the two morphology measures as a deeper level of learning and connections occurred with the participants. This will all be discussed in more depth in the discussion chapter that will follow.

7 Discussion Chapter

7.1 Introduction

The original purpose of the study was to design and evaluate a potential intervention whose goal was to improve participants' reading comprehension by aiding their understanding of morphemes. To this aim the research questions were devised to evaluate the impacts of the intervention designed in this study on reading comprehension and the understanding of morphemes for participants. These questions also aimed to examine the two subgroups, EAL and L1E, for any similarities and differences in the results. This is based on previous work where it has been noted that there are differences in development of reading comprehension skills for EAL and L1E students (e.g., Deacon & Cain, 2011; Geva & Farnia, 2012). However, as there was a shift in the research questions to examine the effectiveness of teaching the semantic meanings of derivational morphemes and exploring the effect of instruction in derivational morphemes on measures of reading comprehension the discussion will likely cover parts of both while attempting to focus more closely on the revised examination.

To this aim the researcher designed an intervention that incorporated both reading and morphology instruction. It was assessed by developing and selecting measures to assess progress and knowledge at three points, giving a baseline period and intervention period with potential differences that could be compared. Follow up interviews with the participants teachers gave additional information about the effect of the intervention on participants' work in class. Additionally, data was collected from the EAL students about their use of language for both speaking and reading to see if there was any connection between perceived language use and the results from the work completed with the students.

7.2 Summary of findings

In this section we will review the findings of this study. The predicted findings for this research were that there would be increases in reading comprehension for all participants with potential for more significant increases for the L1E participants as they might have already begun to develop skills to recognize morphemes and utilize meanings to help when they encountered unfamiliar words. This

would be due to earlier research findings that have shown that the reading comprehension of EALs tends to be less well developed than their monolingual counterparts (Deacon & Cain, 2011). However, there is potential that reading comprehension might increase for both groups and not be primarily dependent upon language background. This has been found in, for instance, O'Connor et al. (2019), in their large sample of 178 EAL and 94 L1E fifth grade students in Canada when tracked over the primary years. The authors report that the differences were not due to being EAL or L1E but due to being good or poor comprehenders to begin with (O'Connor et al., 2019). In many studies, including ones by Lesaux(Lesaux, Crosson, et al., 2010), with Spanish L1 students learning English there were corelations between the Oral Language skills and reading comprehension, this is thought to be due to the significant work with phonics that is part of primary curriculum. This could be similar to the findings with the Single Word Reading Test (SWRT) that indicated gains.

For the whole group, the measures of reading comprehension illustrated that over the intervention period there was statistically significant growth when compared to the change over the baseline period. The measures of morphology additionally found statistically significant positive change in morphological knowledge in comparison to the baseline period for the intervention period. This indicates that there were positive effects from the intervention that went beyond the changes that occurred in regular classroom instruction. The statistically significant change in accuracy was present after the intervention period but not in comparison to the baseline. This could have been caused by effects other than the intervention or potentially due to skills practiced in the intervention that were not the focus of the intervention itself. The changes on the other measures did not meet the threshold of statistical significance, which would indicate that neither the intervention nor the regular classroom instruction affected them to any significant degree over the period of this research.

The finding that for reading comprehension there was a statistically significant finding with a large effect size for all participants, when analysed both by one-way repeated measure ANOVAs for the participant scores and dependent t-tests for the gains, is very important as it illustrates that there was a significant impact on the participants' reading comprehension of the intervention. This is not a surprise, because the participants as part of the intervention had additional practice of reading closely and thinking about what they had read carefully. This, along with the work that the intervention indued them to do with morphemes, could have potentially increased the ability of the participants to understand the passages they read as part of the measure of reading comprehension. As illustrated in

the Morphological Pathways Framework, morphological decoding skills has a significant impact on reading comprehension ability (Levesque et al., 2020). This would lead to what is termed 'the consolidated phase of reading', where letter patterns and larger units of words become more embedded in memory and easier to access, as proposed in earlier research (e.g., Ehri & McCormick, 1998).

If we first look at the growth for the receptive vocabulary, and reading rate where growth was stable and had similar effect sizes, this could potentially be evidence that supports the idea that participant development overall was rather linear and progressed steadily over the course of the intervention study. That accuracy did not show similar, linear and steady development suggests that there are potentially other forces at play. For instance, according to the notes of the test administrator, although the students were not given feedback on any of the measures, but they were able to see the timer for the reading passage. This one potential piece of visible feedback and it could have driven the students to want to beat their previous score and to read faster thus causing them to make more errors and thus reducing their accuracy scores on the second administration. This could have also contributed to the rise in the reading rate. These issues would need to be considered and mitigated during any future research of this type.

If we examine the whole group data first since it is the only group that is fully powered, there were improvements over the entire time of the intervention for all the measures. The use of the baseline growth and the intervention growth could potentially be a new method for in-school research with small groups. However, there would need to be significantly more work undertaken to illustrate that development over time is linear during the course of a school year, and not with some parts of the year showing stronger growth than others. However, for this study it is assumed that the growth is linear and stable over the course of the year.

There are also questions about practice effects that would need to be addressed as the same exact measures were given at all three points. They were separated by ten-week gaps. This should have made it less likely for students to have large practice effects, but it is unlikely that there were no practice effects at all. While changing the passages on the YARC was considered in the design of the study, because there are A and B versions of the passages, there were concerns that the change in passage would be an additional confounding variable and there were additional concerns about the use of different measures at different times. Thus, the decision was made to use the same form for all

three tests. The three measures that are the main focus of this work: Reading comprehension, Morphology measure: prefixes and suffixes and morphology measure: roots. None of these showed statistically significant changes over the baseline period but they did when looking at the whole group over the intervention period. If there were significant practice effects, they most likely would have been present over both periods. This supports the idea that the practice effects, while arguably present, were not significant enough to necessitate the use of additional forms of the measures.

The general trend of the findings that there were more significant effects over the intervention period than the baseline period is not surprising as there would have been no change to the instruction the students had received over that period. The consistency of teachers and the environment is one of the strengths of using the same group of students as their own control as the changes would be the same for the whole group. This suggests that the idea that the changes overserved were primarily the result of the intervention has some merit.

The implications of these effects are that the intervention was successful in improving the participants' reading comprehension and in instructing them in the meaning of morphemes. The question of whether there were additional effects or improvement over time is not known because of the inability to carry out delayed testing and focus groups with the participants as planned. What can be taken from the work that was carried out is cautious optimism that working with students even for a short period of time on a regular basis can influence and improve their knowledge of morphemes which is likely to impact their reading comprehension ability, based on both theories of reading and theories of reading development. This would potentially be due to a speeding up of progression from one phase to the next of reading development by aiding the development of necessary skills to identify morphemes in morphologically complex words. This in turn would activate knowledge in nodes or develop additional nodes of word knowledge in the mental lexicon and thus move a reader from the consolidated alphabetic phase to the automatic phase of reading, as described in Ehri & McCormick (1998) or more explicitly, from the Orthographic phase to the Morphographic literacy stage in Seymour and Duncan's (2001) work. Seymour and Duncan's morphographic literacy in could be seen to be particularly relevant here, as it indicates that if a learner is aware of the morpheme as a unit, they are then able to connect it to others to form words. This awareness of morphology then reduces the effort needed to read more complex words. This move to morphographic literacy would likely also assist with the move from learning to read to reading learn, which is a key transition in the

development of reading. It is also an important development from teachers' perspectives since they are under significant pressure to make sure that students are able to comprehend increasingly complex text and start to undertake their own learning. This can also be related to the shift from a focus on dependence on single word reading lessening over time with comprehension skills becoming automatic, as has been documented in the literature. For instance, Catts, Adlof and Weismer (2006), worked retrospectively with test scores from students in grade 8, 4 and 2 and divided them based on being 57 poor comprehenders, 27 poor decoders, and 98 typical readers. Their study found that the relationships shifted over time and that these changes in priorities could have an impact on students reading comprehension overall. The increases could also reflect the DIER model(Y.-S. G. Kim, 2020) as the increased morpheme knowledge could indirectly affect both the reading comprehension and vocabulary knowledge.

For the EAL participants the two most significant effects from the intervention were the significant gains on reading comprehension and the morphology roots measure. While there were smaller and not statistically significant changes on the other measures. There were larger gains on average over the intervention period for all measures although they did not meet the threshold for statistical significance.

There was a statistically significant positive change on the measure of morphology for roots but not prefixes and suffixes. This could be for a couple of different reasons. The participants might have started with more exposure to prefixes and suffixes and thus they may already have had that knowledge encoded and therefore it would not be considered new knowledge. This could mean that it would take more exposure and or learning effort would have been needed to increase that score with new knowledge. Alternatively, the participants may have already begun to develop the knowledge of roots, and it was consolidated through the intervention and thus it was easier for them to demonstrate a gain in knowledge in the test measures. If so, this can be related to the studies where reaction times are reported as being quicker when morphologically related words were shown one after another and the response times were reduced as the cognitive load had been decreased, (e.g., D. Crepaldi, Rastle, & Davis, 2010; Davide Crepaldi et al., 2013; Jared & Seidenberg, 1990), although this work has mainly been carried out with university aged participants. Nevertheless, it may still be relevant, as the priming effect is likely to be present even in much younger students. In this way, it may have been

the case that the knowledge that was already present was activated and added to, therefore when it was tested, it was more easily retrieved.

When examining the L1E participants' results, there were more statistically significant changes compared to the results of the EAL group. The L1E participants showed statistically significant gains on several measures, such as accuracy, comprehension and, single word reading. That there were no statistically significant gains on the morphology measures is interesting as it is a different result than what occurred with the EAL participants. The L1E participants did show statistically significant gains for reading comprehension, however, which would mean that there likely was some impact from the intervention. These results correspond loosely with work done with younger (age=6 years) students who were given additional practice and tutoring in a study by Vadasy, et al. (2005), where students received approximately 40 additional hours of instruction in phonics and in which one of the groups had additional practice with reading passages aloud. One of the components of the intervention in this study was a reading passage that was read orally by all participants. The participants in Vadasy's study, who had similar additional guided oral reading practice, also demonstrated increases in reading fluency and accuracy post intervention in comparison to a control group and the other intervention group. This was a study with 19 students in each of three groups, involving reading practice, word study, and control. This could indicate that there was some effect purely of the act of receiving instruction in a smaller group or in addition to the regular classroom instruction. This would correspond with the practices of having students read aloud and discuss in class. Smaller groups in a classroom reduces the ability of students to avoid participation and thus requires them to actively participate in the activity and to employ the skills that are being worked on. This variable was not explored in this study, but it could be a possible effect from the work in small groups of between 4-6 students during the intervention, in contrast to their regular classrooms, which comprised between 25-30 students. This is significant as with such a small group each student had opportunities each session to have success with reading and receive positive reinforcement for this work.

It is unique and interesting to note that where both groups started at about the same level, the increase in scores seems to be larger for the participants that would have originally been in the lower 50% of the participants to start with, particularly for the EALs. This certainly would need more exploration to see if it could be replicated with a larger sample and a control group.

7.3 What is indicated by the reading comprehension results?

The results of this study indicate that there was an improvement for reading comprehension on the measures used in this study. They clearly demonstrate that there was a significant and large effect size for all participants over the period of the intervention versus the baseline period, as demonstrated by a dependent (paired samples) t-test (t (21) =-4.349, p=.000, r=0.69). The lack of significance or effect size over the baseline period indicates that repeated use of the measure did not have a significant effect. The results of repeated contrasts revealed that the average score at T2 was not significantly higher than at T1 (F(1,21)=0.953, p=.340, r=.20), and the average score at T3 was significantly higher than at T2 with a large effect size (F(1,21)=70.255, p=.000, r=.88). A cautious conclusion would be that this shows that the intervention improved reading comprehension for the participants as a whole group. This would be supported by previous intervention studies completed with students who were slightly older, fifth grade and sixth grade, that worked with morphemes (e.g., Goodwin, Huggins, et al., 2013). The Goodwin et al. study had a much shorter intervention than the one conducted for this thesis as it was only four sessions long versus the eighteen sessions that were conducted over the intervention for this work. While Goodwin et al. (2013) found negligible effects for reading comprehension with a much larger sample (203 ELLs), it is interesting that the current, much smaller study had a larger effect size. It has been found that effects sizes tend to diminish over time with more research being conducted (c.f., de Bruin & Della Sala, 2015). This does not mean that it is not useful to use effect sizes when comparing data, merely that caution should be taken, and researchers should not over-rely on them. If there are these diminishing returns, then with a larger sample there would be a smaller effect, than on the smaller but, yet fully powered sample present in this work when looking at the whole group data. This would support the idea that work with students in primary grades or the elementary years of education in morphology could potentially result in benefits in reading comprehension for students. This could only be seen as a good thing as understanding what is being read is a key for students since large parts of many curricula are delivered through text, particularly as students' progress to the upper levels of education.

A potential reason for this effect could be the increase in understanding of morphologically complex words and the development of the skills required to break down morphologically complex words and understand at least individual morphemes. This could result in the ability of students to use this knowledge, together with whichever clues they are able extract from the context of the text, and to

combine these in order to make educated guesses as to what meaning the text is attempting to convey. This use of multiple intertwining skills is potentially possible within many of the frameworks discussed in the literature review such as, Levesque et al., (2020) and Perfetti (2007), both of which propose that multiple skills and facets of reading work together to support comprehension. Thus, the idea that the linguistic and orthographic systems are working together with lexical representations and general knowledge all to support comprehension. Therefore, it is arguably possible that readers employ knowledge contained in all these areas working together to interpret even incomplete information in light of what is known by the student. The concept of skills working together and being dependent upon each other would be logical considering how the phases of reading development also build upon skills previously developed. It could indicate the progression and overlap between phases as it is postulated by Ehri and McCormick (Ehri & McCormick, 1998) that students may not just fit neatly into one phase or another. Rather, the authors claim reading development takes along a continuum, with periods of progression overlapping. This continuum of skills development and progression would assist with making more complex and difficult texts more accessible. It would also suggest that skills that are yet to be fully developed can be combined with incomplete knowledge and may allow learners to make "educated guesses" when they are attempting to perform at levels that are slightly above their current level of literacy or development. This may be particularly relevant for EAL or lower SES students who have been reported to be behind students of equivalent age who are L1E or of higher SES with regards to vocabulary (e.g., Buac et al., 2014; Gilkerson et al., 2017; Hart & Risley, 2003). This idea of intertwining skills and improvement in one supporting other would align with theoretical work. The multiple models of processing indicate that the exact balance of skills utilized might be more fluid and this would align with the idea that there is a shift in dependence at certain times on one specific skills in comparison to others as they develop, and the original basic skills become less important as they become more automatic and as readers develop.

The statistically significant and large effects for single word reading and accuracy found in the current study are interesting and need to be considered both together with the reading comprehension's statistically significant and large effect, as well as individually. This is discussed below.

Considering these three measures together illustrates how separate skills can influence others. Being able to read a single word at a time with no additional context or interference from other factors simply shows that a student has reached the phase of consolidation in reading development, as to get

the word correct it must be read completely and correctly on the first try, as tested on the SWRT measure. Thus, the skills involved in sounding out and breaking complex words into smaller chunks would not be applicable. The increase in difficulty as students' progress through the list of words contained in the single word reading test is obvious as they start with very simple words (e.g., see, play, this) and then progress to comparatively more difficult words by the end (e.g., haemorrhage, abysmal, pseudonym). However, when students indicate that they do not think they can continue they are allowed to stop. Many would attempt one more block of words past where they were showing visible signs of struggle as they wanted to compete with their own memory of where they had stopped the previous administration. (This observation is based on comments that the participants made to the researcher as they were taking the test. The students would ask if they had done better and then talk to themselves about where they thought they had gotten to on the previous attempt). They might get a couple more out of the set of ten or row of five. While this would boost their score, it was also a notable struggle. The desire to learn and know what the words were was obvious as many would ask the administrator what a word was or what did it mean. This internal push to do better also demonstrated some confidence by the participants. Each set of ten words indicated a level of passage that they should be able to read for the reading comprehension. At T3 the average score for participants was 42.0 which would mean that on average, participants most likely could read the passage at a year 4 level. Which is not surprising but still slightly better than one might expect to find with students with three months still to go to the end of year 3. This would indicate that while there was a statistically significant improvement it could potentially be due to natural maturation in the participants. This does however raise questions as to whether increases in single word reading really equate to an increase in reading comprehension. If there had been an additional aspect of meaning combined with this, it could potentially contribute more strongly as an aspect of comprehension. As the measure stands it simply indicates that a student is able to decode and state the word orally. With the increasing emphasis on phonics over the recent decades, a student could potentially be able to decode a word and speak it correctly with no knowledge of what meaning it carries. This issue was noted in the first administration at T1 as several students scored well on the single word reading test and it was indicated that they should start with a higher-level passage than they were able to read. Following the administration directions, they were moved to a lower-level passage that they were able to read and comprehend. Should this work be replicated, this issue should be explored or addressed with a modification of the measure or possible substitution with a different measure.

The above leads to a discussion about the applicability and previous performance differences on researcher made measures versus off the shelf standardized measures. In an examination of researcher made tests and standardized commercially made measures, there seems to be better performance observed on researcher made measures (Clemens & Fuchs, 2021). This is likely because the skills taught are the same as the skills that are measured. In contrast, if a commercially available measure is used, it may question in a different format to the one taught, and additionally it may require the use of slightly different skills. While traditionally this has been seen as a weakness of interventions, it could merely be a more rigorous measure that does not measure exactly what was instructed. It has been suggested that a combination of commercially available and standardized measures with a more specific researcher made measure would give a better view as to the efficacy of interventions. The use in this study of a mix of commercially available and researcher made measures should be seen as using a balanced approach to the work. It also increases the rigor for the participants as they are being measured with open ended questions and not multiple choices when it comes to reading. While this might be a partial explanation of why some measures did not have the predicted gains or statistically significant changes, it additionally illustrates that there was an attempt to ensure the measures were both applicable and balanced.

The improvement in accuracy is notable as the scores at T2 were lower than T1 and the scores at T3 were significantly higher. A one way repeated measure ANOVA confirmed this change was statistically significant: $F(2,42)= 18.507$, $p=.000$ for the change between T1 and T3. This also was reflected in the repeated contrast results of ($F(1,21)=39.904$, $p=.000$, $r=.81$) for T2 to T3. This increase is notable and could potentially be caused by multiple factors beyond just the intervention. It is interesting to note that the gains, just the average change in scores, was almost 5 points. The baseline was -3.32(SD=6.357), and intervention was 8.27 (SD=6.143). This would indicate that there were significantly fewer errors in reading on the final administration of the measure. If this was purely due to practice effects, there would likely be similar results for the reading rate scores as well and there are not. The reading rate continually improves whereas accuracy scores did not have the same consistency. What this could indicate is that the skills the participants acquired via the intervention affected how they read whole passages, even when they were not explicitly decoding morphologically complex words, as they took slightly more care to read and did not skip or incorrectly decode words. This could lead to discussions about what constitutes improvement in reading: the speed, accuracy or understanding of what is read. While all three of these components are important, obviously for this study, the understanding or comprehension is the primary focus. The improvement in accuracy is a good addition to the participants' skills but was not predicted or expected by the researcher.

In one meta-analysis of morphology instruction (Reed, 2008), there was a finding that supports these increases in accuracy as there were medium sized effects post intervention. This finding from the analysis of eleven studies with populations ranging from reception to 12th grade, indicates that the increases found in the study reported in this thesis are not abnormal and may even be expected. This demonstrates that there is some interaction between the findings from research with morphology and accuracy in reading that would need further work to investigate the exact nature of this interaction.

One thing that was unexpected was that there were negative gains for some participants over the baseline period. This is shown in Figure 41., which includes all participants' gains for reading comprehension. It can be seen that almost half the participants had a negative gain over the baseline period on reading comprehension. This is interesting as none had a negative gain over the intervention period. The causes behind this are unclear as the testing occurred in the same environment with the same forms and the same person administering the measure each time. All the measures had some portion of the participants' experiencing a negative change over the baseline period. The lack of feedback from the original T1 measures and then the second administration could be a unique variable that has some effect. This is because students in schools in England are given significant amounts of feedback on all work, and the lack of feedback from these measures could have indicated to them that they had answered correctly, and they needed to choose a different answer than they possibly remembered having given previously. However, it is unlikely that the participants remembered both the question and answer given from ten weeks previously. What is more plausible is that students had a vague memory and they potentially second guessed themselves as they were not confident as to what the correct response was. Over the intervention period only four measures showed these negative gains: receptive vocabulary, reading rate, morphology measure of prefixes and suffixes and morphology measure of roots. It is interesting to note that single word reading, accuracy, and reading comprehension only had positive gains over the intervention period. This could be due to natural variation in participant ability, the lack of feedback or even knowledge of scores, or even the relationship with the test administrator. The fact that the participants demonstrated positive changes over the intervention is important as it illustrates that there was some type of positive impact on their single word reading, reading accuracy and reading comprehension over the period of the intervention. There has been some work with the early lexical analysis leading to later gains on reading comprehension in primary education (Carlisle & Fleming, 2003). This could indicate that work examining vocabulary as it is encountered could result in later gains or in the case of this study's

findings, gains in reading comprehension were perhaps partially due to the additional close reading practice.

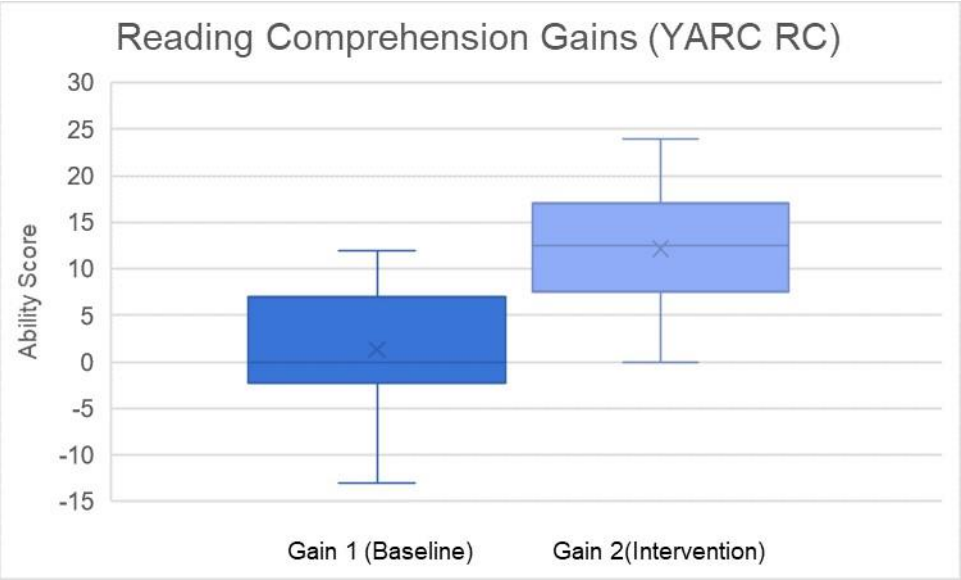

Figure 48:Whole Group Reading Comprehension Gains

The negative gains on the reading rate could be due to the development during the intervention of a relationship with the test administrator, as they also conducted the intervention sessions. This relationship and the students' natural inclination towards wanting praise could together have resulted in a greater desire to perform better on the final measures. The participants may have wanted to do well for the person who had been working with them and who had been encouraging them over the previous 10 weeks. This would mean that they might have slowed their reading speed in order to read more accurately and make fewer mistakes. This could have potentially been avoided if it had been a larger study and multiple researchers or research assistants had been involved and thus someone who was not known to the participants could have administered the measures and thus potentially removed some of the effects of the relationship that had developed between the researcher and participants. However, as this was a small study with no outside funding it was not feasible and additional researchers would have raised more issues with trying to get access to the research schools as they were already concerned with additional people being allowed access to the school and students for safeguarding purposes. This was mitigated by the use of background checks and the researchers' previous work in primary schools, finding additional researchers who met the criteria set forth by the schools would have been difficult and complicated matters. So, while it would have potentially been better to have had additional researchers involved, it was not a practical or reasonable alternative.

7.4 Discussion of the Results from the Morphology Measures

An expected result was the improvement that was observed for taught roots and affixes. This improvement was statistically significant for the measure of roots. This demonstrates that there are gains that can be made on student knowledge of roots over a short period. The lack of transfer from taught roots to untaught roots is not surprising considering that it has not been found in other studies where they have measured over the short term. The goal of this study to additionally investigate long term effects could have added to this had those measures been able to be administered. That the results for roots were stronger than for prefix or suffixes is somewhat surprising, however. This is because a more productive task where participants were given the meaning and had to match the root whereas the prefix and suffix measure was designed to be a more receptive measure and should have been slightly easier as they were provided with the affix and then just had to select the meaning. It is generally acknowledged that receptive vocabulary knowledge develops before productive (Zhou, 2010, p. 15) so the results on these two measures being reversed is slightly unusual. One potential reason is that the intervention possibly connected knowledge that was already beginning to develop within the participants from their own reading or the work they were already completing in school. Helping them to connect the meaning to an already established collection of words that they were familiar with it perhaps reinforced knowledge and supported development that was already in progress. For example, if a student had previously been exposed to the word *astronaut* containing the Latin root astr- and they also had been exposed to astronomy or astronomer in science lessons. They had the seed of a node of knowledge that words with astr- in them had something to do with space or stars. The intervention reinforced this knowledge and consolidated the meaning as *star*. They then encounter this on the measure with the question being asked “Which word below do you think has something to do with stars?” and the knowledge is activated, and they can select astronomy from the three options as it has the same root astr, that they have connected with the meaning stars. This recall could have additionally contributed to the building and reinforcing of the pathway for the information that astr- means something to do with stars. These multiple exposures and usage and then testing could only help to support the learning and then potentially the skill of looking for morphemes and using known meanings to understand what is being read.

As regards earlier research undertaken with the morphology measure of prefixes and suffixes, the original study (WPLT) did not involve an intervention, so it is not directly comparable to the work

done in the current study. However, others have involved interventions and have worked with primary aged students and in these, the reported effects are similar to what is found in this study. For example, Kieffer and Lesaux (2012b) observed, with a much longer intervention of 18 weeks, that there was four months' worth of additional growth in the development of morphological decomposition skills for the 476 EAL and ELL participants who were in grade six in the United States. It has been consistently found that there are increases in knowledge and ability for students post intervention versus the control. In this study that would be comparable to the increases post intervention versus the change or lack of change over the baseline period. The fact that there was no significant change over the baseline for both the whole measure and when broken down into taught and untaught measures illustrates that there was no impact on this knowledge from the regular curriculum instruction. That there were statistically significant changes for both the whole measure and taught affixes demonstrates that there was an impact from the intervention in a positive manner on participants knowledge.

The effects found in this study and in others that have been similarly designed but with slight differences in population are evident. The effect size for the rehit measure with only the second-grade students in Apel and Diem's (2014) intervention study showed that they had a large effect size of d=.86. If we examine the other morphological measures they utilized of relatives (d=1.07) and affix identification (d=1.52), it can be seen that the effect sizes for this present study with the whole sample were medium to large (of r=.88 for the prefix and suffix morphology measure and r= .55 on the roots morphology measure). Taken together, this is growing evidence that instruction in morphology has an impact on student knowledge of morphology even at young ages. The differences in intervention type and the larger focus on inflectional derivation used in the Apel and Diem study should be noted, as instruction in inflectional morphology as a manner of changing word class begins very early in many curricula. One interesting thing to note is the single word reading effect size for this study is much higher at r=.84 when compared with an equivalent measure utilized in the Apel and Diem study of TOWRE: Sight word efficiency, that had an effect size of d=.12. This would be unexpected even with the slight differences in population age, since there is only a difference of one year of schooling. That year of schooling could be where the shifts occur in the phase of reading development that directly impact word reading ability.

One hypothesis that has been mentioned repeatedly is that students with lower morphological knowledge at pre-test stages benefit more over intervention periods than students who start with larger

morphological knowledge. However, when this was modelled by Kieffer and Lesaux (2012b) it was found to not be significant. Instead, they implied that potentially there is a difference when using nonword and real words based on the word experience model from Reichle and Perfetti (2003). This model is similar to the Lexical Quality Hypothesis that states that high quality lexical representations are needed to support reading comprehension. These representations consist of orthography, phonology and semantics. While this study did not employ statistical modelling as it did not have a sufficient sample size, the originally hypothesis could still have potential as the EAL participants did have statistically significant gains on the prefix and suffix measure where the L1E participants did not. This area certainly would warrant more investigation, in order to test these hypotheses more thoroughly. While this idea is counterintuitive as there is a significant amount of theory and work that demonstrates that those who have larger vocabularies increase at a greater rate naturally (c.f., The Matthew Effect), the interventions affecting students with lower starting points more significantly does not seem to follow this pattern. However, if it is investigated further and indicates that an intervention can assist students who are lower ability or have less knowledge to begin with, it would be an important direction for teachers to explore. There is some pressure for teachers to improve test scores for lower ability or lower performing students in many educational systems. The ability to target those students and aide them it would be incredibly important for both teachers and students.

One potential area for continued work would be the development of standardized measures for derivational morphology that are applicable to primary aged students. While the measures used in this study are a good starting point there is room for continued improvement and standardization. This was beyond the scope of this study, but it would be a potential direction for further investigation. The multiple-choice format for the morphology measures, while not as rigorous they could have been, are nevertheless an attempt to measure a construct that has yet to have a measure developed and standardized for primary aged students. Is it the best format for measuring knowledge of meanings of morphemes? Possibly or possibly not but, it is a starting point that can be refined and developed over time. The methods that have been used previously have focused more on the extent to which a participant can either morphologically decompose a word or whether they can use morphology to complete a sentence. For example, one commonly used measure was developed by Carlisle (2000), where participants were given a word and a sentence and were asked to use the word to correctly complete the sentence, for example: “Father refused to give _____ (permit)”. The second commonly used version is an analogy measure where students are given a pair of words and then a third word and

they are expected to complete the analogy, for example: “cried: cry:: drew: ____” (Nunes & Bryant, 2006, p. 14). These two methods would not be valid when trying to assess if a participant can remember the meaning of a morpheme without context. While the measures could have been more balanced, they were originally intended to be balanced between taught and untaught morphemes. If more questions had been piloted and designed, it would have been a better measure. This is a lesson that will be taken and applied to future work to always design and pilot more than will be used as this could have made alternative methods of analysis possible. The balance was thrown off even more due to the pandemic shortening the intervention time by two sessions and morphemes that were tested not being taught.

One of the purposes behind selecting common roots was to be able to tap into this knowledge and ensure that knowledge gained through the intervention would be useful for students over their academic careers. While much research is conducted for theoretical reasons there, the current study aimed also to be of practical benefit as well for participants on a student, teacher and school level. This may not be a commonly held belief amongst academics in Higher Education and research. The majority of the studies read while conducting the literature review for this thesis were focused on developing theories and then testing them. The next logical step for this researcher would be to then test and develop ways to apply these theories to actual classroom settings. Where a practical benefit would be aiding students who are attempting to learn. This leads to the focus of this work being a practical application of theory to an organic school experience. Testing the application of theory relating to learning and teaching to a valid educational context where it could demonstrate real benefit for either students or teachers should be expected and not an exception to the normal practice of research. The current trend of developing a developmental theory and then stating that teachers should teach a concept but with no guidance of how does not feel like a full application of the work. It leaves the practitioner in the field to interpret the work and potentially not apply it as it was intended or posited to work as a theory. This leads to the idea of taking the next step that should be achieved before theories are accepted and researchers move on to the next theory. The intention of this work was to see if there could be an impact on morphological knowledge for students, and then a further impact on reading comprehension due to this morphological knowledge. This was potentially indicated by the positive results on both reading comprehension measures and the morphology measure. This now leads to the additional benefits that were indicated by both the quantitative measures and the qualitative interviews.

The practical impact on the student level could be tied to the increased confidence of weaker students who were observed by their classroom teachers to be willing to work out words that they had previously struggled with. Additionally, the increase confidence could have affected motivation influencing students to read more, gaining more incidental exposure to vocabulary and thus over the long term potentially becoming a stronger reader and less reluctant to participate in reading, spelling or vocabulary activities in class. This increased participation and reading would connect to the work done on the Matthew Effect where those that can easily read more, in fact do so, and thus they are exposed to more vocabulary which this then results in them having larger vocabularies and being able to make use of them when called upon. This is based on the work done with this the theoretical construct of the Matthew effect and the work that has been completed currently that supports this concept (Pfost et al., 2014; Stahl, 2006). Extrapolating this theory to work with morphemes, while it is tentative, it is possible, as there are significant percentages of words in the English language that are morphologically related or word families. If a student gains access to a number of these and can thus build up the relevant networks of nodes in their mental lexicons, presumably they would then have at least partial access to other members of the family that are related. This could exponentially expand a vocabulary and increase the ability to comprehend unknown but morphologically related words when they are encountered. This ability is potentially useful for all students as they are likely to encounter a variety of text types and terms as they progress through academia but being able to connect prior knowledge of words or morphemes to deduce the meaning of a newly encountered word would reduce the level of difficulty when this occurs.

The utility of students' developing the skills and knowledge to be able to morphologically decompose an unknown word when encountered was illustrated by Teacher L in the interview with the student who decomposed autograph and knew it had something to do with writing due to being exposed to that morpheme as part of the word maps completed in the intervention. That this, this EAL student was able to access the knowledge and apply the skill enabled her to understand what was occurring as she read a story even if she did not have a complete understanding of the word. This partial understanding certainly aided comprehension and gave her the confidence to try and understand. What was not explored in this study, as the focus groups with the participants were cancelled, was the changes in student attitude or behaviour that could have come from this intervention. This would have been an interesting area to explore as student motivation can be a key to learning particularly with EAL students. Reading motivation has been described as being made up of "interest, trust and

commitment, interest is related with the initial orientation of the mind, trust is related with the individual's belief in himself/herself, and commitment is related with desire", according to Bozgün & Akın-Kösterelioğlu, (2021, p. 4). This trust component is where the potential additional growth in the participants came about as they were in a small group of peers who were working on similar issues, and they were able to feel successful by being able to complete tasks and discuss ideas to come to conclusions together. This trust then likely fed into a further desire to have more success and thus the observations by the teachers of students being more confident and willing to try and answer more questions in their regular classwork. This would be supported by findings that indicate reading and writing motivation increase as students' academic grit and social-emotional skills increase (Bozgün & Akın-Kösterelioğlu, 2021; Elias & Haynes, 2008). This building of small consistent successes resulting in increases in motivation and confidence would be an interesting additional path of exploration should this intervention be replicated as student motivation is a complex issue and being able to support that in non-traditional manners would be an additional benefit for practitioners.

While much work has been done into how reading comprehension and morphology are theoretically related and could potentially support each other (e.g., Peter N. Bowers, Kirby, & Deacon, 2010; Gersten et al., 2020; Levesque et al., 2018; W. E. Nagy et al., 2014), the reality is that most researchers are not or have not been classroom teachers and thus are not closely acquainted with the realities of teaching. They understand the theory and come up with suggestions that are not practically applicable in a classroom setting. One of the focuses of the design for this intervention was on how to make it practical and applicable to classrooms, not just to theory or research. Therefore, making connections to teacher observations and to the results visible in the classroom was an area of interest. Many articles have at the end almost as an afterthought a paragraph or two of how their work could be applicable to classrooms. Many times, those paragraphs are not connected to the reality of working with twenty to thirty students at a time, following mandated curriculums, and preparing students for high stakes testing. This connection to reality is integral to how this work was designed and implemented with cooperation by the schools involved. Working with them was privilege that also came with a reciprocal relationship as casual hallway discussions occurred that made other teachers in the building interested in morphology and how they could apply knowledge of morphemes in their classrooms. This is part of worldwide teaching culture; the exchange of ideas and resources is common and expected as part of building teaching relationships. Developing these working relationships additionally aided the research when COVID struck, and schools were being more

cautious as to who they were letting in the building even before the UK government closed the doors. It meant that the researcher was able to continue to work with two out of the three sites and modify timings so that the participating classes were still able to prepare students for online learning.

The next sections will address the research questions directly and then be followed with an evaluation of the intervention specifically. As demonstrated thus far there were positive effects for many of the participants on a majority of the measures. These results and how they affect the answers to the research questions and consideration to the loss of one whole school of participants will be reflected upon in this section.

7.4.1. EAL Participants

If we look at the same question for just the EAL participants, this would additionally be a cautious yes. Students demonstrated larger gains on items they were exposed to in the intervention than items they were not. Three out of four items that were tested in the same format and two out of three items that were tested differently showed improvement in comparison to three out of eight that were not taught. Even with small numbers this shows a difference that is more than just chance on the measure of roots.

With regards to the measure of prefixes and suffixes, the untaught items only showed an increase in correct answers on three out of six items. Three out of five items where the morpheme was tested in the same form as taught showed improvement. Six out of seven items where they were tested in a different form than taught showed an increase in correct answers. Potentially this could be evidence for the benefit of instruction of morphology (both morphological deconstruction and analysis skills) for EAL primary students. The fact there is evidence that students were able to identify morphemes that they had been taught in a different context, and that they know the meaning of those morphemes, together shows the beginning of background knowledge being formed. While these results are tentative, they potentially illustrate that this sort of work with primary age EAL participants could be beneficial towards building a mental schema that could later be applied to new contexts. It is unusual and interesting that the morphemes that were taught and tested differently were proportionally more successful than the taught and tested in the same form. This could indicate that the EAL participants were more sensitive to the morpheme than whole word instruction. This could be due to what phase of

reading development they were in at the time tested or that due to functioning in multiple languages they have a more finely developed skill of morphological decomposition and can access more easily morpheme knowledge than whole word knowledge. This could be an additional area for further study. Moving on now to consider the L1E participants where the results were different.

To summarize with the EAL participants there the results indicate a potential positive effect for student morpheme knowledge from the instruction given during the intervention period. This indicates that there is an opportunity to support students through interventions in morphology to understand words that are morphologically related to morphemes that are instructed.

7.4.2. L1E Participants

The final subgroup that needs to be examined for this question is the L1E participants. For this group the data seems to be inconclusive. For the measure of prefixes and suffixes on untaught morphemes four out of six items showed increases in correct responses from T2 to T3. While on items that were tested in the same form as taught only two out of five showed increases in correct responses. On items that were tested in a different format than taught four out of eight showed an increase from T2 to T3 in correct responses.

On the measure of roots the untaught items, three out of eight showed an increase in correct responses from T2 to T3. Whereas three out of four items where they were tested in the same form as taught showed increases in correct responses and, two out of three items where items were tested differently than taught in the intervention showed increases in correct responses from T2 to T3. This evidence is why the answer is not clear for the L1E group. This does not mean that there is not potential benefit, rather it is not as clear as it that which was observed for EAL participants. This leads to additional questions as to why the L1E participants did not seem to respond as strongly to the intervention. It could be that they had previous exposure to the morphemes used in the intervention, so that the exposures from the intervention itself did not increase their background knowledge. Alternatively, it could be that they were already in the process of developing a mental schema for morphemes and the intervention did not significantly increase this. That is, potentially, the EAL students were already primed by their own issues with functioning in a second language academically that they were able to embrace the concept of this being another tool which they could use when they then encountered words that they were not familiar with. Thus, the practice over the nine weeks may have helped to develop their skills such that they were able to begin to apply them when they encountered familiar

morphemes. While there has been some work on order of acquisition for vocabulary and possibly morpheme knowledge this has not been proven(Mochizuki & Aizawa, 2000; K. Nation, 2017). This could be a tentative hypothesis for why L1E participants did not show as much improvement on the measures.

7.5 Is there a difference between EAL and L1 children in the acquisition of the meanings of derivational morphemes that are taught as part of a small group intervention. Is there an effect of individual differences in gender, attendance and age in the effectiveness of teaching the meaning of derivational morphemes?

The mixed regression models and mean scores both illustrated that there was an increase in semantic morphology knowledge over the period of the intervention. The EAL participants scored higher at T1 than the L1E participants, but both showed improvement over the period from T1 to T3 and specifically from T2 to T3 when instruction was given. These observations were supported by the findings of the mixed models. Although there was an overall improvement across the three time periods, the significant effect on scores was seen in the data for the affix knowledge (WPLT) was observed in the large improvement between T1 and T3.

If the Roots measure is examined similar trends are found with an increase in scores over the time from T1 to T3 and significance not being present in the periods from T1 to T2 and T2 to T3. However, in the mixed model and the ANOVAs. When the EAL/ L1E subgroups were examined the change from T2 to T3 was significant for only the EAL participants. This could indicate that they acquired more semantic knowledge from the intervention, though it was not present in the mixed model.

When the models examined age, gender and attendance there were no significant effects for age on either of the measures. However, the measure of affix knowledge was significantly affected by attendance as should be expected. If students are present, they are going to learn more than if they are not present for the instruction as this is just common sense. What was slightly unexpected was that there was a marginally significant effect of gender, the girls surpassed the boys on the measure of affix knowledge. The results were different for the measure of roots as there were no significant effects for any of the characteristics tracked.

That there was no significance found for age could simply indicate that there was not enough variability in the participants age to demonstrate the theorized impact of age based on models of reading development such as Ehri. The average age for EAL students was 92 months and 94 months for the L1E students. The range for EAL students was from 87 to 98 months, which is less than a year difference from the eldest to the youngest. The range for the L1E participants was 83-105 months which is 22 months. This might give enough of a range for differences in development to emerge but, the small sample might not demonstrate this clearly. A much larger age range of participants would be necessary to fully test if there is a clear effect of age on the acquisition of sematic knowledge of morphemes.

Is there a relationship between reading ability and knowledge of derivational morphemes and does this relationship differ between EAL and L1 English children?Each of the participants undertook a number of measures of reading. To control for the potential effects of these scores on measures of morphological knowledge, two mixed regression models were undertaken for each of the morphological knowledge test scores (WPLT and roots).

For the measure of Affixes there were main effects of the Single Word Reading Test (SWRT) and the British Picture Vocabulary Scale (BPVS) The effects on the SWRT were more significant than the BPVS. This is potentially due to the focus on phonics and being able to pronounce words in the current English curriculum. The Key Stage 1 curriculum is focused on students leaning to identify and pronounce phonemes. This could potentially lead to students being able to speak words that they do not know the meanings of successfully. The increasingly complex series of words in the measure would lead to difficulties with the students.

The significance of the measure of affixes at T3 along with the significance of the BPVS and SWRT could contribute to the DIER model from Kim (2020) that posits that there is an interdependence of morphology, phonics along with vocabulary to reading comprehension. It could indicate that an increase in knowledge of sematic meanings of morphemes could increase vocabulary and phonological abilities with these students. If the SWRT is considered indicative of the skill of word reading which Kim (Y. S. G. Kim, 2020)posits is a component of the skills and knowledge that are involved in reading comprehension, this could indicate that the increase in SWRT scores could potentially be an indicator of an increased reading comprehension subskill that will support the overall concept of reading comprehension but may not be significant enough on its own to shift the scores on

the reading comprehension measure itself directly. It might take more than just the increases in word reading and vocabulary to create a significant change in the measure used. These could be instead interpreted as early indicators that there is some small gain of skill over the time that could later influence reading comprehension.

7.6 How do we assess the intervention?

One issue that is being grappled with in this discussion is how to assess properly the question of whether the intervention succeeded. Can it be based purely on numbers? Can it be based on the feedback given by teachers? Does it need to be a blend of both? The origin of this work came from a desire to design something practical. Something that could be easily implemented in a classroom anywhere around the world where students were learning in English, without a dependency on technology or a specific curriculum. In the researcher's previous experience there has always been a reluctance of teachers to teach something they either do not believe in, or do not have enough knowledge to be confident in teaching to students. As there has been a decline in teaching Latin or Greek in schools, the population of teachers who have a basis of knowledge in these areas has also decreased. This would mean that for teachers to be willing to use knowledge of roots with Greek and Latin etymologies, they would need some support in developing their own knowledge. Being able to package the knowledge of roots and affixes in a manner that teachers could pick it up easily and use it without a large amount of study, could reduce this reluctance.

A recent article (Porter, 2021) on how to assess interventions had five guiding questions which are: 1) Who are you trying to help? 2) What is the intervention? 3) How will you measure effects? 4) Compared to what? 5) When will you measure effects? These five questions when applied to this study uncover some interesting answers, which are summarised in the following.

Question 1 is: who are you trying to help? This answer is two-fold as the main focus is to help students with EAL or L1E backgrounds to be able to understand what they are reading more easily. The second is to develop a method which would be practical and easily applicable for teachers to instruct students in meanings of morphemes. . The priority of helping students who are most likely starting from a place where they have a smaller working vocabulary to potentially exponentially increase access to words when encountered would be very useful for both students and teachers. The work that has been completed with students being able to understand 60% of words that are

morphologically related with only knowledge of the root morpheme (White, Power, et al., 1989) would show that there would be a significant increase in access to meanings if even a basic knowledge was acquired.

Question 2 is: what is the intervention? This concept becomes somewhat unclear due to the realities of recruiting participants. While the original focus was still able to be on EAL primary students, trying to recruit a significant number who met the qualifying criteria was difficult. After emailing over 150 schools with follow up emails and phone calls, sending paper invitations through the mail to those schools as well and going to present the proposed research at professional organizations meetings in the target area. Only 6 schools responded with interest in more information or interest in participation. Of these only one was willing to allow the work to be carried out. Then it was found that they would not have enough students on the target year. However, a second school in the area was recruited by the EAL specialist at the school. The Pilot school volunteered to participate in the full study after seeing the results of the pilot. The feedback from the pilot along with the requests of the participant schools to work around their current curriculum and scheduling needs did not lead to the overly modification of the intervention, nor did it affect the goal of being able to work with the students in small groups. During this period of reflection, the researcher designed a number of extensions to the original data collection plan, as it was felt that it would add depth and that it would confirm in a more traditional manner the work that was being completed. This was the planned delayed testing measures and focus group interviews with the participants and interviews with the participants' teachers. There was also a tentative study agreed upon to have two classrooms, one acting as a control with the teacher leading the intervention and one teaching business as usual and then the intervention at a later date. Some of these plans had already gone through the ethics approval procedure at the university and had participants return signed consent forms at the time that the work was cut short by the closure of schools. These plans ultimately did not come to fruition, because of the impact of Covid-19 and they constitute a potential study for a later date. These ideas were to be utilized to add depth and more validity to the original work as it became obvious that the original plan would not recruit enough participants to have fully powered EAL and L1E cohorts or be feasible with the resources available. The extensions would have gathered a variety of data that could have been triangulated to with the originally planned data to give a better picture of the work completed. The use of these mixed methods processes would have given additional validity by confirming the measures by both teacher

observations and feedback from the participants about if they were utilizing what they had learned post intervention. These methods would be suggested should further work be carried out in this area. While quantitative data is useful and widely accepted for evidence of success or failure of work, more qualitative data might give insights into other unexpected areas of improvement that could have occurred. While it is possible to quantify confidence in a classroom based on a survey or observation, the evidence given by teachers who interact with students regularly and can observe trends or even expression and tone changes on a regular basis would give more access with less interference to a classroom and changes to a student's behaviour. This would be more organic and for many teachers feel like a more accurate interpretation of the results.

What appears to be missing from many intervention studies is the personal connection to participants, the more granular and detailed expressions of what worked and what did not. While research needs to be unbiased and attempts to be widely applicable there should also be a balance between the rigidity of science and the art of teaching. With the shifts in many educational systems to the only measures of success being numbers and test results there has been a loss of the humanity of teaching and the heart. While on a day-to-day basis teachers work to keep this in their classrooms the pressure to ensure that students meet certain benchmarks by certain dates has meant that there has been a push for more quantitative methods and the replication of these methods to only be exact and timed to the second. This results in more scepticism and a lack of respect for experience and intuition when teachers work with students.

Question 3 is: how will you measure effects? The measures were a mix of qualitative and quantitative with both bespoke and standardized measures being utilized. This mix is supported by some of the work that has been explored in the literature review (e.g., Clemens & Fuchs, 2021). The use of both gives a different level of validation as both the concept that was anticipated being affected and other standard skills involved in reading are evaluated. The use of teacher interviews and observations gives a level of depth that is not usually found in pure quantitative data. Where there could be biases in the qualitative data it is a useful manner to gather observations of where participants could be utilizing the knowledge in their regular classwork. This gives support to some additional benefits that were not planned from the intervention but could still be considered positive outcomes.

Question 4 is: compared to what? By comparing the students' performance in one time period with their own performance at a slightly later time period allowed this study to have a perfectly matched sample. It additionally meant that potentially compounding independent variables could be disregarded as all these variables would be the same for all students. Such variables include participants' not being from the same language background, or not having the same amount of experience in an English classroom setting, or motivation, reading level or even previous teachers and knowledge are. The only changes being time and the intervention, and this limiting of these variables is unique to this study and a potential avenue for further research. This is because matching control samples to test samples in education with the heterogeneity of many school populations could prove to be difficult as the world has become much more transient, and students and families immigrate and settle in new and different countries. The average classroom in England currently has a population that contains approximately a third of the students who are classified as EAL. This should and potentially does have an impact on the instruction methods used and how curricula are structured on a day-to-day level. This population has different needs to many other populations present in a classroom. Some of the work that has been done as illustrated that the levels of long-term instruction for EALs could be similar to students who come from lower SES backgrounds and thus methods that could be applicable for one could potentially transfer to another. This could additionally mean that students who are both EAL and/or are from a lower SES background could be disadvantaged in a way that is both different to these separate groups and the same. Finding methods that can reach and aid these groups will continue to be necessary as these students' progress through education and continuing conflicts in the world encourage immigration and settling in English language dominant countries.

Question 5 is: when will you measure effects? The effects were measured earlier than planned. The plan had been for students to have a week after the end of the intervention before testing started. The issues with the shorting of the intervention and not wanting to be without any data due to the closing of schools were considered and debated. Reflecting now over a year later with what eventually happened the utmost care was taken to try and safeguard the participants of the schools who were willing to allow the researcher in. For the final measures rooms were found where windows could be opened, and any materials students touched were then sanitized or left in sealed plastic folders for over two weeks. The loss of the delayed testing did indeed affect some of the possible analysis that was planned. It also affected the work with the students and the inability to gather qualitative data

from them. The use of the data that was gathered while it is not as sound as was planned it is still useful and gives evidence for potential further study.

To summarize the answers to these questions, the work completed with the intervention was moderately successful as it worked with the teachers and participants and some positive results on the final measures of reading comprehension and morphological knowledge were observed. The unmeasured successes that were observed by the participants teachers were additional positives for the evaluation of the work. The comparison of students' progress over two periods while experimental, has the potential to be refined and employed as a method for studying small populations. Overall, the intervention has the potential to be used to support primary aged students with reading comprehension and understanding of morphemes.

This combination of findings provides some support for the premise that instruction and work with morphemes and their meanings with small groups of both EAL and L1E students can support their knowledge of the taught morphemes when they are encountered in both the same form they were taught and in new contexts. The findings additionally support the idea that there was a significant improvement in reading comprehension for the EAL students over the period of the intervention. This would suggest that the intervention in some manner has the potential to improve reading comprehension for EAL students over at least a short term. The intervention also had additional anecdotal effects on the students work in their main classrooms, as they were more willing to participate and seemed to have developed more confidence. They were additionally willing to employ the skills taught in the intervention in order to understand new words when encountered in class.

7.7 If this study was to be replicated, what would need to be changed to validate the results?

The double baseline design, while theoretically possible, still needs additional validation to provide evidence of its effectiveness in use. There are multiple factors that would need to be modified if this study was to be replicated for the results to be more easily accepted as valid. The paragraphs below will consider several of them.

The first and most significant change that would be needed would be made involves the inclusion of aa specific control group that is not exposed to the instruction which was undertaken by the present study participants. Such a control group would, however, need to be tested at the same three times that

all the current participants were tested. The control group would need to be composed of a similar group of both EAL and L1 students as the participants in the intervention. While it might be difficult to build a control group with identical or highly matched language backgrounds as the participants, if it was possible, this would strengthen the use of a control more significantly.

The second change that would assist in the validation of the results would be the development of a third set of passages on grade level for the YARC. As previously, discussed the YARC was developed with two sets of passages for use as a pre and post test. An alternative measure of similar format that was tested and standardized with at least three sets of passages including both fiction and nonfiction passages could be considered, as this would reduce or even potentially eliminate the possibility of practice effects, as participants would not see the same passage more than once.

The third change that would likely benefit this study would be the use of a nonword measure in addition to or in replacement for the morphology measures. Using nonwords which contain the morphemes that were being taught and tested could help with questions about the vocabulary knowledge of the participants that was not accounted for in original design. While this was considered during the original planning of this study, the idea was dropped due to both time that would have taken to select the best measure from several that have been used in similar studies and the time it would take to administer them in addition to the measures that were already being administered. As discussed previously, the amount of time students were pulled out of their regular classes was a difficult issue to navigate. Adding an additional measure to the time they were out of class for testing likely would have added additional interruptions and more days of testing in the schools. The current set of measures took approximately 30-40 minutes per student for the measures given individually. Then the two morpheme measures that were given in groups took approximately 15-25 minutes (or a whole afternoon break period for one school). While this is a significant issue, if one could be found that could be administered in less than five minutes and in a group setting, this could be added without too much of an additional time burden on the schools.

Fourthly, more attention could be paid to the selection of morphemes both for the measures and the instruction. Focusing more intently on the occurrences and frequency of the word families could produce stronger results, as this would also benefit from likely incidental exposure to the morphemes outside of the instruction and intervention. Although it was useful in the present study, the use of lists

that had already been validated and determined could be revised to use more current work as school texts have likely evolved since the work was originally completed in the last century.

While this is not an exhaustive examination of what could be changed, should the study be replicated these three suggestions would likely most significantly impact to the questions of validity that have been raised. The use of a control group would be the most impactful, followed by different reading passages for the assessment and finally the use of a measure with nonwords to control for vocabulary knowledge.

8. Conclusions

This study set out to examine the possible connections between instruction in morphology and reading comprehension. The results indicate that this is a potential area for further work and study. The effects of instruction for EAL students in particular were significant and could be a fruitful area of future work. While the results for the L1E participants were not as conclusive, they nevertheless did indicate that there was some improvement over the period of the intervention on reading comprehension and knowledge of morphemes. This study found that there are further potential positive effects for primary aged students' reading comprehension skills from instruction of semantic meanings of derivational morphemes in small group settings. There are still questions to be examined about the utility of such instruction with larger groups or whole classes.

The findings from this study make several contributions to the current literature. First, they contribute a potential new methodology, that could play a role in future research with difficult to access populations. Secondly, it contributes a possible new measure of morphological knowledge with particular emphasis on instruction of semantic meanings of derivational morphemes. Thirdly, it proposes a method for instruction of derivational morphemes that has demonstrated potential positive effects for both EAL and L1E students in year 3 of primary school. Fourth, it adds to the growing evidence that instruction in morphology can affect reading comprehension, and this supports the theoretical work that has hypothesised that morphological awareness plays an important role in the development of reading comprehension.

The methodology used to compare a baseline period with an intervention period using the same participants is a potential area for further exploration, particularly as it allows for the study of populations that are difficult to access. With schools moving toward more proactively safeguarding their students from outside influences, gaining access to significant numbers of students is becoming more difficult. Establishing a methodology that enables researchers to utilize to the best capability smaller populations would aid early career researchers in particular, with being able to carry out work in realistic environments that are true to contexts where education occurs. To illustrate its full utility, it would need more study using a traditional control group methodology comparison to establish its reliability. This would potentially establish this method as a manner to study smaller populations that would be difficult to find adequate samples to fully power studies.

The morphology measure of root knowledge, while it was purposely made for this individual study, could be transferred to additional work in primary grade settings. It is a starting point for a measure of student knowledge as the measures that have been previously employed focus more on inflectional morphology. This area is ripe for development and this measure has potential to be further developed and expanded upon.

The findings from the intervention suggest that this method of instruction for derivational morphemes shows significant potential for students' acquisition of the meaning of taught morphemes. The results indicating that morphemes that were instructed were able to have their meaning identified over the short term would suggest that students can acquire morphological knowledge from the intervention. The examples from the interviews with teachers illustrate that the instruction in morphological decomposition has positive effects on students' ability to identify morphemes and deduce their meanings to take "educated guesses" as to the meaning of a previously unencountered word. This could have utility for students as they progress through school and encounter more morphologically complex words. This increase in encounters would theoretically necessitate an ability to identify morphemes and connect them with prior knowledge as to the meaning of the morpheme. The evidence from both the quantitative measures of morphology and the interviews with the teachers combine to support that there was some positive effect for the students from the participation in the intervention.

The results of this research support the idea that quality lexical representation of a word can aid comprehension. The improvements from the reading comprehension measures and the increases on the two measures of morphological knowledge indicate that there is a connection between this knowledge. This would support the work that has been completed that suggests that increasing skills in the area of morphological analysis could support comprehension of texts containing morphologically complex terms. While more specific work needs to be carried out in this area the results of this study indicate that there is potential for this area of work.

This section described the contributions that this study made to the overall field of morphology and reading. It described the cautious positive effects of an intervention and additional support for current theoretical work. It additionally has contributed a potential new methodology for working with small or difficult to access populations and a potential new measure of morphological knowledge. The following sections will address two areas of additional importance, implications for instructional settings and the limitations of this study. This study was carried out with the intention of being able to

contribute not only to current theoretical work with derivational morphology but, additionally aide teachers in designing instruction that can best utilize instructional time to give students the greatest possible benefit from limited instructional time that is available. The limitations will address the issues with sample size and conducting and collecting data during a pandemic. The effects of this cannot be avoided when discussing the results from this study and need to be considered thoughtfully as to potential impacts that occurred.

9. Implications for instructional settings

As far as impact on instructional settings is concerned, there are several different ways the experience and information from this work could be employed in instructional settings. First, the individual components of the intervention could be used individually to introduce and reinforce vocabulary or morphological knowledge. Secondly, the method or lesson cycle used could be applied to an area of study and target morphemes could be focused on in addition to the area of study. Thirdly, the age or year level of in which instruction could or should take place would need to be considered as there has been some discussion about when it would be most impactful, and this would need additional study. Finally, the setting for implementing the intervention is considered with relation to group size.

There were some positive effects from the intervention. This means that some of the parts could be easily adapted to regular classroom instruction or for small group instruction sessions. The word maps part of the intervention could be used with any text that is being read that contains morphologically complex words. It could be displayed on a smart board or white board with the class filling it in together with them later finding additional morphologically complex words in the text and mapping them out for individual practice. This use of word mapping is similar to that used by Baumann (2003) in his work with social studies texts. It could be used in conjunction with reading anthologies (used in many reading programs in the United States), social studies or science texts as a method for students to understand the morphologically complex vocabulary that is present in these sometimes very complex texts. The follow up activity to the word mapping could be sentence writing with the words or members of the same word family of the words that had been mapped. This would continue to build and deepen the connections with the morphemes present in the words and their meanings and potential uses for them. Word sorts could be used when a significant number of morphologically complex words had been encountered. If a lesson contains 14-16 words they could be printed on small cards and students could sort them by affixes, roots or meanings. These could be combined with words from additional lessons to give a challenge to students that would help them to remember and revise words that had been exposed earlier.

There is potential that a target morpheme could be chosen for a unit of science or another subject, and students could hunt for this morpheme as they are reading and list all of the words, they find that it occurs in, to build the connections between words that are in a similar area of study. This could potentially help to build networks of nodes of morphological knowledge in the mental lexicon and which would make it easier for students to recall where they have seen a morpheme previously and thus to be able to connect it to the knowledge they possess when they encounter unknown words that contain morphemes they do know. This would allow for instruction in morphemes and morphological knowledge without a large impact upon class time as it would be embedded in the work that is already being covered. It would be able to be differentiated for students performing on different levels. Stronger students could find and map more complex words while students who might be struggling could have some support in the word map where they are given parts of the map already filled in with the focus being on the target morpheme. Similar support could be given with sentences and sentence starters where a struggling student would just need to complete the sentence with the correct morphologically complex word and a phrase. This type of differentiation would make the work accessible to students who might not have fully transitioned to the morphological or automatic phase of reading.

The grade or year in which instruction could begin is variable, as it would need to depend upon individual student development and if they have begun to notice morphemes in words. If we look at the models of reading development, they suggest that year or grade 3 would be a likely time to begin instruction as students should be making the transition or have transitioned to the final stages of reading development. While year or grade 4 would be a time when more focus could be placed on morphological knowledge as students should have fully transitioned out of the learning to read phase and be more focused on reading to learn. This would allow students to have additional skills to draw upon while reading and encountering new words that they might not already have in their vocabulary. This addition of skills through regular practice embedded in instruction already occurs in a majority of classrooms. The spread of this to subjects other than reading would connect and reinforce the skills that are being developed. This connection could support students who might be struggling as they do not receive as much exposure to English if they speak another language at home. This support might help them to make additional

connections and develop a curiosity about words and how they are constructed, and meanings being carried by each morpheme.

Instruction on derivational morphemes should, as stated above, start in the later primary years and continue through the end of formal education. The use of derivational morphemes in morphologically complex words in academic language necessitates a solid basis of knowledge for comprehension of such terms. The earlier students can start building this knowledge the more secure a foundation they will have as they progress. Students that may start with a lower level of vocabulary can only benefit from having access to a wider range of vocabulary by using morphological knowledge to access related words. Thus, developing this knowledge by connecting it with other knowledge and basing it simpler concepts and then scaffolding it to the higher-level concepts can aide these students' success.

The intervention itself could be utilized as if for both small group and whole class instruction as it is designed. If is it as successful for whole class as it was for small group instruction has not been demonstrated by this study. However, it was designed by an experienced teacher to function in either setting so there is a strong likelihood that it would be able to be utilized for both. The small group interventions could be led by the regular classroom teacher, a school EAL specialist, or a teaching assistant. If there was a need to pull out small groups, it would be a possible intervention for them to improve reading comprehension.

The next section will focus on the limitations of this work. While there were attempts to mitigate various limitations, there are several that were not planned for and need to be taken into consideration when examining the results in particular the subgroups of EAL and L1E participants. This will be covered with more depth in the following section.

10. Limitations of findings

This section will cover the main limitations that need to be considered when considering applying the findings of this work to other areas. There are several and as they have been considered throughout this work and with particular concern with the conclusions and implications for instruction. This section will examine them and potential methods for future work so as to not be confronted by them in the future.

While there was significant thought and planning to attempt to make the results from this study as widely applicable and valid as possible there were unforeseen circumstances that have resulted in more limited findings and a need for caution when applying the results to a wider context. There are several limitations for the findings in this study. They will be addressed in detail in this section. The first is the final small sample size as it meant that the study was only fully powered when both EAL and L1E groups were combined. The second is the impact of the impending United Kingdom lockdown due to the COVID 19 pandemic to the results of the final time 3 tests. The third is the impact of the United Kingdom lockdown due to the COVID 19 pandemic on the long-term measures and the ability to gather qualitative data from the participants about their experiences.

The small sample size was both planned and unplanned. The planned aspect of the small sample size was a result of there being only one person who could both conduct the testing and implement the intervention. This in addition with the need to work with schools as partners and to be accommodating to the needs of each school for instruction made the reality of conducting the research very time limited. The unplanned aspect of the small sample size was the loss of almost half of the participants as the result of the school they attended taking precautions to limit the students' exposure to non-members of the school community in the weeks leading up to the first UK lockdown. While this is understandable, the researcher is grateful that the other participating schools were willing to allow her to come in and with extra precautions in place collect data from the students that were present. This loss of participants left the researcher with a very small but still fully powered whole group sample. However, while the whole group was balanced equally between the subgroups, the subgroups as individual samples did not meet the threshold to be fully powered. This means that while the whole group data can be trusted to accurately reflect the population, the subgroup data should be interpreted

with caution and should be seen more as a pilot for further work with larger groups. This leads to the second limitation of the possible effects of the first UK lockdown on the time 3 measures.

The time 3 measures took place over the week prior to the first UK lockdown in March of 2020. The researcher in conjunction with the advice of her supervisors shortened the intervention by one week and started testing earlier than planned. This was beneficial as the last day of testing was the day the national lockdown was announced by the Prime Minister. As has been studied there are effects on student performance due to test anxiety and stress (Putwain, 2007; Von Der Embse et al., 2013). These factors tend to inhibit memory recall of learned knowledge and increase off task behaviours during testing. There were documented instances of 'off task' behaviours during the time 3 measures. The notes from testing indicate that a majority of participants tested showed physical indications of stress during the time 3 measures as compared to none during either time 1 or time 2. As was part of the protocol for testing, breaks were allowed if students asked or appeared to be anxious. Several of the participants fidgeted more and had a noticeable change in breathing which indicated increased stress. This would potentially indicate that there was increased stress and that the measures might not be an accurate reflection of learning as students might be inhibited by anxiety.

The third limitation is the lack of delayed measures that were planned for ten weeks post intervention and follow-up focus groups. These were planned and designed to see if there were delayed effects from the intervention or long-term improvement. Additionally, the focus groups would have given participants an opportunity to reflect upon their experience and discuss if any of the skills or knowledge gained from the intervention were being applied in class. However, with the continuing governmental lockdowns these were unable to proceed.

Additionally, this study drew from a population of both EAL and L1E students who were being educated as part of the regular curriculum in England. This would mean that while there is potential for this intervention to be transferred to other English dominant countries the results could be different as there are some differences in instruction and curriculum. The age group while consistent with all participants means that the results of this work are only applicable to students of this age. While the intervention is designed so that it could be utilized in a range of years it has not been evaluated for efficacy on this range currently. While this study could be replicated at a later time with both a control group and a wider range of ages as the data and findings stand there is a limited range and tentative applicability for the results.

11. Future Directions for Research

While this study was a valiant attempt it needs to be replicated with the addition of a control group if its findings are to be verified. The work while exploratory and largely positive, needs refinement before it could be firmly established as a reasonable method for instruction in derivational morphemes. This study has established that it has potential, but it needs to be expanded upon with a wider group of participants. The replication would additionally benefit from the planned but unexecuted work with long term effects and follow-up interviews with all participating classroom teachers and participants. It is understood and accounted for that this study was affected by the pandemic and that its results need to be taken with caution.

In addition to replicating the small group interventions, it would likely be interesting to carry out the planned whole classroom exploratory work as this has the potential to be more useful for teachers as instruction for the entire class could be integrated in subjects beyond basic English Language Arts work. This integration would benefit both students and teachers as a method for supporting vocabulary acquisition and using instructional time wisely to cover more than one topic.

Further work into the effects of instruction in derivational morphemes and Greek and Latin roots could benefit both EAL and L1E primary students as it has been tentatively established that morphological knowledge contributes to reading comprehension in some manner. The question as to how that instruction should be conducted is still unclear. There are potential methods that are being developed by teachers to aid in the fulfilling of the requirements of teaching morphemes to aid spelling in England and to assist with vocabulary knowledge in the United States. These causes are supported by connecting the knowledge acquired to reading comprehension could only add to the connections being formed by students mentally and thus aiding their understanding of a wider range of words. Work directly involving teachers in designing interventions and methods could benefit over both short term and long time periods as it would ensure methods are not only based in research but also practical for the reality that is teaching in a classroom. This inclusion of experienced teachers could additionally produce opportunities for researchers who do not have the experience of working in a primary classroom setting with opportunities to experience this setting personally and then utilize this experience to conduct research that would result in practical applications to improve reading or vocabulary instruction. The inclusion would help to strengthen the applicability of research to

instruction as much research focuses on theory and is detached from the practical realities of day-today classroom instruction.

An additional area where this type of cooperation between researchers and teachers could potentially be useful would be with international schools worldwide, where there are consistent mixes of both EAL and L1E students. The EAL students in many of these schools participate in instruction in content areas in English on a regular basis as one of the goals for many parents in sending their children to these schools is for the children to be bilingual and able to attend a university in an English dominate country. The differences in exposure to English for participants would be reduced and the students theoretically would be very motivated to improve their reading comprehension and be able to expand their vocabulary. The need to access content area vocabulary in areas where there are increasing occurrences of morphologically complex vocabulary could encourage interest in participation in this type of research.

12. Appendixes

12.1. Word Part Levels Test

Test instructions

This test has 1 section: Meaning section

There are a total of 19 questions

You cannot go back and change your answers

You must choose an answer that is most likely to be correct if you don't know it

There are no penalties for wrong answers

Meaning Section

This section is about word part meanings.

For each item, a word part is presented with two example words. (The word part may also appear in other words.)

You must choose the meaning of the word part from four choices.

Here is an example.

-ed (walked; played)

(1) past

(2) not

(3) many

(4) person

The first part of this section is about prefixes (word parts that attach to the beginning of a word).

1. co- (co-worker; co-exist)

(1) person/thing

(2) direction

(3) together

(4) main

2. multi- (multinational; multicultural)

(1) opposite

(2) person/thing

(3) many

(4) small

3. pre- (pre-war; pre-school)

(1) direction

(2) bad

(3) before

(4) person/thing

4. un- (unable; unlikely)

(1) the furthest

(2) not

(3) a state of

(4) female

5. mono- (monotone; monoculture)

(1) person/thing

(2) not

(3) one

(4) into another state/place

6. dis- (disappear; disorder)

(1) not

(2) person

(3) new

(4) main

7. auto- (autobiography; autopilot)

(1) not

(2) too much

(3) direction

(4) self

8. anti- (anti-social; anti-government)

(1) toward

(2) a state of

(3) against

(4) person

9. uni- (unisex; unicycle)

(1) one

(2) person/thing

(3) not

(4) under

10. bi- (biplane; biannual)

(1) person/thing

(2) two

(3) away from

(4) in advance

11. mid- (midnight; midsummer)

(1) halfway

(2) opposite

(3) in advance

(4) person/thing

12. trans- (transform; transplant)

(1) two

(2) main

(3) into another state/place

(4) too much

13. post- (postwar; postmodern)

(1) after

(2) person

(3) again

(4) opposite

The second part of this section is about suffixes (word parts that attach to the end of a word) 14. -ible

(accessible; convertible)

(1) can be

(2) person/relating to

(3) after

(4) times

15. -less (endless; useless)

(1) before

(2) without

(3) the furthest

(4) person

16. -or (editor; governor)

(1) not

(2) person/thing

(3) number

(4) made of

17. -ist (specialist; artist)

(1) against

(2) person

(3) two (4) not

18. -able (acceptable; predictable)

(1) person (2) not

(3) can be

(4) one

19. -er (manager; leader)

(1) small

(2) person/thing

(3) not

(4) under

12.2. Test of Root Knowledge

Measure of root knowledge

Directions: Circle the best answer based on the roots underlined in the answers.

Example: Which word below do you think has something to do with **looking** at something?

a. Telegraph
b. Telephone
c. Telescope

1. Which word below do you think has something to do with **stars**?
    1. Economy
    2. Astronomy 3. Autonomy

2. Which word below do you think has something to do with **sending** something?

a. Tension
b. Erosion
c. Mission

3. Which word below do you think has something to do with the **earth**?

a. Biologist
b. Geologist
c. Zoologist

4. Which word below do you think has something to do with **light**?

a. Photograph
b. Bibliography
c. Paragraph

5. Which word below do you think has something to do with **speaking**?

a. Dictate
b. Climate
c. Migrate

6. Which word below do you think has something to do with **carrying** something?

a. Reliable

b. Probable
c. Portable

7. Which word below do you think has something to do with **measuring**?
a. Chronoscope
b. Chronograph
c. Chronometer

8. Which word below do you think has something to do with **self**?
a. Economy
b. Astronomy
c. Autonomy

9. Which word below do you think has something to do with **sending across?**
a. Intermit
b. Transmit
c. Recommit

10. Which word below do you think has something to do with **moving** something?
a. Promote
b. Profile
c. Promise

11. Which word below do you think has something to do with **breaking**?
a. Pressure
b. Juncture
c. Fracture

12. Which word below do you think has something to do with **time**?
a. Aerobic
b. Organic
c. Chronic

13. Which word below do you think has something to do with **one**?
a. Monograph
b. Paragraph
c. Autograph

14. Which word below do you think has something to do with **building** something?
a. Construct
b. Condition
c. Conclude

15. Which word below do you think has something to do with **hearing**?
a. Visible
b. Audible
c. Posible

Answer Key

1. B*
2. C
3. B*
4. A*
5. A
6. C*
7. C
8. C
9. B
10. A
11. C*
12. C
13. A*
14. A*
15. B
16. C

Denotes taught roots = *

12.3. Example word sort cards

1	2	3	
Bi	**Tri**	**Uni**	**Mono**
bicycle	triangle	unite	monotone
bifocals	trilogy	unicorn	monolingual
bimonthly	trio	unicycle	monopoly
biweekly	triplets	uniform	monorail
binary	triathlon	unify	monologue
billion	tricycle	unibrow	monotonous

12.4. Example sentences from students

Morphologically complex words are in bold with the target morpheme underlined. None have been corrected for grammar or spelling.

1. An **astronom<u>er</u>** look at the stars.
2. At **<u>mid</u>night** I had a snack.
3. At **<u>mid</u>night** I watched a meteor shower.
4. There is an **end<u>less</u>** mode when I was playing a game on my tablet.
5. The junk was **worth<u>less</u>**.
6. I am **<u>un</u>afraid** of spiders.
7. I had a **<u>dis</u>agreement** with someone I know.
8. I am a perfect **lawy<u>er</u>** and I study law.
9. I was a **teach<u>er</u>** who taught everyone in the afternoon.
10. I know what a **<u>pre</u>fix** is.

12.5. Interview schedule for teachers

1. How do you teach vocabulary (for both reading and spelling) currently?

2. Do you use affixes (prefixes and suffixes) and roots for teaching meaning currently?
3. Why or why not?
4. Have you noticed changes in the students who participated in the intervention?
5. If so,what type of changes?
6. Have there been improvements in their reading or spelling or writing? Can you be specific about changes and which students?
7. (Only for teachers at schools with both EAL and L1 participants) Were there any noticeable differences between your EAL and L1 students' responses to the intervention that you observed?
8. What is your opinion about the structure of the intervention, with students being pulled out twice a week for 10 weeks?
9. Was it disruptive?
10. Was the length fine or was it too long for the students to be missing class?
11. Would you have liked more information about the intervention or an opportunity to have observed a session?
12. If you did observe a session, what is your opinion about what you saw?
13. If you were given a program that would teach about the meanings of affixes and roots, would you be willing or interested in using it with your students?

12.6. Baseline Data from Schools

Whole Group Means and Standard Deviations for measures

	School A*		School B		School C	
	M	*SD*	*M*	*SD*	*M*	*SD*
BPVS T1	72.50	15.68	90.43	11.96	87.33	12.59
BPVS T2	81.25	15.33	98.86	8.93	96.67	12.52
SWRT T1	28.38	6.03	34.43	5.22	31.61	8.21
SWRT T2	32.63	6.99	37.14	5.55	33.89	7.98
YARC Reading Rate Ability Score T1	47.06	19.56	66.14	11.33	54.00	13.21
YARC Reading Rate Ability Score T2	57.50	15.57	67.86	6.99	60.50	12.61
YARC Accuracy Ability Score T1	41.31	5.56	49.86	6.07	45.78	7.04
YARC Accuracy Ability Score T2	42.75	6.42	46.43	3.41	43.72	6.78
YARC Comprehension Ability Score T1	51.00	7.17	50.57	8.42	51.33	4.31

YARC Comprehension Ability Score T2	45.44	7.39	53.00	6.83	52.11	6.58
WPLT T1	5.31	1.92	6.14	2.12	6.61	3.26
WPLT T2	5.00	1.32	7.00	2.45	7.11	3.72
Roots T1	6.94	1.57	7.29	2.63	7.06	2.58
Roots T2	7.00	1.26	6.86	1.07	6.94	3.35
	*School A was not able to participate in testing at T3 so its data was dropped from the final analysis					

12.7. Baseline Data by Schools and Subgroups

	School A*				School B		School C			
	EAL		L1E		EAL		EAL		L1E	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
BPVS T1	68.13	14.52	76.88	16.51	90.43	11.96	90.33	13.61	86.73	12.80
BPVS T2	76.50	12.63	86.00	17.10	98.86	8.93	97.00	11.27	96.60	13.12
SWRT T1	27.63	5.73	29.13	6.62	34.43	5.22	33.33	5.51	31.27	8.76
SWRT T2	31.88	6.08	33.38	8.16	37.14	5.55	34.33	6.35	33.80	8.46
YARC Reading Rate Ability Score T1	42.88	15.02	51.25	23.55	66.14	11.33	50.67	8.39	54.67	14.11
YARC Reading Rate Ability Score T2	55.38	12.09	59.63	19.06	67.86	6.99	60.00	11.36	60.60	13.22
YARC Accuracy Abilty Score T1	40.13	5.11	42.50	6.07	49.86	6.07	45.33	3.06	45.87	7.67
YARC Accuracy Abilty Score T2	41.88	6.77	43.63	6.39	46.43	3.41	43.33	3.79	43.80	7.33
YARC Comprehension Ability Score T1	50.25	8.36	51.75	6.25	50.57	8.42	51.33	4.16	51.33	4.48
YARC Comprehension Ability Score T2	43.13	9.69	47.75	3.33	53.00	6.83	52.67	5.69	52.00	6.92
WPLT T1	5.00	2.45	5.63	1.30	6.14	2.12	9.33	2.08	6.07	3.22
WPLT T2	5.50	1.20	4.50	1.31	7.00	2.45	9.00	4.36	6.73	3.63
Roots T1	6.88	1.46	7.00	1.77	7.29	2.63	8.67	3.06	6.73	2.46
Roots T2	6.63	1.51	7.38	0.92	6.86	1.07	7.00	2.00	6.93	3.61

*School A was not able to participate in testing at T3 so its data was dropped from the final analysis

12.8 Child Language Use Survey

Child language use PN:

1. Do you speak the same language at home and at school?

   Mark only one oval.

   - Yes
   - No

2. Do you speak the same language with your parents as with your teachers?

   Mark only one oval.

   - Yes
   - No

3. Do you speak only English with your friends?

   Mark only one oval.

   - Yes
   - No
   - A mix of English and something else

4. If you have brothers and sisters what language do you speak with them?

   Mark only one oval.

   - English
   - Something else
   - A mix of English and something else

5. What language do you prefer to speak in?

   Mark only one oval.

   - English
   - Something else
   - No preference

6. What language do you prefer to read in?

   Mark only one oval.

   - English
   - Something else
   - No preference

7. What languages do you know?

   _____

8. What percentage of the time do you read in a language other than English?

   Mark only one oval.

   0% of the time (never)

   - 1
   - 2
   - 3
   - 4
   - 5
   - 6
   - 7
   - 8
   - 9
   - 10

   100% of the time (always)

9. What percentage of the time do you speak in a language other than English?

   Mark only one oval.

   0% of the time (never)

   - 1
   - 2
   - 3
   - 4
   - 5
   - 6
   - 7
   - 8
   - 9
   - 10

   100% of the time (always)

10. Do you have classes or lessons for your other language?

    _____

12.9 Intervention General Plan

Week # Theme	**1 Numbers**	**2 Not**	**3 Science**	**4 Person Who**	**5 look/ light**	**6 Position**	**7 Communication**	**8 States**	**9 Build Break or Movement**	**10 Useful**
Session #1 Reading 1 Word Map	bi-* tri-	dis-* $+◉ un-* $+◉	Geo-* Bio-*	-er*$+ -or*$+	Scope-* photo-*	Pre- *$+ post-* mid-*$	Phone/ *phon-	-able*$◉►+ -ible*$◉+	-Rupt* -Fract*	de-*$+ co-*
Session #2 Reading 2 Word Sort	uni-* mono-*	Review: Word Sort Sentences	Geo- * Astro-*	Review: Word Sort Sentences	Review: Word Sort Sentences	Review: Word Sort Sentences	Scribe-* Graph-*	-ness*$◉+ -less*$◉+	-Struct* -port*	Review: Word Sort Sentences

Composition of Intervention Prefixes= 20 Suffixes=8	**Key for support of affixes and roots:** (*)Manyak, P. C., Baumann, J. F., & Manyak, A.-M. (2018). Morphological Analysis Instruction in the Elementary Grades: Which Morphemes to Teach and How to Teach Them. *The Reading Teacher*, 72(3), 289–300. Retrieved 3 December 2018 from https://doi.org/10.1002/trtr.1713 ($) Marzanno, R. J., & Sims, J. A. (2011). A Six-Step Process for Vocabulary Instruction. In *Vocabulary for the Common Core*. (+)White, T. G., Sowell, J., & Yanagihara, A. (1989). *Teaching Elementary Students to Use Word-Part Clues*.

	Source: The Reading Teacher (Vol. 42). Retrieved 27 November 2018 from https://www.jstor.org/stable/pdf/20200115.pdf?refreqid=excelsior%3Ae566ff47d443c79e92e4f45f05d6f070 (◉) Mentioned in the UK Curriculum (►) Mentioned in CCSS

13. References

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

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the Simple View of Reading Include a Fluency Component? *Reading and Writing*, *19*(9), 933–958. https://doi.org/10.1007/s11145-006-9024-z

Anglin, J. M. . (1993). Vocabulary Development: A Morphological Analysis. *Monographs of the Society for Research in Child Development*, *58*(10), i–186. https://doi.org/10.2307/1166112

Apel, K., & Diehm, E. (2014). Morphological Awareness Intervention With Kindergarteners and First and Second Grade Students From Low SES Homes: A Small Efficacy Study. *Journal of Learning Disabilities*, *47*(1), 65–75. https://doi.org/10.1177/0022219413509964

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, *6*(4), 253–279. https://doi.org/10.1093/ijl/6.4.253

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., Kame'enui, E. J., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal*, *40*(2), 447–494. https://doi.org/10.3102/00028312040002447

Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Edward, J. K., & Olejnik, S. (2002). Teaching Morphemic and Contextual Analysis to Fifth-Grade Students. *Reading Research Quarterly*, *37*(2), 150–176. internal-pdf://200.178.208.122/748155.pdf

Baumann, J. F., & Kameenui, E. J. (2004). *Vocabulary Instruction: Research to Practice*. Guilford Press. https://books.google.co.uk/books?id=T3U3NKZi2poC

Beck, I. L., & Mckeown, M. G. (1983). Learning Words Well: A Program to Enhance Vocabulary and Comprehension. *The Reading Teacher*, *36*(7), 622–625. internal-pdf://133.126.197.239/20198294 (1).pdf

Beck, I. L., & McKeown, M. G. (2007). Increasing Young Low-Income Children's Oral Vocabulary Repertoires through Rich and Focused Instruction. *The Elementary School Journal*, *107*(3), 251–271. https://doi.org/10.1086/511706

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing Words to Life, Second Edition: Robust Vocabulary Instruction* (2nd;2;). Guilford Publications Inc. M.U.A.

Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, *39*(2), 141–163. https://doi.org/10.1007/s10936-009-9130-6

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62. https://doi.org/10.1037/0022-0663.98.1.44

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498–520. https://doi.org/10.1037/0022-0663.93.3.498

Blachowicz, C. L. Z., Fisher, P. J. L., Ogle, D., & Watts-Taffe, S. (2006). Theory and Research into Practice: Vocabulary: Questions from the Classroom. *Reading Research Quarterly*, *41*(4), 524–539. http://www.jstor.org/stable/4151816

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills. *Review of Educational Research*, *80*(2), 144–179. https://doi.org/10.3102/0034654309359353

Buac, M., Gross, M., & Kaushanskaya, M. (2014). The Role of Primary Caregiver Vocabulary Knowledge in the Development of Bilingual Children's Vocabulary Skills. *Journal of Speech, Language, and Hearing Research*, *57*, 1804–1816. https://doi.org/10.1044/2014_JSLHR-L-13-0055

Cangelosi, M., Barichello, C., Dijkstra, T., & Palladino, P. (2023). How SES May Affect Reading Comprehension and Vocabulary in Language Minority Bilingual and Monolingual Children. *Reading and Writing Quarterly*. https://doi.org/10.1080/10573569.2023.2181246

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, *9*(03), 247–266. https://doi.org/10.1017/S0142716400007839

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, *12*(3), 169–190. https://doi.org/10.1023/a:1008131926604

Carlisle, J. F., & Fleming, J. (2003). Lexical Processing of Morphologically Complex Words in the Elementary Years. *Scientific Studies of Reading*, *7*(3), 239–253. https://doi.org/10.1207/S1532799XSSR0703_3

Carlisle, J. F., McBride-Chang, C., Nagy, W., & Nunes, T. (2010). Effects of Instruction in Morphological Awareness on Literacy Achievement: An Integrative Review. *Reading Research Quarterly*, *45*(4), 464–487. https://doi.org/10.1598/RRQ.45.4.5

Carver, R. P. (1994). PERCENTAGE OF UNKNOWN VOCABULARY WORDS IN TEXT AS A FUNCTION OF THE RELATIVE DIFFICULTY OF THE TEXT - IMPLICATIONS FOR INSTRUCTION. *JOURNAL OF READING BEHAVIOR*, *26*(4), 413–437. https://doi.org/https://doi.org/10.1080/10862969409547861

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Consortium, L. and R. R. (2015). Learning to Read: Should We Keep Things Simple? *Reading Research Quarterly*, *50*(2), 151–169. https://doi.org/https://doi.org/10.1002/rrq.99

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*(2), 213–238. https://doi.org/10.2307/3587951

Crosson, A. C., & McKeown, M. G. (2016). Middle School Learners' Use of Latin Roots to Infer the Meaning of Unfamiliar Words. *Cognition and Instruction*, *34*(2), 148–171. https://doi.org/10.1080/07370008.2016.1145121

Crosson, A. C., McKeown, M. G., Moore, D. W., & Ye, F. (2018). Extending the bounds of morphology instruction: teaching Latin roots facilitates academic word learning for English Learner adolescents. *Reading and Writing*, *32*(3), 1–39. https://doi.org/10.1007/s11145-018-9885-y

Cummins, J. (2001). *Bilingual Children's Mother Tongue: Why Is It Important for Education?* http://www.iteachilearn.com/cummins/mother.htm

Deacon, H., & Cain, K. (2011). What we have learned from 'learning to read in more than one language'. *Journal of Research in Reading*, *34*(1), 1–5. https://doi.org/10.1111/j.1467-9817.2010.01487.x

Deacon, S. H., Campbell, E., Tamminga, M., & Kirby, J. (2010). Seeing the harm in harmed and harmful: Morphological processing by children in Grades 4, 6, and 8. *Applied Psycholinguistics*, *31*(4), 759–775. https://doi.org/10.1017/S0142716410000238

Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The Relation Between Morphological Awareness and Reading Comprehension: Evidence From Mediation and Longitudinal Models. *Scientific Studies of Reading*, *18*(6), 432–451. https://doi.org/10.1080/10888438.2014.926907

Deacon, S. H., Tong, X., & Francis, K. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension. *Journal of Research in Reading*, *40*(1), 1–16. https://doi.org/10.1111/1467-9817.12056

Department for Education. (2015). *Reading: the next steps* (Issue March).

Department for Education (2017), School census autumn 2017 to summer 2018: guide for schools and LAs.

Dewaele, J. M. (2018). Why the Dichotomy 'L1 Versus LX User' is Better than 'Native Versus Non-native Speaker'. *Applied Linguistics*, *39*(2), 236–240. https://doi.org/10.1093/applin/amw055

Duff, D., Bruce Tomblin, J., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a matthew effect. *Journal of Speech, Language, and Hearing Research*, *58*(3), 853–864. https://doi.org/10.1044/2015_JSLHR-L-13-0310

Duff, D., & Brydon, M. (2020). Estimates of individual differences in vocabulary size in English: how many words are needed to 'close the vocabulary gap'? *Journal of Research in Reading*, *43*(4), 454–481. https://doi.org/https://doi.org/10.1111/1467-9817.12322

Duncan, L. G., & Seymour, P. H. K. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, *91*(2), 145–166. https://doi.org/https://doi.org/10.1348/000712600161736

Ehri, L. C., & McCormick, S. (1998). PHASES OF WORD LEARNING: IMPLICATIONS FOR INSTRUCTION WITH DELAYED AND DISABLED READERS. *Reading & Writing Quarterly*, *14*(2), 135–163. https://doi.org/10.1080/1057356980140202

Ehri, L. C., & Wilce, L. S. (1987). Cipher Versus Cue Reading: An Experiment in Decoding Acquisition. *Journal of Educational Psychology*, *79*(1), 3–13. https://doi.org/10.1037/0022-0663.79.1.3

Ellis, E. (1992). *LINCS : a starter strategy for vocabulary learning* (1st ed.).

Fan, M. (2000). How Big Is the Gap and How To Narrow It? an Investigation Into the Active and Passive Vocabulary Knowledge of L2 Learners. *RELC Journal*, *31*(2), 105–119. https://doi.org/10.1177/003368820003100205

Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2006). Understanding the Language Demands of Schooling: Nouns in Academic Registers. *Journal of Literacy Research*, *38*(3), 247–273. https://journals.sagepub.com/doi/pdf/10.1207/s15548430jlr3803_1

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*(2), 234–248. https://doi.org/10.1111/desc.12019

Galloway, E. P., & Lesaux, N. K. (2015). Reading Comprehension Skill Development and Instruction for Adolescent English Language Learners : A Focus on Academic Vocabulary Instruction. In K. L. Santi & D. K. Reed (Eds.), *Improving reading comprehension of middle and high school students* (2015th ed., Vol. 10, pp. 153–178). Springer International Publishing. https://doi.org/10.1007/978-3-319-14735-2

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Read Writ*, *25*, 1819–1845. https://doi.org/10.1007/s11145-011-9333-8

Gilbert, J. K., Goodwin, A. P., Compton, D. L., & Kearns, D. M. (2014). Multisyllabic Word Reading as a Moderator of Morphological Awareness and Reading Comprehension. *Journal of Learning Disabilities*, *47*(1), 34–43. https://doi.org/10.1177/0022219413509966

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Goodwin, A. P. (2016). Effectiveness of word solving: Integrating morphological problem-solving within comprehension instruction for middle school students. *Reading and Writing*, *29*(1), 91–116. https://doi.org/10.1007/s11145-015-9581-0

Goodwin, A. P., & Ahn, S. (2013). A Meta-Analysis of Morphological Interventions in English: Effects on Literacy Outcomes for School-Age Children. *Scientific Studies of Reading*, *17*(4), 257–285. https://doi.org/10.1080/10888438.2012.689791

Goodwin, A. P., Gilbert, J. K., & Cho, S. J. (2013a). Morphological contributions to adolescent word reading: An item response approach. *Reading Research Quarterly*, *48*(1), 39–60. https://doi.org/10.1002/rrq.037

Goodwin, A. P., Gilbert, J. K., & Cho, S.-J. (2013b). Morphological Contributions to Adolescent Word Reading: An Item Response Approach. *Reading Research Quarterly*, *48*(1), 39–60. https://doi.org/10.1002/rrq.037

Goodwin, A. P., Gilbert, J. K., Cho, S.-J. J., & Kearns, D. M. (2014). Probing Lexical Representations: Simultaneous Modeling of Word and Reader Contributions to Multidimensional Lexical Representations. *Journal of Educational Psychology*, *106*(2), 448–468. https://doi.org/10.1037/a0034754

Goodwin, A. P., Huggins, A. C., Carlo, M. S., August, D., & Calderon, M. (2013). Minding morphology: How morphological awareness relates to reading for English language learners. *Reading and Writing*, *26*(9), 1387–1415. https://doi.org/10.1007/s11145-012-9412-5

Goodwin, A. P., Petscher, Y., Carlisle, J. F., & Mitchell, A. M. (2017). Exploring the dimensionality of morphological knowledge for adolescent readers. *Journal of Research in Reading*, *40*(1), 91–117. https://doi.org/10.1111/1467-9817.12064

Green, T. M. (2014). *The Greek & Latin Roots of English*. Rowman & Littlefield Publishers.

Griffin, G., & Harley, T. A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics*, *17*(4), 443–460. https://doi.org/10.1017/s0142716400008195

Halliday, M. A. K. (1993). Some grammatical problems in scientific English. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing Science: Literacy And Discursive Power* (pp. 69–85). University of Pittsburgh Press. https://doi.org/10.4324/9780203209936

Hammond, L. (2015). Early childhood educators' perceived and actual metalinguistic knowledge, beliefs and enacted practice about teaching early reading. *Australian Journal of Learning Difficulties*, *20*(2), 113–128. https://doi.org/10.1080/19404158.2015.1023208

Harris, M. L., Schumaker, J. B., & Deshler, D. D. (2011). The effects of strategic Morphologic Analysis Instruction on the Vocabulary performance of Secondary students with and without disabilities. In *Source: Learning Disability Quarterly* (Vol. 34, Issue 1). https://www.jstor.org/stable/23053294

Hart, B., & Risley, T. R. (1989). The Longitudinal Study of Interactive Systems. *Education and Treatment of Children*, *12*(4), 347–358.

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., Smith, G., Bowyer-Crane, C., & Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*(8), 820–827. https://doi.org/10.1111/j.1469-7610.2005.01559.x

Hedrick, W. B., Harmon, J. M., & Linerode, P. M. (2004). Teachers' Beliefs and Practices of Vocabulary Instruction With Social Studies Textbooks in Grades 4-8. Reading Horizons: A. In *Journal of Literacy and Language Arts* (Issue 2).

Hiebert, E. H., Goodwin, A. P., & Cervetti, G. N. (2018). Core Vocabulary: Its Morphological Content and Presence in Exemplar Texts. *Reading Research Quarterly*, *53*(1), 29–49. https://doi.org/10.1002/rrq.183

Hoff, E. (2003). The Specificity of Environmental Influence : Socioeconomic Status Affects Early Vocabulary Development via Maternal Speech. *Child Development*, *74*(5), 1368–1378. https://doi.org/https://doi.org/10.1111/1467-8624.00612

Hogan, T. P., Adlof, S. M., Alonzo, C. N., Hogan, T. P., Adlof, S. M., & Alonzo, C. N. (2016). *On the importance of listening comprehension On the importance of listening comprehension*. *9507*(May), 199–207. https://doi.org/10.3109/17549507.2014.904441.On

Hoover, W. A., & Gough, P. B. (1990). The Simple View of Reading. *Reading and Writing*, *2*, 127–160. https://link.springer.com/content/pdf/10.1007%2FBF00401799.pdf

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403–430. https://doi.org/http://hdl.handle.net/10125/66973

Jared, D., & Seidenberg, M. S. (1990). Naming Multisyllabic Words. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 92–105. https://doi.org/10.1037/0096-1523.16.1.92

Jarmulowicz, L., & Taran, V. L. (2007). Exploration of lexical-semantic factors affecting stress production in derived words. *Language, Speech, and Hearing Services in Schools*, *38*(4), 378–389. https://doi.org/10.1044/0161-1461(2007/039)

Jeon, E. H. (2011). Contribution of Morphological Awareness to Second-Language Reading Comprehension. *The Modern Language Journal*, *95*(2), 217–235. https://doi.org/10.1111/j.1540-4781.2011.01179.x

Kearns, D. M., & Hiebert, E. H. (2021). The Word Complexity of Between First and Third Grade in Widely Used Curricula. *Reading Research Quarterly*, *0*(0), 1–31. https://doi.org/10.1002/rrq.429

Kieffer, M. J., & Box, C. D. (2013). Derivational morphological awareness, academic vocabulary, and reading comprehension in linguistically diverse sixth graders. *Learning and Individual Differences*, *24*, 168–175. https://doi.org/10.1016/J.LINDIF.2012.12.017

Kieffer, M. J., & Lesaux, N. K. (2010). Morphing Into Adolescents: Active Word Learning for English-Language Learners and Their Classmates in Middle School. *Journal of Adolescent & Adult Literacy*, *54*(1), 47–56. https://doi.org/10.1598/JAAL.54.1.5

Kieffer, M. J., & Lesaux, N. K. (2012a). Direct and Indirect Roles of Morphological Awareness in the English Reading Comprehension of Native English, Spanish, Filipino, and Vietnamese Speakers. *Language Learning*, *62*(4), 1170–1204. https://doi.org/10.1111/j.1467-9922.2012.00722.x

Kieffer, M. J., & Lesaux, N. K. (2012b). Effects of Academic Language Instruction on Relational and Syntactic Aspects of Morphological Awareness for Sixth Graders from Linguistically Diverse Backgrounds. *The Elementary School Journal*, *112*(3), 519–545. https://doi.org/10.1086/663299

Kieffer, M. J., & Lesaux, N. K. (2012c). *Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade*. *25*(2), 347–373. https://doi.org/10.1007/s11145-010-9272-9

Kim, Y. S. G. (2020). Toward Integrative Reading Science: The Direct and Indirect Effects Model of Reading. *Journal of Learning Disabilities*, *53*(6), 469–491. https://doi.org/10.1177/0022219420908239

Kim, Y. S. G. (2023). Executive Functions and Morphological Awareness Explain the Shared Variance between Word Reading and Listening Comprehension. *Scientific Studies of Reading*. https://doi.org/10.1080/10888438.2023.2195112

Kim, Y.-S. (2023). Simplicity Meets Complexity: Expanding the Simple View of Reading With the Direct and Indirect Effects Model of Reading (DIER). In *Handbook on the Science of Early Literacy* (pp. 9–22). https://www.researchgate.net/publication/369528411

Kim, Y.-S. G. (2020). Hierarchical and Dynamic Relations of Language and Cognitive Skills to Reading Comprehension: Testing the Direct and Indirect Effects Model of Reading (DIER). *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000407.supp

Kintsch, W., & Intsch, W. K. (1988). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. In *Psychological Review* (Vol. 95, Issue 2).

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Services in Schools*, *40*(3), 341–351. https://doi.org/10.1044/0161-1461(2008/08-0009)

Kuo, L., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist*, *41*(3), 161–180. https://doi.org/10.1207/s15326985ep4103_3

Laberge, D., & Samuels, S. J. (1974). Toward a Theory of Automatic Information Processing in Reading'. In *COGNITIVE PSYCHOLOGY* (Vol. 6).

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). *Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension*. *22*(1). http://nflrc.hawaii.edu/rfl

Lawrie, B., Dodd, B., & Nancollis, A. (2005). Phonological awareness Intervention and the Acquisition of Literacy Skills in Children From Deprived Social backgrounds. *Language Speech and Hearing Services in Schools*, 325–336.

Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven Profiles: Language Minority Learners' Word Reading, Vocabulary, and Reading Comprehension Skills. *Journal of Applied Developmental Psychology*, *31*(6), 475–483. https://doi.org/10.1016/J.APPDEV.2010.09.004

Lesaux, N. K., & Kieffer, M. J. (2012). Development of Morphological Awareness and Vocabulary Knowledge in Spanish-Speaking Language Minority Learners: A Parallel Process Latent Growth Curve Model. *Applied Psycholinguistics*, *33*(1), 23–54. https://doi.org/https://doi.org/10.1017/S0142716411000099

Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools. *Reading Research Quarterly*, *45*(2), 196–228. https://doi.org/10.1598/RRQ.45.2.3

Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2020). How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development. *Journal of Research in Reading*. https://doi.org/10.1111/1467-9817.12313

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of Experimental Child Psychology*, *160*, 1–20. https://doi.org/10.1016/j.jecp.2017.02.015

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2018). Inferring Meaning From Meaningful Parts: The Contributions of Morphological Skills to the Development of Children's Reading Comprehension. *Reading Research Quarterly*, *54*(1), 63–80. https://doi.org/10.1002/rrq.219

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the Simple View of Reading With Elementary School Children: Still Simple After All These Years. *Remedial and Special Education*, *39*(5), 260–273. https://doi.org/10.1177/0741932518764833

Lubliner, S., & Hiebert, E. H. (2011). An Analysis of English-Spanish Cognates as a Source of General Academic Language. *Bilingual Research Journal*, *34*(1), 76–93. https://doi.org/10.1080/15235882.2011.568589

Manyak, P. C., Baumann, J. F., & Manyak, A.-M. (2018). Morphological Analysis Instruction in the Elementary Grades: Which Morphemes to Teach and How to Teach Them. *The Reading Teacher*, 72(3), 289–300. Retrieved 3 December 2018 from https://doi.org/10.1002/trtr.1713

Marzanno, R. J., & Sims, J. A. (2011). A Six-Step Process for Vocabulary Instruction. In *Vocabulary for the Common Core*.

McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly*, *29*(4), 520–530. https://doi.org/10.1016/J.ECRESQ.2014.06.002

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, *159*(3810), 56 LP – 63. https://doi.org/10.1126/science.159.3810.56

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*(1), 134–147. https://doi.org/10.1037/0022-0663.98.1.134

Nagy, W. E., Anderson, R. C., Nagy, W. E., & Anderson, R. C. (1984). How Many Words Are There in Printed School English? *Reading Research Quarterly*, *19*(3), 304–330. https://doi.org/10.2307/747823

Nagy, W. E., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological Knowledge and Literacy Acquisition. *Journal of Learning Disabilities*, *47*(1), 3–12. https://doi.org/10.1177/0022219413509967

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. internal-pdf://127.225.26.4/1022_001.pdf

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2009). *Common-Core State Standards Initiative Standards-Setting Criteria*. http://www.corestandards.org/

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *English Language Arts Standards | Common Core State Standards Initiative*. National Governors Association Center for Best Practices, Council of Chief State School Officers. http://www.corestandards.org/ELA-Literacy/

National Reading Panel (U.S.) & National Institute of Child Health and Human Development (U.S.). (2000). *Report of the National Reading Panel: Teaching children to read : an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction : reports of the subgroups*.

Newton, J. (2018). Teachers as learners: The impact of teachers' morphological awareness on vocabulary instruction. *Education Sciences*, *8*(4). https://doi.org/10.3390/educsci8040161

Nunes, T., & Bryant, P. (2006). *Improving Literacy by Teaching Morphemes* (T. Nunes, P. Bryant, U. Pretzlik, & J. Hurry, Eds.). Routledge.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, *18*(4), 443–468. https://doi.org/10.1080/01690960344000008

O'Connor, M., Geva, E., & Koh, P. W. (2019). Examining Reading Comprehension Profiles of Grade 5 Monolinguals and English Language Learners Through the Lexical Quality Hypothesis Lens. *Journal of Learning Disabilities*, *52*(3), 232–246. https://doi.org/10.1177/0022219418815646

Ownby, K. L. (2017). *Teacher Perceptions of Teaching Morphology and Greek and Latin Roots to Primary Grade Learners of English as a Second Language*. University of Glasgow.

Ozturk, M. (2015). Vocabulary growth of the advanced EFL learner. *Language Learning Journal*, *43*(1), 94–109. https://doi.org/10.1080/09571736.2012.708053

Panel, N. R. (2000). *Report of the National Reading Panel: Teaching Children to Read*. National Institute of Child Health and Human Development. https://www.ncbi.nlm.nih.gov/pubmed/

Perfetti, C. A. (1999). Comprehending Written Language: a blueprint of the reader. In C. M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 167–208). Oxford University Press.

Perfetti, C. A. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The Lexical Quality Hypothesis. In L. T. Verhoeven, C. Elbro, P. (Pieter) Reitsma, P. Reitsma, L. (Ludo) Verhoeven, L. Verhoven, & I. Ebrary (Eds.), *Precursors of functional literacy [electronic resource]* (pp. 189–213). J. Benjamins Pub.

Perfetti, C. A., & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual Differences in Reading Development: A Review of 25 Years of Empirical Research on Matthew Effects in Reading. *Review of Educational Research*, *84*(2), 203–244. https://doi.org/https://doi.org/10.3102/0034654313509492

Pittman, R. T., Lindner, A. L., Zhang, S., Binks-Cantrell, E., & Malatesha Joshi, R. (2022). What do teacher educators know about English spelling? *Reading and Writing*, *35*(9), 2049–2074. https://doi.org/10.1007/s11145-022-10326-x

Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish-English bilingual children in elementary school. *Reading and Writing*, *25*(7), 1635–1664. https://doi.org/10.1007/s11145-011-9336-5

Reed, D. K. (2008). A Synthesis of Morphology Interventions and Effects on Reading Outcomes for Students in Grades K–12. *Learning Disabilities Research & Practice*, *23*(1), 36–49. https://doi.org/10.1111/j.1540-5826.2007.00261.x

Reichle, E. D., & Perfetti, C. A. (2003). Scientific Studies of Reading Morphology in Word Identification : A Word- Experience Model That Accounts for Morpheme Frequency Effects. *Scientific Studies of Reading*, *7*(3), 219–237. https://doi.org/10.1207/S1532799XSSR0703

Sangster, P., Anderson, C., & O'Hara, P. (2013). Perceived and actual levels of knowledge about language amongst primary and secondary student teachers: Do they know what they think they know? *Language Awareness*, *22*(4), 293–319. https://doi.org/10.1080/09658416.2012.722643

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. In *Language Teaching Research* (Vol. 12, Issue 3). https://doi.org/10.1177/1362168808089921

Schmitt, N. (2014). Size and Depth of Vocabulary Knowledge: What the Research Shows. *Language Learning*, *64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., Jiang, X., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *Modern Language Journal*, *95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in Language Teaching* (2nd ed.). Cambridge University Press. https://doi.org/DOI: 10.1017/9781108569057

Schreuder, R., & Baayen, R. H. (1995). Modeling Morphological Processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Lawrence Erlbaum Associates.

Seymour, P. H. K., & Duncan, L. G. (2001). Learning to Read in English. *Psycology: The Journal of the Hellenic Psycological Society*, *8*(3), 281–299. https://doi.org/https://doi.org/10.12681/psy_hps.24122

Shanahan, T., & August, D. (2008). *Developing Reading and Writing in Second-Language Learners: Lessons from the Report of the National Literacy Panel on Language-Minority Children and Youth* (D. August & T. Shanahan, Eds.). Routledge.

Snowling, M., Stothard, S., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K., & Hulme, C. (2009). YARC York Assessment of Reading for Comprehension Passage Reading. In *GL Publishers. (2009)*. GL Publishers. http://discovery.ucl.ac.uk/1328127/

Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early First-Language Reading and Spelling Skills Predict Later Second-Language Reading and Spelling Skills. *Journal of Educational Psychology*, *100*(1), 162–174. https://doi.org/10.1037/0022-0663.100.1.162

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds. *Child Development*, *90*(4), 1303–1318. https://doi.org/10.1111/cdev.13072

Stanovich, K. E. (1986a). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Reading Research Quarterly*, *21*(4), 360–407. http://www.jstor.org.libproxy.york.ac.uk/stable/747612

Stanovich, K. E. (1986b). Matthew Effects in Reading : Some Consequences of Individual Differences in the Acquisition of Literacy Author (s): Keith E . Stanovich Published by : International Literacy Association and Wiley Stable URL : https://www.jstor.org/stable/747612 REFERENC. *International Literacy Association*, *21*(4), 360–407.

Storch, S. A., & Whitehurst, G. J. (2002). Oral Language and Code-Related Precursors to Reading: Evidence From a Longitudinal Structural Model. *Developmental Psychology*, *38*(6), 934–947. https://doi.org/10.1037/0012-1649.38.6.934

Strand, S., Malmberg, L., & Hall, J. (2015). *English as an Additional Language (EAL) and educational achievement in England: An analysis of the National Pupil Database* (Issue January).

Swart, N. M., Muijselaar, M. M. L. L., Steenbeek-Planting, E. G., Droop, M., de Jong, P. F., & Verhoeven, L. (2017). Differential lexical predictors of reading comprehension in fourth graders. *Reading & Writing*, *30*(3), 489–507. https://doi.org/10.1007/s11145-016-9686-0

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(4), 745–765. https://doi.org/10.1080/02724980343000477

Tyler, A., & Nagy, William. (1989). The Acquisition of English Derivational Morphology - ProQuest. *Journal of Memory and Language ;*, *28*(6), 649–667. internal-pdf://217.232.68.160/tyler-nagy- acquisitionof English derivational.pdf

Uhry, J. K., & Ehri, L. C. (1999). Ease of Segmenting Two- and Three-Phoneme Words in Kindergarten: Rime Cohesion or Vowel Salience? In *Journal of Educational Psychology* (Vol. 91, Issue 4, pp. 594–603). https://doi.org/10.1037/0022-0663.91.4.594

Van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic Knowledge, Processing Speed, and Metacognitive Knowledge in First- and Second-Language Reading Comprehension: A Componential Analysis. *Journal of Educational Psychology*, *96*(1), 19–30. https://doi.org/10.1037/0022-0663.96.1.19

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*(2), 217–234. https://doi.org/https://doi.org/10.1017/S0142716401002041

von Koss Torkildsen, J., Bratlie, S. S., Kristensen, J. K., Gustafsson, J. E., Lyster, S. A. H., Snow, C., Hulme, C., Mononen, R. M., Næss, K. A. B., López-Pedersen, A., Wie, O. B., & Hagtvet, B. (2021). App-Based Morphological Training Produces Lasting Effects on Word Knowledge in Primary School Children: A Randomized Controlled Trial. *Journal of Educational Psychology*, *114*(4), 833–854. https://doi.org/10.1037/edu0000688

Waring, R. (1997). A Study of Receptive and Productive Vocabulary Learnring from Word Cards. *Studies in Foreign Languages and Literature*, *21*(1), 94–114.

Webb, S. (2008). RECEPTIVE AND PRODUCTIVE VOCABULARY SIZES OF L2 LEARNERS. *Studies in Second Language Acquisition*, *30*(1), 79–95. https://doi.org/10.1017/S0272263108080042

Webb, S., & Rodgers, M. P. H. H. (2009). The Lexical Coverage of Movies. *Applied Linguistics*, *30*(3), 407–427. https://doi.org/10.1093/applin/amp010

White, T. G., Power, M. A., & White, S. (1989). Morphological analysis : Implications for teaching and understanding. *Reading Research Quarterly*, *24*(3), 283–304. https://doi.org/10.2307/747771

White, T. G., Sowell, J., & Yanagihara, A. (1989). Teaching Elementary Students to Use Word-Part Clues. *Source: The Reading Teacher*, *42*(4), 302–308. https://www.jstor.org/stable/pdf/20200115.pdf?refreqid=excelsior%3Ae566ff47d443c79e92e4f45f05d6f070

Wysocki, K., & Jenkins, J. R. (1987). Deriving Word Meanings through Morphological Generalization. In *Source: Reading Research Quarterly* (Vol. 22, Issue 1). Winter. https://doi.org/10.2307/747721

Zhang, J., & Shulley, L. (2017). Poor comprehenders in English-only and English language learners: influence of morphological analysis during incidental word learning. *Journal of Research in Reading*, *40*(2), 169–183. https://doi.org/10.1111/1467-9817.12100

Zucker, T. A., Cabell, S. Q., & Pico, D. L. (2021). Going Nuts for Words: Recommendations for Teaching Young Students Academic Vocabulary. *Reading Teacher*, *74*(5), 581–594. https://doi.org/10.1002/trtr.1967